GENERATING MULTI-STEP CHEMICAL REACTION PATHWAYS WITH BLACK-BOX OPTIMIZATION

Danny Reidenbach[1], Connor W. Coley[2], Kevin Yang[1]
[1]University of California Berkeley, [2]Massachusetts Institute of Technology
dreidenbach@berkeley.edu

ABSTRACT

The practical usability of de novo small molecule generation depends heavily on the synthesizability of generated molecules. We propose BBO-SYN, a generative framework based on black-box optimization (BBO), which predicts diverse molecules with desired properties together with corresponding synthesis pathways. Given an input molecule A, BBO-SYN employs a state-of-the-art BBO method operating on a latent space of molecules to find a reaction partner B, which maximizes the property score of the reaction product C, as determined by a pre-trained template-free reaction predictor. This single-step reaction (A+B$\rightarrow$C) forms the basis for an optimization loop, resulting in a synthesis tree yielding products with high property scores. Empirically, the sampling and search strategy of BBO-SYN outperforms comparable baselines on four synthesis-aware optimization tasks (QED, DRD2, GSK3β, and JNK3), increasing product diversity by 37% and mean property score by 25% on our hardest JNK3 task.

1 INTRODUCTION

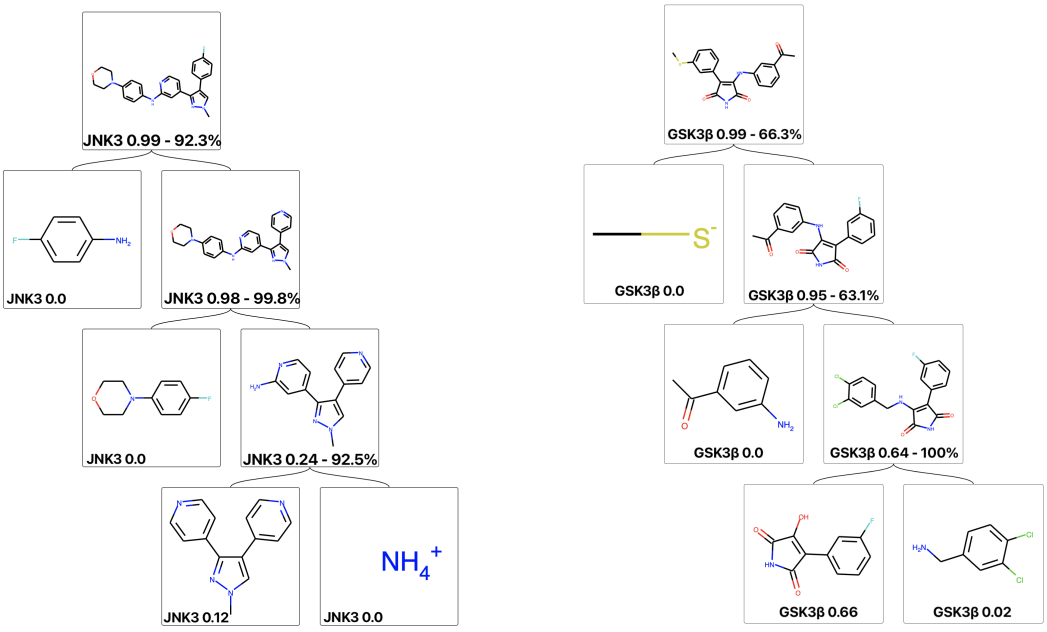
Strong automated methods for molecular design have the potential to greatly accelerate early-stage drug discovery and molecular optimization. However, while several current strategies can accelerate the filtering of 10^{60} possible drug-like molecules (Reymond & Awale, 2012) or generate novel molecules with desired properties (Kim et al., 2020b; Engkvist et al., 2021), many methods may overlook molecular synthesizability constraints, resulting in output molecules which are challenging or impossible to synthesize in practice. Despite recent work on automated retrosynthetic planners (Law et al., 2009; Segler et al., 2018a; Coley et al., 2018a; 2019a), finding viable and economically feasible synthesis pathways is still a labor- and time-intensive process.[1]

As many de novo generation methods can optimize for arbitrary properties given a scorer, some works have designed rule- or model-based heuristic synthesizability scores to guide optimization towards synthesizable molecules (Ertl & Schuffenhauer, 2009; Coley et al., 2018b; Segler et al., 2018a). While such approaches are plausible in theory, generative models often exploit these heuristics in practice. Additionally, heuristics only address half the problem: even given a perfect heuristic for synthesizability, the corresponding chemical reaction steps would still be unknown.

In this work, we design a model for the task of *synthesizability-constrained molecular design*, which we define as generating not only synthesizable molecules optimized for desired chemical properties but also corresponding reaction pathways for actually creating those molecules (Gottipati et al., 2020). In doing so, we can significantly reduce the difficulty of physically synthesizing the predicted molecules in practice.

We propose BBO-SYN, which leverages black-box optimization (BBO) to generate accurate synthesis trees with final products possessing high desired property scores. Fig. 2 illustrates the underlying workflow of BBO-SYN and how it is used to generate a synthesis tree, like in Fig. 1. BBO-SYN uses a Monte Carlo Tree Search (MCTS) based latent space partitioning algorithm, LaP^3 (Yang et al., 2021), to find effective reactants for building synthesis trees. BBO-SYN improves over DAGs (Bradshaw et al., 2020), a template-free synthesis planner with discrete reactants, by converting the reactant selection problem to an optimization problem over a continuous latent space. In this

[1]Retrosynthesis: Given a product molecule, predict the reaction tree and reactants to synthesize it.



(a) JNK3 (b) GSK3β

Figure 1: Example synthesis trees generated by BBO-SYN for JNK3 and GSK3β properties. Left-side nodes are reactants found directly in the latent space via BBO-SYN . All products correspond to the top-1 Molecular Transformer prediction, with associated reaction probabilities shown.

way, BBO-SYN can easily handle any number of potential reactants and is agnostic to the chosen molecular representation. We use the same model-based chemical reaction predictor, Molecular Transformer (Schwaller et al., 2019), as DAGs to encourage the production of viable reaction steps, as depicted in Fig. 1. Fig. 1 illustrates the synthesis trees generated by BBO-SYN when optimizing for two distinct chemical properties. BBO-SYN progresses by adding more reactants with each step up the tree until reaching the final product. Finally, we show that BBO-SYN outperforms DAGs in four synthesis-aware property optimization tasks (QED, DRD2, GSK3β, and JNK3), increasing the product diversity by 37% and the mean property score by 25% on our hardest JNK3 task.

2 Related Work

2.1 Black Box Optimization

BBO methods constitute a flexible class of approaches that optimize a given function with little to no assumptions on its internal structure. Classical approaches such as CEM (Rubinstein, 1999) and CMA-ES (Hansen, 2006) learn a local model around promising trajectories; however, both greedily focus on promising regions of the search space and may get trapped in local optima. Other recent approaches, such as VOOT (Kim et al., 2020a) and DOO (Munos, 2011), use a recursive region partitioning scheme to alleviate the aforementioned issues. LA-MCTS (Wang et al., 2020a) and LaP^3 (Yang et al., 2021), a latent-space-based extension, further improve upon prior methods by adaptively partitioning the search regions based on sampled function values.

2.2 Chemical Reaction Prediction

A strong chemical reaction prediction system is critical for predicting viable synthesis pathways. There are two distinct methodologies for chemical reaction prediction: template-based and template-free. Template-based methods use chemical reaction rules based on subgroup pattern matching scraped from literature (Bøgevig et al., 2015; Szymkuć et al., 2016; Chen & Jung, 2021; Dai et al., 2019; Coley et al., 2017; Zhang et al., 2022b). These methods provide approximations for feasible reactions but are limited by the availability and specificity of applicable chemical reaction templates.

Template-free methods directly model chemical reactions to generalize to unseen reactions but can struggle with prediction accuracy compared to template-based methods especially when subject to out-of-domain data. Some methods rely on editing the graphical representation of molecules (Coley et al., 2019b; Sacha et al., 2021), others model the problem as a sequence-to-sequence generation problem (Schwaller et al., 2019; Lin et al., 2020; Duan et al., 2020), and more recent works leverage both representations for more efficient reaction prediction (Tu & Coley, 2022). We employ template-free methods for reaction prediction because such methods can in principle generalize beyond a limited number of available reaction templates. BBO-SYN takes advantage of such generalizability to optimize over its continuous molecular latent space.

2.3 Continuous Molecule Representations

As it is computationally infeasible to enumerate every drug-like molecule for de novo generation and molecular design, many methods choose to optimize over a fixed-sized continuous vector molecule representation. Such representations unlock a wide array of complex and in some cases fully differentiable optimization techniques that are intractable over a discrete set of molecules (Gómez-Bombarelli et al., 2018). For example, both SMILES-VAE (Gómez-Bombarelli et al., 2018) and MolMIM (Reidenbach et al., 2022) are models that generate novel molecules, optimized for molecular properties directly in their respective latent spaces.

Following this logic, to take advantage of recent advancements in BBO, BBO-SYN employs LaP^3 over a continuous latent space. We chose to work with HierVAE's (Jin et al., 2020) molecular representation as it has been well-benchmarked on several property-guided optimization tasks. HierVAE breaks down input 2D molecule graphs into common subgraph motifs, building a hierarchical autoencoder for autoregressive molecule generation. BBO-SYN expands on the prior idea of direct latent optimization by leveraging BBO in an iterative mechanism designed explicitly for synthesizability-constrained generation.

2.4 Synthesizability-Constrained Generation

As synthesizability-constrained generation is rooted in reaction prediction, there exist both template-based and template-free methods. Several early methods, such as SYNOPSIS (Vinkers et al., 2003) and DOGS (Hartenfeller et al., 2012), combine discrete synthetic building blocks for molecular design. RL methods such as PGFS (Gottipati et al., 2020), REACTOR (Horwood & Noutahi, 2020), and SynNet (Gao et al., 2022) use reaction templates to form a discrete action space for an actor-critic algorithm to generate optimal synthesis trees. Popular template-free methods include ChemBO (Korovina et al., 2020), which uses Bayesian optimization, and MoleculeChef (Bradshaw et al., 2019), which leverages latent gradients over fused reactant embeddings. Building on MoleculeChef, DAGs (Bradshaw et al., 2020) uses an iterative RL finetuning scheme over whole synthesis tree embeddings for synthesizability-constrained molecular property optimization. While DAGs limits its entire generative process to a small discrete set of reactants, BBO-SYN uses LaP^3, an MCTS-based latent search algorithm, to locate optimal reactants for building synthesis trees. Unlike DAGs, BBO-SYN can be scaled to handle extremely large reactant sets with little increase in computational cost due to its use of BBO and continuous molecular representations.

Due to the inherent tree structure of synthesizability-constrained generation, several methods use MCTS to generate plausible synthesis trees. AutoSynRoute (Lin et al., 2020) and SMC (Zhang et al., 2022b) use MCTS to explore pathways in template-free retrosynthesis and generate template-based reaction networks based on stacks of linear reactions. Several methods combine MCTS with RL to create dynamic synthesis solutions. Segler et al. (2018b) uses MCTS with an expansion policy network to guide retrosynthetic pathways toward buyable reactants. Similarly, Wang et al. (2020b) integrates a learned value function with MCTS to discover efficient and safe synthesis pathways. In contrast, BBO-SYN employs MCTS in a fundamentally different role in the *inner* optimization loop for reactant selection. We note that modeling entire synthesis trees via MCTS is orthogonal to our proposed method; combining BBO-SYN with existing MCTS approaches would likely further improve performance in exchange for higher compute costs.

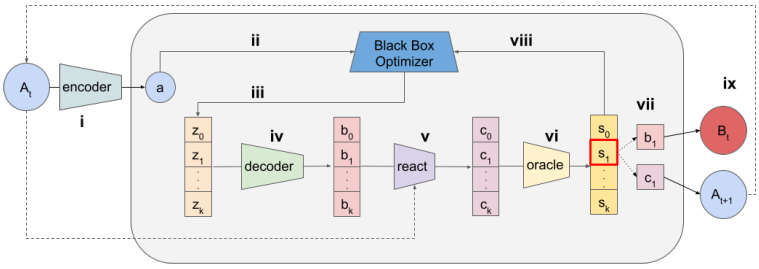


Figure 2: Internal diagram of BBO-SYN : (i) The input molecule A_t is encoded and then (ii) passed to the black-box optimizer. (iii) The optimizer generates k+1 latent points, which are then (iv) decoded into SMILES. (v) The input and latent reactants are reacted together where (vi) each product is scored by a given oracle function. (vii) The highest-scoring product and latent reactant are stored. (viii) The scores for all products are returned to the BBO to score the latent points generated in step iii to update for the next iteration. (ix) The optimal reactant B_t and product A_{t+1} are returned after all BBO iterations are complete and A_{t+1} becomes the input for the next synthesis iteration.

3 METHODS

3.1 OUTER SYNTHESIZABILITY OPTIMIZATION LOOP

We define our synthesis framework, BBO-SYN, as follows. BBO-SYN breaks down the task of molecule generation into two distinct optimization steps: an outer iterative loop and an inner BBO loop for latent reactant selection. The outer loop is where synthesis trees are built one node at a time, as shown in Fig. 1, to enforce strict model-based synthesizability constraints. BBO-SYN begins by encoding an input molecule and passing it to the black-box optimizer where LaP^3 is used in the inner loop to generate reactant options directly in the latent space (Fig. 2 i-iii). Independent of the underlying BBO method, BBO-SYN scores proposed latent reactants z for each input molecule A according to: oracle(react(A, decode(z))) (Fig. 2 iv-vi). Every potential latent solution is decoded and reacted with our input molecule via template-free reaction prediction. After each sub-iteration of the black-box optimization, the product and accompanying reactant yielding the highest property score are saved, and then all scores are returned to the optimizer to update and continue the next sub-iteration (Fig. 2 vii-viii).[2]

Once the outer procedure is complete, the best product is used as the input reactant for the next BBO-SYN outer iteration, and the associated best reactant partner is returned (Fig. 2 ix). We show that we can greedily build synthesis trees by reusing the intermediate products as new inputs to generate a final optimal product for the desired chemical property. While greedy optimization might not necessarily lead to truly optimal synthesis trees, we found it worked well enough in practice.

3.2 INNER BLACK-BOX OPTIMIZATION LOOP

The inner optimization loop uses BBO to select reactants from a continuous latent space for the building of optimal synthesis trees. Specifically, BBO-SYN uses LaP^3, which iteratively samples latent points to learn a recursive space partition focusing on good regions while also still exploring bad regions using an upper confidence bound. While prior methods utilize MCTS to model entire synthesis trees (corresponding to our outer loop), LaP^3 uses MCTS to learn a space partitioning function which is used to produce optimal latent reactants for individual chemical reaction steps.

As synthesizability-constrained generation is driven by the choice of available reactants, we opted to use a continuous molecular representation to efficiently handle a variable number of candidate reactants without major changes to underlying methodology (Jin et al., 2020). As it is challenging to generate a smooth molecular latent space (Gómez-Bombarelli et al., 2018; Zhang et al., 2022a), we use BBO to alleviate the structural and optimization difficulties of working in a non-smooth space (Yang et al., 2021; Wang et al., 2020a). BBO thus provides us with a powerful solution to the

[2]We used the default number of LaP^3 iterations according to `https://github.com/yangkevin2/neurips2021-lap3`.

reactant selection component of template-free synthesizability-constrained generation. Furthermore, we stress BBO-SYN's modularity as it does not depend on any one choice in reaction predictor, molecular latent space, or optimization goal. Discrete methods such as DAGs must be retrained from scratch for each alteration. BBO-SYN, on the other hand, can seamlessly take advantage of future advances in adjacent areas of BBO, chemical reaction prediction, and latent molecular representation.

4 EXPERIMENTS

4.1 BASELINES AND TASK SETUP

We compare BBO-SYN to various DAGs' fine-tuned DoG-Gen models (Bradshaw et al., 2020), one per tested property holding fixed the initial set of starting molecules and the chemical reaction predictor. DAGs was chosen as it is the most expressive template-free forward synthesis planner to date that has been successfully applied to property-guided molecule generation. For the initial set of starting molecules, we use a subset of the starting molecules from DAGs' published validation set of crafted synthesis trees.[3] Both methods use DAGs' pre-trained Molecular Transformer model weights for template-free chemical reaction prediction to ensure a fair comparison. Both methods also only consider the top-1 Molecular Transformer generated products with no threshold for model confidence. BBO-SYN limits the depth of generated synthesis trees to 4, as each optimization step is computationally expensive.[4]

Given that BBO-SYN's generated synthesis trees are conditioned on a specified starting point, we filter the final DoG-Gen products to keep only the highest-scoring synthesis tree for each of the shared starting points.[5] This alignment step is necessary as DoG-Gen is fine-tuned by repeatedly re-training on its top-k seen trees and can only return a sorted list of every synthesis tree encountered during its iterative refinement. In this way, we can condition the outputs of both methods on the same discrete set of starting molecules.

Due to the different definitions and design choices of the components of the synthesis-aware generation task, a true head-to-head comparison is difficult to create. As a result, DoG-Gen is only used as an anchor point to understand BBO-SYN's performance. As BBO-SYN explores a continuous reactant latent space, it is not confined to the same set of discrete reactants as DoG-Gen. Our benchmarks are designed to mitigate the differences of continuous vs. discrete spaces as much as possible and provide extensive ablations to understand how BBO presents a robust and scalable solution to the synthesis-aware generation problem. Since the BBO-SYN latent spaces are trained to approximate the distribution of the discrete DoG-Gen reactants, and all synthesis trees start with one of the discrete reactants, we opt to use the initial starting molecule as the main equalizing criteria in the later comparisons.

4.2 REACTANT LATENT SPACE SET UP

We trained two distinct HierVAE (Jin et al., 2020) models with a 32-dimensional latent space. The first was trained solely on the DAGs published building blocks (4,343 molecules). The second latent space was trained on all unique reactant and product molecules in the USPTO_MIT data set (Jin et al., 2017) that was used to train the aforementioned Molecular Transformer. The second latent space also included the above building blocks filtered for SMILES length [3,45], resulting in a total of 404,898 molecules or potential reactants for forward synthesis. These two latent spaces are henceforth referred to as the small and large latent spaces, respectively. We applied BBO-SYN to optimize for various chemical properties over both latent spaces to understand how our framework would operate in various environments.

[3] We use SMILES with length in [5,25], totaling 2246 molecules.

[4] 99% of DoG-Gen trees had depth <= 4 with max depth of 10. Best DoG-Gen JNK3 trees were depth 4.

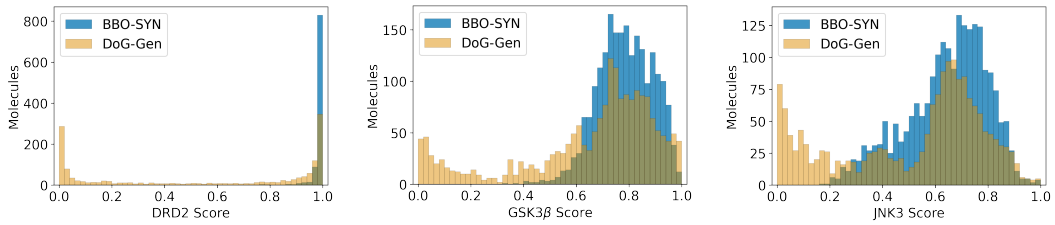
[5] Synthesis trees are scored by the property score of the final product molecule.

4.3 METRICS

We focus on each method's ability to generate diverse final product molecules with high property scores. As such, we report the property score distributions of the final product molecules of all generated synthesis trees. For all experiments, we used the TDC package (Huang et al., 2021) for property oracle functions (QED, DRD2, JNK3, GSK3β). We also report the internal diversity (IntDiv$_p$), defined as $1 - \sqrt[p]{\frac{1}{|G|^2}\sum_{m_1,m_2 \in G} T(m_1, m_2))^p}$ for a set of molecules G and Tanimoto similarity T taken from MOSES (Polykovskiy et al., 2020). A low percentage of diverse molecules illustrates a method's collapse to a select few solutions, i.e., a lack of generative robustness.

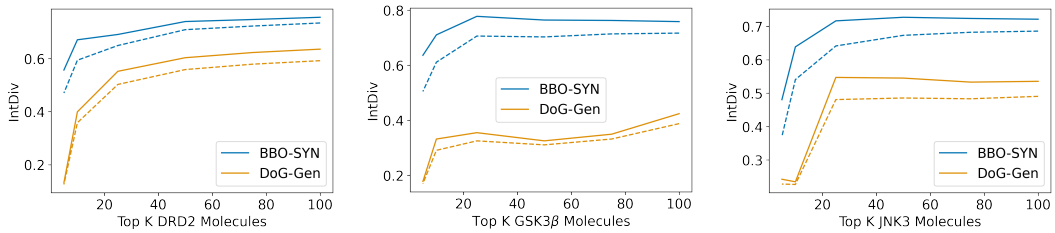
4.4 RESULTS

Here we present the respective optimized product distributions for both DoG-Gen and BBO-SYN for DRD2, GSK3β, and JNK3.



(a) DRD2 (b) GSK3β (c) JNK3

Figure 3: Property scores of final predicted molecules for BBO-SYN and DoG-Gen property-guided optimization. Both methods use the same initial building blocks and chemical reaction predictor.



(a) DRD2 (b) GSK3β (c) JNK3

Figure 4: Internal diversity of top 5, 10, 25, 50, 75, 100 molecules. Solid and dashed lines correspond to IntDiv$_1$ and IntDiv$_2$.

Fig. 3 shows that BBO-SYN significantly eliminates weaker products while maintaining top-end performance. We see in Fig. 4 BBO-SYN produces more unique high-scoring molecules for all properties. We hypothesize that the large increase in diversity (37% for top 100 JNK3 molecules) is due to BBO-SYN's independent optimization for each input. Compared to DoG-Gen's bulk fine-tuning, BBO-SYN searches the reactant space in parallel to find the best reactants for each starting molecule.

BBO-SYN finds a unique optimal product for nearly every input which is desirable when developing novel molecules. By actively searching for promising reactants, BBO-SYN avoids converging to a small set of solutions, as seen in DoG-Gen. Although BBO-SYN has a higher computational cost due to the inner loop LaP^3 optimization steps, we observed that giving DoG-Gen additional training iterations to equalize the property oracle budget of BBO-SYN resulted in no discernible difference in the resulting property distributions. We suspect that because DoG-Gen is repeatedly fine-tuned on its top encountered synthesis trees, increasing DoG-Gen's oracle budget only further increases the apparent mode collapse. We also acknowledge the flaws of template-free reaction prediction as seen in the low model confidence reactions in Fig. 1. While it does not impact the comparison between

BBO-SYN and DoG-Gen as they use the same reaction prediction, in the future, reaction confidence can be directly optimized by incorporating it into the BBO scoring function.

4.5 Analysis and Ablations

Below we analyze the effect of the choice of BBO-SYN's BBO method as well as the latent space size on the property scores and diversity of the generated synthesis trees. Specifically, we compare two BBO methods, a simple CMA-ES, and LaP^3, over both the small and large latent spaces for a series of property optimization tasks. We utilized a SMILES length penalty on all CMA-ES proposed reactants to prevent exploding sequence lengths due to sampling from non-smooth regions of the latent space.[6] No length penalties were needed for LaP^3 .

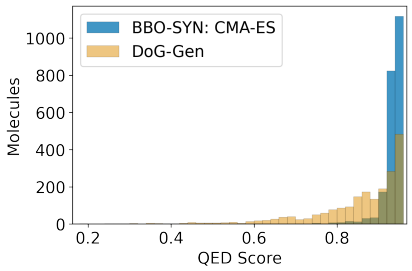


Figure 5: Unique QED Synthesis Product Distributions. Even when LaP^3 is replaced with a weaker optimization method, CMA-ES, BBO-SYN outperforms DoG-Gen when using the same set of starting points and chemical reaction predictor.

Fig. 5 demonstrates that BBO-SYN needs only a simple CMA-ES to achieve strong performance for QED. Compared to DoG-Gen, BBO-SYN increases the diversity of the top 100 molecules by 5%. However, the story changes when we compare CMA-ES and LaP^3 on the more challenging properties: DRD2, GSK3β, and JNK3.

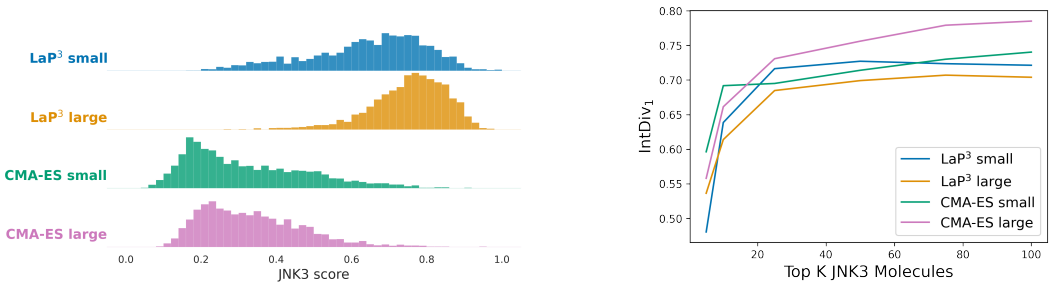
CMA-ES struggles to generate molecules with JNK3 greater than 0.6, whereas that is where most of the LaP^3 optimized results are located (Fig. 6(a)). Similar behavior can be seen for GSK3β and DRD2 in Appendix Fig. 9(a)- 10(a) as CMA-ES tends to generate more broad distribution whereas LaP^3 is more concentrated around high scoring molecules. Fig. 6(b) shows the internal diversity for the analyzed generated products. We point to the significant gap in property optimization performance as the reason for CMA-ES achieving higher diversity, i.e., it is easy to generate diverse products when they are not strongly optimized for a specific property.

For all three tested properties, the property scores and diversity of generated molecules also depend on the number of reactants considered or, in the case of BBO-SYN , the size of latent space used. Fig. 6(a) illustrates the impact of the number of available reactants on generating high-scoring synthesis trees. LaP^3 achieves significantly better property scores when given the large latent space that was trained with 100x the molecules as the small. Interestingly CMA-ES seems to prefer the small latent space for top-end performance. Fig. 6(b) shows how the internal diversity is correlated with the width of the property distribution, and as a result, CMA-ES and LaP^3 on the large latent space result in the best and worst, respectively. Similar results can be found in Appendix Fig. 9-10 for GSK3β and DRD2.

We note that a significant advantage of BBO-SYN over DoG-Gen is its ability to consider 100x more reactants by using a different latent space, with negligible increase in computational cost. It is infeasible to run DAGs with over 400k potential reactants.

However, LaP^3 's improved performance on the large latent space may come at the cost of complete synthesizability. When using the large latent space, it is possible that chosen reactants may actually be intermediate products, resulting in *convergent synthesis*: molecule X can react with another intermediate Y, rather than requiring Y to be a starting material. Convergent synthesis poses new difficulties but may also increase the potential flexibility of the method. For example, by introducing

[6]Property scores of product SMILES of length ≥ 70 and reactant SMILES of length ≥ 55 were reduced by a factor of 10.



(a) (b)

Figure 6: (a) JNK3 property score distributions of generated molecules for various latent space sizes and BBO methods. (b) Internal diversity (IntDiv_1) of top 5, 10, 25, 50, 75, 100 JNK3-optimized molecules.

simple RL as seen in DAGs and PGFS, one could at each step allow BBO-SYN to choose between adding to the tree directly (as we do currently) or picking two latent reactants to produce a convergent reactant (Gottipati et al., 2020). While we do not explicitly attempt this, we expect that this procedure should be feasible because BBO-SYN's large latent space is trained on USPTO reactants and their single reaction step products.

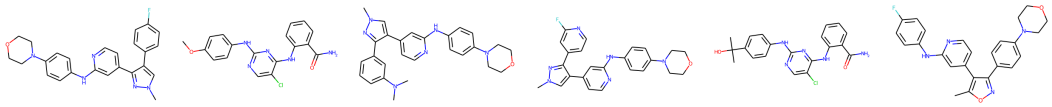


Figure 7: Top 6 BBO-SYN generated products optimized for JNK3.

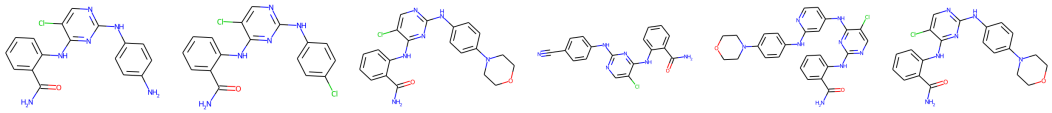


Figure 8: Top 6 DoG-Gen generated products optimized for JNK3.

Lastly, we provide a qualitative comparison of the diversity of the top 6 generated products for JNK3 in Fig. 7-8. Here it can be seen that DoG-Gen mostly makes minor updates to the end of the same underlying molecular scaffold, whereas BBO-SYN generates more geometrically different molecules. Similar results for GSK3β and DRD2 can be found in Appendix Fig. 12 - 15.

5 Conclusion

In this work, we introduce BBO-SYN, a synthesizability-constrained generative framework that leverages template-free chemical reaction prediction to build property-guided synthesis trees. BBO-SYN uses LaP^3 over a latent space of viable reactant molecules to select optimal reactants to produce products with high desired property scores. We show that BBO-SYN achieves state-of-the-art performance on QED, DRD2, JNK3, and GSK3β guided synthesis tasks by substantially increasing product diversity while maintaining high property scores. BBO-SYN can uniquely handle any number of reactants with relative ease compared to discrete reactant systems. Furthermore, BBO-SYN's latent space, black-box optimization method, and reaction predictor can be easily swapped out for orthogonal future advancements. The sampling efficiency of black-box optimization methods could also be further explored and is a key component to extending BBO-SYN to more complex tasks such as protein docking.

REFERENCES

Anders Bøgevig, Hans-Jürgen Federsel, Fernando Huerta, Michael G. Hutchings, Hans Kraut, Thomas Langer, Peter Löw, Christoph Oppawsky, Tobias Rein, and Heinz Saller. Route design in the 21st century: The icsynth software tool as an idea generator for synthesis prediction. *Organic Process Research & Development*, 19(2):357–368, 02 2015. doi: 10.1021/op500373e. URL https://doi.org/10.1021/op500373e.

John Bradshaw, Brooks Paige, Matt J. Kusner, Marwin H. S. Segler, and José Miguel Hernández-Lobato. A model to search for synthesizable molecules, 2019.

John Bradshaw, Brooks Paige, Matt J Kusner, Marwin Segler, and José Miguel Hernández-Lobato. Barking up the right tree: an approach to search over molecule synthesis dags. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6852–6866. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/file/4cc05b35c2f937c5bd9e7d41d3686fff-Paper.pdf.

Shuan Chen and Yousung Jung. Deep retrosynthetic reaction prediction using local reactivity and global attention. *JACS Au*, 1(10):1612–1620, 10 2021. doi: 10.1021/jacsau.1c00246. URL https://doi.org/10.1021/jacsau.1c00246.

Connor W. Coley, Luke Rogers, William H. Green, and Klavs F. Jensen. Computer-assisted retrosynthesis based on molecular similarity. *ACS Central Science*, 3(12):1237–1245, 12 2017. doi: 10.1021/acscentsci.7b00355. URL https://doi.org/10.1021/acscentsci.7b00355.

Connor W. Coley, William H. Green, and Klavs F. Jensen. Machine learning in computer-aided synthesis planning. *Accounts of Chemical Research*, 51(5):1281–1289, 2018a. doi: 10.1021/acs.accounts.8b00087. URL https://doi.org/10.1021/acs.accounts.8b00087. PMID: 29715002.

Connor W. Coley, Luke Rogers, William H. Green, and Klavs F. Jensen. Scscore: Synthetic complexity learned from a reaction corpus. *Journal of Chemical Information and Modeling*, 58(2):252–261, 02 2018b. doi: 10.1021/acs.jcim.7b00622. URL https://doi.org/10.1021/acs.jcim.7b00622.

Connor W. Coley, Dale A. Thomas, Justin A. M. Lummiss, Jonathan N. Jaworski, Christopher P. Breen, Victor Schultz, Travis Hart, Joshua S. Fishman, Luke Rogers, Hanyu Gao, Robert W. Hicklin, Pieter P. Plehiers, Joshua Byington, John S. Piotti, William H. Green, A. John Hart, Timothy F. Jamison, and Klavs F. Jensen. A robotic platform for flow synthesis of organic compounds informed by ai planning. *Science*, 365(6453):eaax1566, 2019a. doi: 10.1126/science.aax1566. URL https://www.science.org/doi/abs/10.1126/science.aax1566.

Connor W. Coley, Wengong Jin, Luke Rogers, Timothy F. Jamison, Tommi S. Jaakkola, William H. Green, Regina Barzilay, and Klavs F. Jensen. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem. Sci.*, 10:370–377, 2019b. doi: 10.1039/C8SC04228D. URL http://dx.doi.org/10.1039/C8SC04228D.

Hanjun Dai, Chengtao Li, Connor Coley, Bo Dai, and Le Song. Retrosynthesis prediction with conditional graph logic network. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/file/0d2b2061826a5df3221116a5085a6052-Paper.pdf.

Hongliang Duan, Ling Wang, Chengyun Zhang, Lin Guo, and Jianjun Li. Retrosynthesis with attention-based nmt model and chemical analysis of "wrong" predictions. *RSC Adv.*, 10:1371–1378, 2020. doi: 10.1039/C9RA08535A. URL http://dx.doi.org/10.1039/C9RA08535A.

Ola Engkvist, Josep Arús-Pous, Esben Jannik Bjerrum, and Hongming Chen. Chapter 13 molecular de novo design through deep generative models. In *Artificial Intelligence in Drug Discovery*, pp. 272–300. The Royal Society of Chemistry, 2021. ISBN 978-1-78801-547-9. doi: 10.1039/9781788016841-00272. URL http://dx.doi.org/10.1039/9781788016841-00272.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics*, 1(1):8, Jun 2009. ISSN 1758-2946.

Wenhao Gao, Rocío Mercado, and Connor W. Coley. Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=FRxhHdnxt1.

Sai Krishna Gottipati, Boris Sattarov, Sufeng Niu, Yashaswi Pathak, Haoran Wei, Shengchao Liu, Karam J. Thomas, Simon Blackburn, Connor W. Coley, Jian Tang, Sarath Chandar, and Yoshua Bengio. Learning to navigate the synthetically accessible chemical space using reinforcement learning. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020.

Rafael Gómez-Bombarelli, Jennifer N. Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, Ryan P. Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 2018. doi: 10.1021/acscentsci.7b00572. URL https://doi.org/10.1021/acscentsci.7b00572. PMID: 29532027.

Nikolaus Hansen. The cma evolution strategy: A comparing review, 2006.

Markus Hartenfeller, Heiko Zettl, Miriam Walter, Matthias Rupp, Felix Reisen, Ewgenij Proschak, Sascha Weggen, Holger Stark, and Gisbert Schneider. DOGS: reaction-driven de novo design of bioactive compounds. *PLoS Comput Biol*, 8(2):e1002380, February 2012.

Julien Horwood and Emmanuel Noutahi. Molecular design in synthetically accessible chemical space via deep reinforcement learning. *ACS Omega*, 5(51):32984–32994, December 2020.

Kexin Huang, Tianfan Fu, Wenhao Gao, Yue Zhao, Yusuf Roohani, Jure Leskovec, Connor W Coley, Cao Xiao, Jimeng Sun, and Marinka Zitnik. Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development. *NeurIPS Datasets and Benchmarks*, 2021.

Wengong Jin, Connor W. Coley, Regina Barzilay, and Tommi Jaakkola. Predicting organic reaction outcomes with weisfeiler-lehman network. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pp. 2604–2613, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964.

Wengong Jin, Regina Barzilay, and Tommi S. Jaakkola. Hierarchical generation of molecular graphs using structural motifs. *CoRR*, abs/2002.03230, 2020. URL https://arxiv.org/abs/2002.03230.

Beomjoon Kim, Kyungjae Lee, Sungbin Lim, Leslie Kaelbling, and Tomas Lozano-Perez. Monte carlo tree search in continuous spaces using voronoi optimistic optimization with regret bounds. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(06):9916–9924, Apr. 2020a. doi: 10.1609/aaai.v34i06.6546. URL https://ojs.aaai.org/index.php/AAAI/article/view/6546.

Hyunho Kim, Eunyoung Kim, Ingoo Lee, Bongsung Bae, Minsu Park, and Hojung Nam. Artificial intelligence in drug discovery: A comprehensive review of data-driven and machine learning approaches. *Biotechnol. Bioprocess Eng.*, 25(6):895–930, 2020b.

Ksenia Korovina, Sailun Xu, Kirthevasan Kandasamy, Willie Neiswanger, Barnabás Póczos, Jeff Schneider, and Eric P. Xing. Chembo: Bayesian optimization of small organic molecules with synthesizable recommendations. In *AISTATS*, pp. 3393–3403, 2020. URL http://proceedings.mlr.press/v108/korovina20a.html.

James Law, Zsolt Zsoldos, Aniko Simon, Darryl Reid, Yang Liu, Sing Yoong Khew, A. Peter Johnson, Sarah Major, Robert A. Wade, and Howard Y. Ando. Route designer: A retrosynthetic analysis tool utilizing automated retrosynthetic rule generation. *Journal of Chemical Information and Modeling*, 49(3):593–602, 03 2009. doi: 10.1021/ci800228y. URL https://doi.org/10.1021/ci800228y.

Kangjie Lin, Youjun Xu, Jianfeng Pei, and Luhua Lai. Automatic retrosynthetic route planning using template-free models. *Chem. Sci.*, 11:3355–3364, 2020. doi: 10.1039/C9SC03666K. URL http://dx.doi.org/10.1039/C9SC03666K.

Rémi Munos. Optimistic optimization of a deterministic function without the knowledge of its smoothness. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K.Q. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011. URL https://proceedings.neurips.cc/paper/2011/file/7e889fb76e0e07c11733550f2a6c7a5a-Paper.pdf.

Daniil Polykovskiy, Alexander Zhebrak, Benjamin Sanchez-Lengeling, Sergey Golovanov, Oktai Tatanov, Stanislav Belyaev, Rauf Kurbanov, Aleksey Artamonov, Vladimir Aladinskiy, Mark Veselov, Artur Kadurin, Simon Johansson, Hongming Chen, Sergey Nikolenko, Alan Aspuru-Guzik, and Alex Zhavoronkov. Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models. *Frontiers in Pharmacology*, 2020.

Danny Reidenbach, Micha Livne, Rajesh K. Ilango, Michelle Gill, and Johnny Israeli. Improving small molecule generation using mutual information machine, 2022. URL https://arxiv.org/abs/2208.09016.

Jean-Louis Reymond and Mahendra Awale. Exploring chemical space for drug discovery using the chemical universe database. *ACS Chemical Neuroscience*, 3(9):649–657, 09 2012. doi: 10.1021/cn3000422. URL https://doi.org/10.1021/cn3000422.

Reuven Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology And Computing In Applied Probability*, 1(2):127–190, 1999. doi: 10.1023/A:1010091220143. URL https://doi.org/10.1023/A:1010091220143.

Mikołaj Sacha, Mikołaj Błaż, Piotr Byrski, Paweł Dąbrowski-Tumański, Mikołaj Chromiński, Rafał Loska, Paweł Włodarczyk-Pruszyński, and Stanisław Jastrzębski. Molecule edit graph attention network: Modeling chemical reactions as sequences of graph edits. *Journal of Chemical Information and Modeling*, 61(7):3273–3284, 2021. doi: 10.1021/acs.jcim.1c00537. URL https://doi.org/10.1021/acs.jcim.1c00537. PMID: 34251814.

Philippe Schwaller, Teodoro Laino, Théophile Gaudin, Peter Bolgar, Christopher A. Hunter, Costas Bekas, and Alpha A. Lee. Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Central Science*, 5(9):1572–1583, 09 2019. doi: 10.1021/acscentsci.9b00576. URL https://doi.org/10.1021/acscentsci.9b00576.

Marwin H. S. Segler, Mike Preuss, and Mark P. Waller. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604–610, 2018a. doi: 10.1038/nature25978. URL https://doi.org/10.1038/nature25978.

Marwin H. S. Segler, Mike Preuss, and Mark P. Waller. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604–610, 2018b. doi: 10.1038/nature25978. URL https://doi.org/10.1038/nature25978.

Sara Szymkuć, Ewa P. Gajewska, Tomasz Klucznik, Karol Molga, Piotr Dittwald, Michał Startek, Michał Bajczyk, and Bartosz A. Grzybowski. Computer-assisted synthetic planning: The end of the beginning. *Angewandte Chemie International Edition*, 55(20):5904–5937, 2016. doi: https://doi.org/10.1002/anie.201506101. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201506101.

Zhengkai Tu and Connor W. Coley. Permutation invariant graph-to-sequence model for template-free retrosynthesis and reaction prediction, 2022. URL https://openreview.net/forum?id=LLHwQh9zEb.

H Maarten Vinkers, Marc R de Jonge, Frederik F D Daeyaert, Jan Heeres, Lucien M H Koymans, Joop H van Lenthe, Paul J Lewi, Henk Timmerman, Koen Van Aken, and Paul A J Janssen. SYNOPSIS: SYNthesize and OPtimize system in silico. *J Med Chem*, 46(13):2765–2773, June 2003.

Linnan Wang, Rodrigo Fonseca, and Yuandong Tian. Learning search space partition for black-box optimization using monte carlo tree search. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20, Red Hook, NY, USA, 2020a. Curran Associates Inc. ISBN 9781713829546.

Xiaoxue Wang, Yujie Qian, Hanyu Gao, Connor W. Coley, Yiming Mo, Regina Barzilay, and Klavs F. Jensen. Towards efficient discovery of green synthetic pathways with monte carlo tree search and reinforcement learning. *Chem. Sci.*, 11:10959–10972, 2020b. doi: 10.1039/D0SC04184J. URL http://dx.doi.org/10.1039/D0SC04184J.

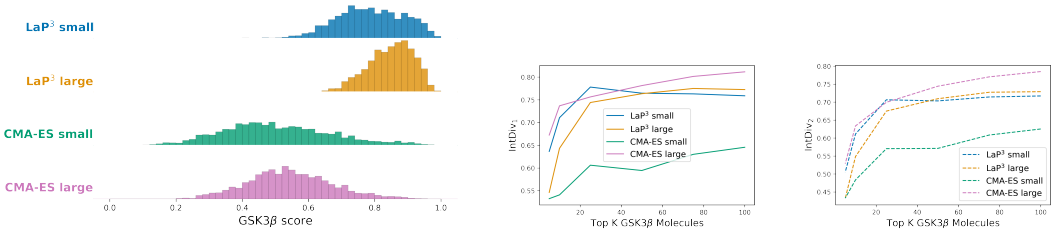
Kevin Yang, Tianjun Zhang, Chris Cummins, Brandon Cui, Benoit Steiner, Linnan Wang, Joseph E. Gonzalez, Dan Klein, and Yuandong Tian. Learning space partitions for path planning. In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan (eds.), *NeurIPS*, 2021.

Jiying Zhang, Xi Xiao, Long-Kai Huang, Yu Rong, and Yatao Bian. Fine-tuning graph neural networks via graph topology induced optimal transport, 2022a.

Qi Zhang, Chang Liu, Stephen Wu, and Ryo Yoshida. Bayesian sequential stacking algorithm for concurrently designing molecules and synthetic reaction networks, 2022b.

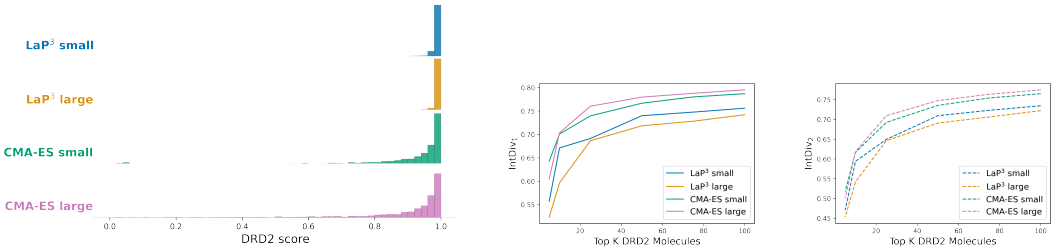
6 APPENDIX

Below we present additional figures for the GSK3β and DRD2 black-box method and latent space size ablations.



(a) (b) (c)

Figure 9: Optimization of GSK3β for various latent space sizes and BBO methods. All use the same initial starting molecules. (bc) Internal Diversity of top 5, 10, 25, 50, 75, 100 GSK3β optimized molecules. Solid and dashed lines correspond to IntDiv$_1$ and IntDiv$_2$.



(a) (b) (c)

Figure 10: (a) Optimization of DRD2 for various latent space sizes and BBO methods. All use the same initial starting molecules. (bc) Internal Diversity of top 5, 10, 25, 50, 75, 100 DRD2 optimized molecules. Solid and dashed lines correspond to IntDiv$_1$ and IntDiv$_2$.

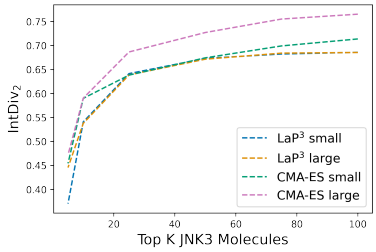


Figure 11: Internal Diversity of top 5, 10, 25, 50, 75, 100 JNK3 optimized molecules. Dashed lines correspond to IntDiv$_2$.

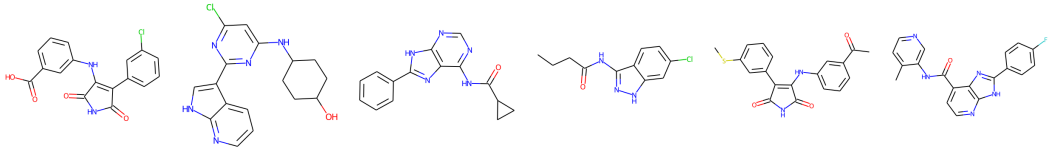


Figure 12: Top 6 BBO-SYN generated products optimized for GSK3β.

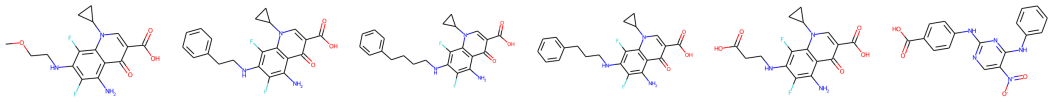


Figure 13: Top 6 DoG-Gen generated products optimized for GSK3β.

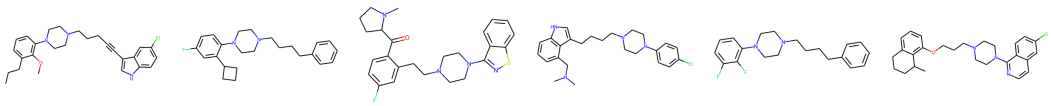


Figure 14: Top 6 BBO-SYN generated products optimized for DRD2.

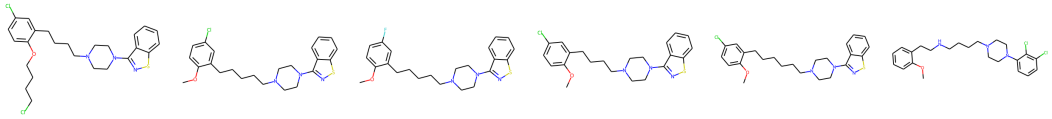


Figure 15: Top 6 DoG-Gen generated products optimized for DRD2.