Fragment-Based Sequential Translation for Molecular Optimization

Anonymous Author(s)
Affiliation
Address
email

Abstract

Search of novel molecular compounds with desired properties is an important problem in drug discovery. Many existing generative models for molecules operate on the atom level. We instead focus on generating molecular fragments–meaningful substructures of molecules. We construct a coherent latent representation for molecular fragments through a learned variational autoencoder (VAE) that is capable of generating diverse and meaningful fragments. Equipped with the learned fragment vocabulary, we propose **Fra**gment-based **S**equential **T**ranslation (FaST), which iteratively translates model-discovered molecules into increasingly novel molecules with high property scores. Empirical evaluation shows that FaST achieves significant improvement over state-of-the-art methods on benchmark single-objective/multi-objective molecular optimization tasks.

1 Introduction

Molecular optimization is a challenging task for drug discovery, and has been explored in previous work through several different generation methods, including VAE [5, 15], GAN [10], and RL [6]. These character-by-character (for SMILES/SELFIES strings) and node-by-node (for molecular graphs) models generate valid molecules, but can struggle to explore the complex chemical space under multiple property constraints. More recent works attempt to generate molecules using a predefined set of fragments [20, 13, 19] and achieve impressive empirical results. However, the fixed fragment vocabulary limits the generative capabilities of the models.

Shifting away from previous frameworks, we learn a distribution of molecular fragments using vector-quantized variational autoencoders (VQ-VAE) [18]. We then generate molecular graphs through addition and deletion of molecular fragments from the learned distributional fragment vocabulary, enabling the generative model to span a much larger chemical space than models with a fixed fragment vocabulary. Considering atomic edits as primitive actions, the idea of using fragments can be thought of as *options* [17, 16] as a way to simplify the search problem.

There are two primary generation schemes from previous works: (1) generating from scratch [20, 19] and (2) translating from known active molecules [8, 7]. Generation under the first scheme is usually very challenging because the set of molecules with high property score is typically a very small subspace of the entire chemical space. It is, in general, much easier to generate molecules satisfying desired properties under the translation scheme, being able to start from a prior over “good" molecules. However, this generation scheme suffers from generating molecules too similar to those in the active set, which is undesirable as this precludes the ability of the model to produce novel molecules for drug discovery applications.

To this end, we bridge the gap between the two aforementioned generation paradigms by introducing a novel *sequential translation* scheme. We start the molecular search by translating from known active

Submitted to NeurIPS 2021 AI for Science Workshop.

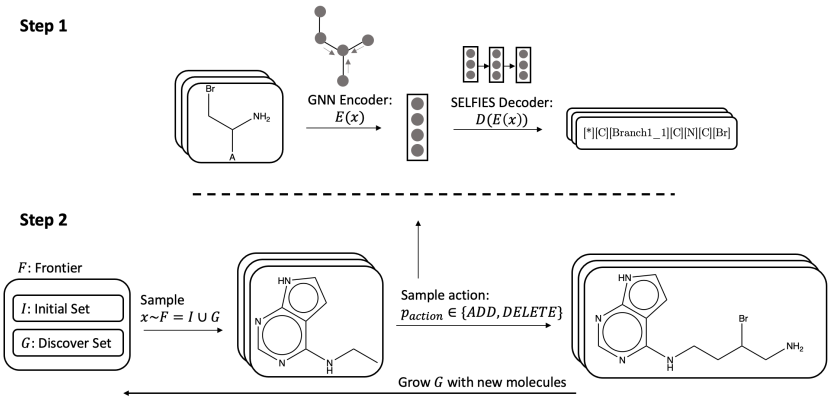

Figure 1: Overview of **Fra**gment-based **S**equential **T**ranslation (FaST), which consists primarily of two component steps. In the first step, we train a VQ-VAE that embeds molecular fragments. In the second step, we train a search policy that uses the learned embeddings as an action space. The search policy starts from the *frontier* set F, which consists of an initial set of good molecules (I), and good molecules discovered by the policy (G).

molecules, and store the discovered molecules as new potential initialization states for subsequent searches. As monotonic expansion of molecular graphs will end up producing undesirable, large molecules, we also include the deletion of fragments as a possible action. This enables our method to backtrack to good molecular states, and iteratively improve generated molecules during the sequential translation process. Our proposed framework is (1) highly efficient in finding molecules that satisfy property constraints since the model stay close to the high-property-score chemical manifold; and (2) able to produce highly novel molecules because the sequence of fragment-based translation can lead to very different and diverse molecules compared to the known active set.

2 Methods

Molecular Optimization as a Markov Decision Process. We model the molecular optimization problem as a Markov decision process (MDP), defined by the 5-tuple $\{\mathcal{S}, \mathcal{A}, p, r, \rho_0\}$, where the state space $\mathcal{S}$ is the set of all molecules. The goal of molecular optimization is to find a set of molecules $G \subset \mathcal{S}$ that has high quality (success), novelty, and diversity (detailed in Section 3). In order to achieve this goal, we introduce novel designs over the action space $\mathcal{A}$ (and the corresponding transition model $p : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$), the reward function r and the initial state distribution ρ_0. In summary, our action space $\mathcal{A}$ is based on molecular fragments learned by a VQ-VAE, while r and ρ_0 interact with policy learning to implement the proposed sequential translation optimization scheme. An illustration of our model is in Figure 1.

2.1 Fragment-based Molecular Generation

VQ-VAE Encoder/Decoder We first pretrain a VQ-VAE on molecular fragments, which uses a GNN encoder. GNNs are suitable for describing actions on the molecular state, as they explicitly parametrize the representations of each atom and bond. Meanwhile, the decoder architecture is a recurrent network that decodes a SELFIES string representation of a molecule. We choose a recurrent network for the decoder, because we do not need the full complexity of a graph decoder. Due to the construction scheme (see Appendix A.2), the fragments are rooted trees and all have a single attachment point. As our fragments are small in molecular size (≤ 10 atoms), the string grammar is simple to learn, and we find the SELFIES decoder works well empirically.

Adding and deleting fragments as actions. At each step of the MDP, the policy network first takes the current molecular graph as input and produce a Bernoulli distribution on whether to add or delete a fragment. Equipped with the fragment VQ-VAE, we define the *Add* and *Delete* actions at the fragment-level:

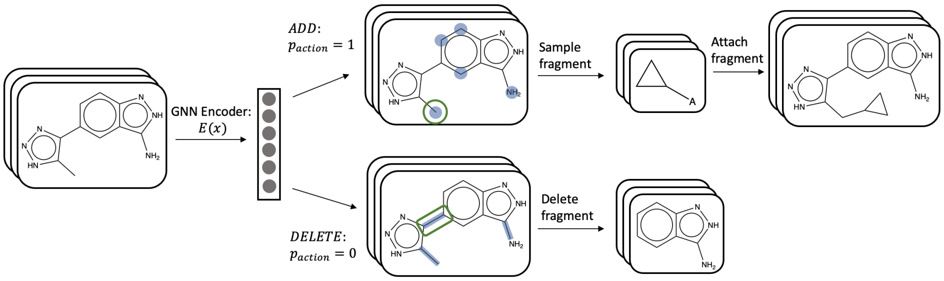

Figure 2: Each episode starts from a molecule sampled from the frontier. The molecule is encoded by a GNN, which is then used to predict either an *Add* or *Delete* action. When the *Add* action is selected, the model predicts and samples an atom as the attachment point, and subsequently predicts a fragment to attach to that atom. When the *Delete* action is selected, the model samples a directed edge, indicating the molecular fragment to be deleted.

- Fragment Addition. The addition action is characterized by a probability distribution over the atoms: $p_{add}(v_i) = \sigma[\text{MLP}(h_v)]$, where h_v is the output atom embedding of the GNN. Conditioned on the attachment point atom v_{add} sampled from p_{add}, we predict a categorical latent vector that is fed to the decoder: $z_{add} = \sigma[\text{MLP}([h_{v_{add}}; h_x])]$, where h_x is the embedding of the input molecular graph. The fragment to add is then obtained by decoding z_{add} through the learned fragment decoder.

- Fragment Deletion. The deletion action acts over the directed edges of the molecule. A probability distribution over deletable edges is computed with a MLP: $p_{del}(e_{ij}) = \sigma[\text{MLP}(h_{e_{ij}})]$, where $h_{e_{ij}}$ is the final edge embedding for edge e_{ij}. One edge is then sampled and deleted; since the edges are directed, the directionality specify the the molecule to keep and the fragment to be deleted.

With the action space $\mathcal{A}$ defined as above, the transition model for the MDP is simply $p(s'|s, a) = 1$ if applying the addition/deletion action a to s results in the molecule s', and $p(s'|s, a) = 0$ otherwise. We terminate an episode when the molecule fails to satisfy the desired property or when the episode exceeds 10 steps. The fragment-based action space is powerful as it (1) is powered by the enormous distributional vocabulary learned by the fragment VQ-VAE, thus spans a diverse set of editing operations over molecular graphs; (2) exploits the meaningful latent representation of fragments, since the representation of similar fragments are grouped together. These advantages greatly simplify the molecular search problem. An illustration of the two types of actions is given in Figure 2.

2.2 Discover Novel Molecules through Sequential Translation

We propose sequential translation that incrementally grows the set of discovered novel molecules, and uses the model-discovered molecules as starting points for further search episodes. This regime of starting exploration from states reached in previous episodes was also explored under the setting of RL from image inputs [2]. More concretely, we implement sequential translation with a reinforcement learning policy that operates under the fragment-based action space defined in Section 2.1, while using a moving initial state distribution ρ_0, which is a distribution over the *frontier* set F – the union of the initial set and good molecules that are discovered by the RL policy. We gradually expand the discovered set G by adding *qualified* molecules found in the RL exploration within the MDP. A molecule is qualified if it satisfies the desired properties and is novel compared to molecules currently in the frontier F, measured by fingerprint similarity. We use a simple binary reward of $+1$ for a transition that results in a molecule qualified for the set G, and a reward of 0 otherwise. We further discourage the model from producing invalid molecules by adding a reward of -0.1 for a transition that produces an invalid molecular graph. We further use an upper-confidence-bound (UCB) score to select good initial molecule from the frontier set. More implementation details of the method are included in the appendix.

Table 1: Results on our model (FaST) against multiple baselines. FaST outperforms all the baselines on both single-property optimization and multi-property optimization.

Model	GSK3β				GSK3β+QED+SA			
	SR	Nov	Div	PM	SR	Nov	Div	PM
Rationale-RL	1.00	.534	.888	.474	.699	.402	.893	.251
GA+D	.846	1.00	.714	.600	.891	1.00	.628	.608
JANUS	1.00	.829	.884	.732	-	-	-	-
MARS	1.00	.840	.718	.600 (± .04)	.995	.950	.719	.680 (± .03)
MARS+Rationale	.995	.804	.746	.597 (± .07)	.981	.800	.807	.632 (± .07)
FaST	1.00	1.00	.905	**.905 (± .000)**	1.00	1.00	.861	**.861 (± .001)**

Model	JNK3				JNK3+QED+SA			
	SR	Nov	Div	PM	SR	Nov	Div	PM
Rationale-RL	1.00	.462	.862	.400	.623	.376	.865	.203
GA+D	.528	.983	.726	.380	.857	.998	.504	.431
JANUS	1.00	.426	.895	.381	-	-	-	-
MARS	.988	.889	.748	.660 (± .04)	.913	.948	.779	.674 (± .02)
MARS+Rationale	.976	.843	.780	.642 (± .04)	.634	.779	.787	.386 (± .08)
FaST	1.00	1.00	.905	**.905 (± .001)**	1.00	.866	.856	**.741 (± .001)**

Model	GSK3β+JNK3				GSK3β+JNK3+QED+SA			
	SR	Nov	Div	PM	SR	Nov	Div	PM
Rationale-RL	1.00	.973	.824	.800	.750	.555	.706	.294
GA+D	.847	1.00	.424	.360	.857	1.00	.363	.311
JANUS	1.00	.778	.875	.681	1.00	.326	.821	.268
MARS	.995	.753	.691	.520 (± .08)	.923	.824	.719	.547 (± .05)
MARS+Rationale	.976	.843	.780	.642 (± .04)	.654	.687	.724	.321 (± .09)
FaST	1.00	1.00	.863	**.863 (± .001)**	1.00	1.00	.716	**.716 (± .011)**

3 Experiments

Datasets, evaluation, and baselines. We use benchmark datasets for molecular optimization, which aims to generate ligand molecules for inhibition of two proteins: glycogen synthase kinase-3 beta (GSK3β) and c-Jun N-terminal kinase 3 (JNK3). We also optimize for quantitative estimate of drug-likeliness (QED) [1] and synthetic accessibility (SA) [3] as done in previous work. Following previous works, we evaluate our generative model on three target metrics, success, novelty and diversity. The metric scores are computed from 5,000 molecules generated by the model. Our model is initialized with molecule rationales as obtained in Jin et al. [8]. We compare to state-of-the-art molecular optimization methods including **Rationale-RL** [8] (molecular rationales as intialization + atom-by-atom RL compeletion); **GA+D & JANUS** [11, 12] (genetic algorithms); and **MARS** [19] & **MARS+Rationale** (MCMC sampler intialized with/without rationales). More details on the datasets, evaluation metrics, and baseline methods are included in the appendix.

Performance FaST outperforms all the baselines on all tasks including both single-property and multi-property optimization. For the most challenging task, GSK3β+JNK3+QED+SA, our model improves upon the previous best model by over 30% in the product of the three evaluation metrics. The MARS+Rationale model, which uses the same rationale molecules as the initialization for their search algorithm, does not perform well compared to the original implementation, which initializes each search with a simple “C-C" molecule. Our model is able to efficiently search for molecules that stay within the constrained property space, and discover novel and diverse molecules by sequentially translating known active molecules.

4 Conclusion

We propose a new framework for molecular optimization, which leverages a learned representation of molecular fragments to search the chemical space efficiently. We demonstrate that our search method, which adaptively grows a set of promising molecular candidates, can achieve high performance on single-property and multi-property optimization tasks.

References

[1] G Richard Bickerton, Gaia V Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L Hopkins. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012. 4

[2] Adrien Ecoffet, Joost Huizinga, Joel Lehman, Kenneth O Stanley, and Jeff Clune. First return, then explore. *Nature*, 590(7847):580–586, 2021. 3

[3] 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):1–11, 2009. 4

[4] Anna Gaulton, Louisa J Bellis, A Patricia Bento, Jon Chambers, Mark Davies, Anne Hersey, Yvonne Light, Shaun McGlinchey, David Michalovich, Bissan Al-Lazikani, et al. Chembl: a large-scale bioactivity database for drug discovery. *Nucleic acids research*, 40(D1):D1100–D1107, 2012. 7

[5] 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. 1

[6] Gabriel Lima Guimaraes, Benjamin Sanchez-Lengeling, Carlos Outeiral, Pedro Luis Cunha Farias, and Alán Aspuru-Guzik. Objective-reinforced generative adversarial networks (organ) for sequence generation models. *arXiv preprint arXiv:1705.10843*, 2017. 1

[7] Wengong Jin, Kevin Yang, Regina Barzilay, and Tommi Jaakkola. Learning multimodal graph-to-graph translation for molecule optimization. In *International Conference on Learning Representations*, 2019. URL `https://openreview.net/forum?id=B1xJAsA5F7`. 1

[8] Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Multi-objective molecule generation using interpretable substructures. In *International Conference on Machine Learning*, pages 4849–4859. PMLR, 2020. 1, 4, 6, 7

[9] Yibo Li, Liangren Zhang, and Zhenming Liu. Multi-objective de novo drug design with conditional graph generative model. *Journal of cheminformatics*, 10(1):1–24, 2018. 6

[10] Łukasz Maziarka, Agnieszka Pocha, Jan Kaczmarczyk, Krzysztof Rataj, Tomasz Danel, and Michał Warchoł. Mol-cyclegan: a generative model for molecular optimization. *Journal of Cheminformatics*, 12(1):1–18, 2020. 1

[11] AkshatKumar Nigam, Pascal Friederich, Mario Krenn, and Alán Aspuru-Guzik. Augmenting genetic algorithms with deep neural networks for exploring the chemical space. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020. URL `https://openreview.net/forum?id=H1lmyRNFvr`. 4, 7

[12] AkshatKumar Nigam, Robert Pollice, and Alan Aspuru-Guzik. Janus: Parallel tempered genetic algorithm guided by deep neural networks for inverse molecular design. *arXiv preprint arXiv:2106.04011*, 2021. 4, 7

[13] Marco Podda, Davide Bacciu, and Alessio Micheli. A deep generative model for fragment-based molecule generation. In *International Conference on Artificial Intelligence and Statistics*, pages 2240–2250. PMLR, 2020. 1

[14] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. 6

[15] Martin Simonovsky and Nikos Komodakis. Graphvae: Towards generation of small graphs using variational autoencoders. In *International conference on artificial neural networks*, pages 412–422. Springer, 2018. 1

[16] Martin Stolle and Doina Precup. Learning options in reinforcement learning. In *International Symposium on abstraction, reformulation, and approximation*, pages 212–223. Springer, 2002. 1

[17] Richard S. Sutton, Doina Precup, and Satinder Singh. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112(1): 181–211, 1999. 1

[18] Aäron van den Oord, Oriol Vinyals, and Koray Kavukcuoglu. Neural discrete representation learning. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 6306–6315, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/7a98af17e63a0ac09ce2e96d03992fbc-Abstract.html. 1

[19] Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, and Lei Li. {MARS}: Markov molecular sampling for multi-objective drug discovery. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=kHSu4ebxFXY. 1, 4, 6, 7

[20] Jiaxuan You, Bowen Liu, Zhitao Ying, Vijay S. Pande, and Jure Leskovec. Graph convolutional policy network for goal-directed molecular graph generation. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 6412–6422, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/d60678e8f2ba9c540798ebbde31177e8-Abstract.html. 1

A Appendix

A.1 Reinforcement learning algorithm details

We detail the specifics of our reinforcement learning algorithm in Algorithm 1. To bias the initial state distribution to favor molecules that can derive more novel high quality molecules, we keep an upper-confidence-bound (UCB) score for each initial molecule in the frontier F. We record the number of times we initiate a search $N(x,t)$ from a molecule $x \in F$, and the number of molecules qualified for adding to G that are found in episodes starting from x: $R(x,t)$. Here $t = \sum_{x \in \rho_0} N(x)$ is the total number of search episodes. The UCB score of the initial molecule m is calculated by:

$$UCB(x,t) = \frac{R(x,t)}{N(x,t)} + \frac{\sqrt{\frac{3}{2}\log(t+1)}}{N(x,t)} \tag{1}$$

The probability of a molecule in the initalization set being sampled as the starting point of a new episode is then computed by a softmax over the UCB scores: $p_{init}(x,t+1) = \frac{\exp(UCB(x,t))}{\sum_{x \in I} \exp(UCB(x,t))}$.

We train our RL policy using the Proximal Policy Optimization (PPO, Schulman et al. 14) algorithm. We find the RL training robust despite both the reward function r and the initial state distribution ρ_0 are non-stationary (i.e., changing during the course of RL training). We construct the initial set of molecules for our search algorithm from the rationales extracted from [8]. These rationales are obtained through a sampling process, Monte Carlo Tree Search (MCTS), on the active molecules that tries to minimize the size of the rationale subgraph, while maintaining their inhibitory properties. Rationales for multi-property tasks (GSK3β+JNK3) are obtained by combining the rationales for single-property tasks.

A.2 Experimental setup

Datasets. The dataset, originally extracted from ExCAPE-DB, contains 2665 and 740 actives for GSK3β and JNK3 respecitvely. Each target also contains 50,000 negative ligand molecules. Following previous works [9, 8, 19], we adopt the same strategy of using a random forest trained on these datasets as the oracle property predictor. QED is a quantitative score that assesses the quality of a molecule through comparisons of its physicochemical properties to approved drugs. SA is a score that accounts for the complexity of the molecule in the context of easiness of synthesis, thereby providing an auxiliary metric for the feasibility of the compound as a drug candidate.

Algorithm 1 Molecule Search through Fragment-based Sequential Translation (FaST)

```
Input: $N$ is the desired number of discovered new molecules
Input: $I$ is the initial set of molecules
Input: $D$ is the decoder pretrained using the VQ-VAE
Input: $T : X \to \{0, 1\}$ is the episode termination criterion function given an input molecule $x$
Input: $C : X \to \{0, 1\}$ is a function that returns 1 if the input $x$ satisfies the desired properties.
Let $G = \emptyset$ be the discovered set of molecules
Let $F = I \cup G$ be the frontier where search is initialized from
Let $t = 0$ be the number of episodes
while $|G| \leq N$ do
    Let $t = t + 1$
    Update $UCB(x, t) \forall x \in F$ according to Equation (1)
    Sample initial molecule $x_0 = (V, E)$ from $p_{init} = \sigma[UCB(x, t)] \forall x \in F$
    Let $x = x_0$
    while $T(x) = 0$ do
        Sample action type $a$ from $p_{action} = \sigma[\text{MLP}(h_x)] \in \{\text{ADD}, \text{DELETE}\}$
        if $a = \text{ADD}$ then
            Sample $v_{add}$ from $p_{add}(v) = \sigma[\text{MLP}(h_v)]\ \forall v \in V$
            propose fragment encoding as action $f(x, v_{add}) = \text{MLP}([h_x; h_{v_{add}}])$
            Decode fragment $y = D(f(x, v_{add}))$
            Add fragment $y$ to molecule: $x \leftarrow x + y$
        else
            Sample $e$ from $p_{del}(e) = \sigma[\text{MLP}(h_{e_{ij}})]\ \forall e \in E$
            Let $y$ be the fragment designated by $e$, delete fragment $x \leftarrow x - y$
        if $C(x) = 1$ then
            $G \leftarrow G \cup \{x\}$
            $F \leftarrow I \cup G$
```

Molecular Fragments are extracted from molecules in the ChEMBL database [4]. For each molecule, we randomly sample fragments by extracting subgraphs that contain 10 or fewer atoms that have a single bond attachment to the rest of the molecule. We then use a VQ-VAE to encode these fragments into a meaningful latent space. The use of molecular fragments simplifies the search problem, while the variable-sized fragment distribution maintains the reachability of most molecular compounds. Because our search algorithm ultimately uses the latent representation of the molecules as the action space, we find that using a VQ-VAE with a categorical prior instead of the typical Gaussian prior makes RL training stable and provides performance gains.

Evaluation metrics. Following previous works, we evaluate our generative model on three target metrics, success, novelty and diversity. 5,000 molecules are generated by the model, and the metric scores are computed as follows: **Success** measures the proportion of generated molecules that fit the desired properties. For inhibition of GSK3β and JNK3, this is a score of at least 0.5 from the pretrained predictor. QED has a target score of $\geq .6$ and SA has a target score of ≤ 4. **Novelty** measures how different the generated molecules are compared to the set of actives in the dataset, and is the proportion of molecules whose fingerprint similarity is at most .4 to any molecule in the active set. **Diversity** measures how different the generated molecules are compared to each other. Here, diversity is computed as an average of pairwise fingerprint similarity across all generated compounds.

Baseline methods. **Rationale-RL** [8] extracts rationales of the active molecules and then uses RL to train a completion model that add atoms to the rationale in a sequential manner to generate molecules satisfying the desired properties. **GA+D & JANUS** [11, 12] are two genetic algorithms that use random mutations of SELFIES strings to generate promising molecular candidates; JANUS leverages a two-pronged approach, accounting for mutations towards both exploration and exploitation. **MARS** [19] uses Markov Chain Monte Carlo (MCMC) sampling to iteratively build new molecules by adding or removing fragments, and the model is trained to fit the distribution of the active molecules. We additionally include a baseline **MARS+Rationale** that initialize the MARS algorithm with the same starting initial rationale set used in Rationale-RL and our method in order to provide better comparisons of the methods. Where possible, we use the numbers from the original corresponding paper.