Symmetry-Constrained Generation of Diverse Low-Bandgap Molecules with Monte Carlo Tree Search

Akshay Subramanian
Massachusetts Institute of Technology
akshay_s@mit.edu

James Damewood
Massachusetts Institute of Technology

Juno Nam
Massachusetts Institute of Technology

Kevin P. Greenman
Massachusetts Institute of Technology

Avni P. Singhal
Massachusetts Institute of Technology

Rafael Gómez-Bombarelli
Massachusetts Institute of Technology
rafagb@mit.edu

Abstract

Organic optoelectronic materials are a promising avenue for next-generation electronic devices due to their solution processability, mechanical flexibility, and tunable electronic properties. In particular, near-infrared (NIR) sensitive molecules have unique applications in night-vision equipment and biomedical imaging. Molecular engineering has played a crucial role in developing non-fullerene acceptors (NFAs) such as the Y-series molecules, which have significantly improved the power conversion efficiency (PCE) of solar cells and enhanced spectral coverage in the NIR region. However, systematically designing molecules with targeted optoelectronic properties while ensuring synthetic accessibility remains a challenge. To address this, we leverage structural priors from domain-focused, patent-mined datasets of organic electronic molecules using a symmetry-aware fragment decomposition algorithm and a fragment-constrained Monte Carlo Tree Search (MCTS) generator. Our approach generates candidates that retain symmetry constraints from the patent dataset, while also exhibiting red-shifted absorption, as validated by TD-DFT calculations.

1 Introduction

Organic optoelectronic materials have emerged as promising candidates for next-generation electronic devices, particularly in the fields of organic photodiodes (OPDs) and organic photovoltaics (OPVs). These materials offer unique advantages such as solution processability, mechanical flexibility, and tunable electronic properties, making them attractive alternatives to traditional inorganic semiconductors [58, 37, 54]. Of particular interest are low-bandgap π-conjugated organic molecules and polymers sensitive to near-infrared (NIR) wavelengths, which can harvest a broader portion of the solar spectrum, extending beyond the visible range. These NIR-sensitive materials have found diverse applications, from military night-vision equipment to biomedical imaging devices and semi-transparent photovoltaics [65, 38, 36, 64]. The optimization of molecular structures, especially the design of improved donor–acceptor (D-A) systems, has led to substantial progress in the power conversion efficiency (PCE) of OPV devices. Over the past decade, PCE has increased from around 5% to over 19% for single-junction devices [42]. This advancement is largely attributed to the

development of non-fullerene acceptors (NFAs), which offer more tunability of energy levels and absorption spectra through molecular engineering, than fullerene acceptors (FAs) [66]. The Y-series family of NFAs (Y1 through Y6) has been a particularly promising class of molecules [67], which has sparked several follow-up designs and derivatives [24]. These improvements highlight the role of molecular design in enhancing device performance and the potential for further advancements through innovative approaches to materials discovery.

Over the past half-decade, generative models have received tremendous attention in molecular design [52, 14], using a variety of algorithms and molecular representations [53, 5, 21, 30, 48, 43, 27, 16]. However, without special attention to synthetic accessibility, off-the-shelf models often generate non-synthesizable designs [18]. Several existing approaches incorporate knowledge about synthesizability into generative models, either implicitly or explicitly.

Implicit approaches typically work by 1) biasing the objective function for property optimization with an additional synthesizability term, 2) applying a synthesizability filter post-hoc to generated candidates, or 3) constraining the training dataset using combinatorial enumerations of molecular fragments (for example, symmetric attachments as a proxy for synthetic accessibility). 1 and 2 often use approximate score functions such as SAscore [15], SCScore [10], or retrosynthesis tools like ASKCOS [11] or AiZynthFinder [20] to quantify synthesizability. In 1, multi-objective optimization is sensitive to the choice of the loss function's functional form, which if not done appropriately can lead to a compromise in other property objectives. In 2, a key assumption is that a sufficiently large portion of generated candidates are synthesizable, an assumption often proven false, as demonstrated in Figure S10 of Gao and Coley [18]. Additionally, the scoring functions themselves have specific definitions that make them more suitable for certain situations than others, limiting their universality. SAscore for instance is designed for biologically relevant molecules such as drug-like molecules, and might be less applicable for other chemical spaces such as organic electronics [15]. In 3, a downside is that the heuristics used to design the training dataset might not be reflected in the generations unless more explicit constraints are injected into the models.

Explicit methods of incorporating synthesizability on the other hand, do not suffer from limitations of scoring functions, or generations not satisfying prior constraints. They typically either embed pre-specified synthetic reaction templates [19, 23, 61, 59], or machine-learned reaction prediction algorithms [6, 7] into the generation grammar to constrain the scope of generation. While both reduce the possibility of chemically infeasible outcomes, reaction templates significantly limit the breadth of the search space, and reaction prediction algorithms could be prone to error propagation in multi-step synthetic routes [18]. This case of multi-step pathways is common in organic electronics systems [56]. Some recent works at the intersection of organic electronics and synthesizability include Kwak et al. [31], Li and Tabor [35], Westermayr et al. [63].

In this paper, we develop a simple, and more universally applicable approach to incorporate an approximate form of synthesizability into the generation process, and demonstrate its utility for organic electronics applications by generating low-bandgap molecules. Building on previous work on extracting datasets from published patents, our method is based on the fact that molecules reported in published patents are likely to be synthetically accessible, and when extracted based on domain-focused keyword queries made on the text, are also high-performance for domain-relevant properties [57]. The chemical fragments that make up a molecule, and the symmetry in their attachments contain key information about the precursors and synthetic pathways used to make the molecule. To utilize this information effectively during generation, we develop:

1. A fragment decomposition algorithm that extracts fragments and their reactive positions from patent-mined structures, preserving symmetry information.

2. A generation algorithm based on Monte Carlo Tree Search (MCTS) [28, 12, 60] that constrains the generated molecules to contain only the previously extracted fragments attached at their reactive positions, while satisfying symmetry requirements.

We evaluate our method on two chemical spaces of different sizes, a smaller, focused space of Y6-derivatives [67], and a much larger space of patent-mined fragment derivatives [57]. We employ Time-Dependent Density Functional Theory (TD-DFT) to obtain our "ground-truth" bandgaps for training and evaluation. Our generated candidates retain symmetry constraints from the patent dataset while also exhibiting red-shifted absorption.

2 Methods

2.1 Fragment Identity, Connectivity and Symmetry

Previous work by Subramanian et al. [57] developed an automated pipeline to mine USPTO patents and create domain-focused unsupervised datasets of molecular structures. These patent-derived molecules likely possess two important characteristics: high performance (justifying the investment in a patent application) and synthetic accessibility. Consequently, the fragments present in these structures, along with their reactive positions (points of connectivity), potentially correlate with both synthetic accessibility and domain-specific properties. Moreover, the symmetry of reactive positions relative to each other encodes information about the synthetic pathways used to create these molecules. To effectively leverage this structural and synthetic information, we have developed a two-stage method: **Fragment Decomposition:** We designed an algorithm that extracts fragments and their reactive positions from patent-mined structures, while marking each reactive position in a way that faithfully represents the identity of the substructure attached to it. This preserves important contextual information about molecular connectivity. It is this connectivity and identity-preserving property that we refer to as "symmetry" throughout this paper.

Constrained Generation: We implemented a generation algorithm based on MCTS that constrains the output to molecules composed solely of patent-extracted fragments. These fragments are attached at their corresponding reactive positions while adhering to the symmetry requirements observed in the original patent structures. This approach aims to maintain synthetic feasibility while exploring novel molecular configurations. By combining these two components, we aim to create a framework for generating molecules that are not only likely to be synthetically accessible but also tailored to the specific property requirements of organic electronic materials. This method leverages the information embedded in patent-derived structures to guide the exploration of chemical space towards promising and realizable candidates.

2.1.1 Fragment Decomposition

We developed a fragment decomposition algorithm that preserves connectivity and symmetry information from patent molecules. We achieve this through a recursive application of RDKit reaction SMARTS [32]. More details are provided in the following paragraphs, and an illustrative example is shown in Figure 1(b).

A reaction SMARTS string is first defined to indicate what types of bonds must be deleted during decomposition. For our purpose, we defined this so that single bonds connecting an aromatic ring with an aliphatic chain, or an aromatic ring to another aromatic ring are matched.
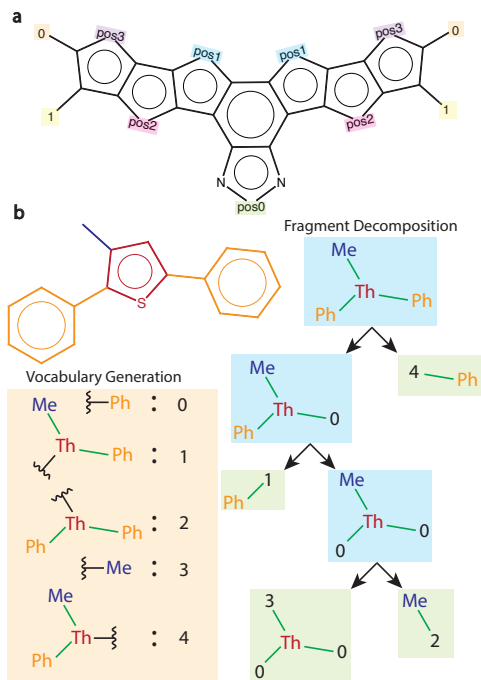

Figure 1: **Fragments to initialize MCTS** (a) Editable Y6 core with marked positions. (b) Fragment decomposition algorithm to obtain fragments from patent dataset. Vocabulary dictionary is first created by breaking one-bond at a time, and labeling the resulting fragments with unique integer values. Recursive decomposition of the starting molecule is then performed. Reactive positions are labeled with values corresponding to the broken fragment's identity. The leaf nodes (shaded in green) represent the final set of fragments obtained after decomposition. Me, Th, and Ph represent methyl, thiophene, and phenyl groups, respectively.

The algorithm consists of two steps: 1) creation of fragment vocabulary, 2) fragment decomposition and assignment of reactive sites.

In step 1, the reaction SMARTS is applied to the molecule to get all possible fragments that are obtainable from deletion of one bond in the molecule. The vocabulary of unique fragments is created, and each is assigned a unique ID. For the IDs, we use numbers $[0, N)$ where N is the number of fragments in the vocabulary. Details on how this is actually represented in our code implementation are given in SI Section A.5.

In step 2, the original molecule is recursively decomposed with one-bond deletions to form a binary tree. After every application of the reaction SMARTS, the reactive positions (positions at which bond deletion was performed) on the fragment are labeled using the ID of the fragment which was attached before the bond was broken. This preserves information about the identity of the fragment which was connected at the reactive position. The two fragments are then checked for a match with the reaction SMARTS to see if further decomposition is possible. If possible, the same series of steps described above is performed on the fragment. Otherwise, the fragment is a leaf node on the binary tree, and is added to the list of final fragments.

Finally, memoization [44] can be used to prevent redundant decompositions of the same fragment. We memoize computations in a hash table during tree creation, and reuse previously performed computations. Figure 1 illustrates this process of fragment decomposition on an example molecule.

2.1.2 Generation with MCTS

We frame the problem of designing a molecule as a series of building block/molecular fragment choices. MCTS is an efficient method of navigating large exploration spaces, that learns to optimize a chosen reward function by iteratively updating value estimates for action choices in the tree [28, 12, 60, 55]. Since our objective is to design low-bandgap molecules, our reward function of choice is the negative bandgap as predicted by Chemprop [25], a message-passing neural network (MPNN) architecture that we trained on labels obtained from TD-DFT calculations. While more details about TD-DFT calculations are provided in Section 2.4.1, we note here that our use of the term "bandgap" throughout this paper refers to the lowest-energy singlet vertical excitation energies. Each iteration of the MCTS algorithm consists of four sequential steps: selection, expansion, rollout, and backpropagation. The **selection step** starts with the root node and traverses the already explored portions of the tree to reach a leaf node. In our work, the path from the root node to the leaf node represents the construction of a molecular structure, where each node corresponds to a partial or complete structure. We follow the standard Upper Confidence Bound applied to Trees (UCT) policy [28] to choose an action at each timestep of traversal.

$$a_t = \operatorname*{argmax}_{a \in A_t} \left[Q_{t-1}(a) + c\sqrt{\frac{\log(N_{t-1}^p)}{N_{t-1}(a)}} \right]$$

where a_t is the action at timestep t, A_t is the set of action choices available at timestep t, $Q_{t-1}(a)$ is the action value function of action a before timestep t, $N_{t-1}(a)$ is the number of times the action a has been chosen before timestep t, and N_{t-1}^p is the number of times the parent node has been visited before timestep t. The exploration coefficient c controls the exploration-exploitation trade-off, i.e., the larger its value, the higher the exploration.

If the leaf node is not a terminal action, the **expansion step** enumerates possible next states and expands the tree with new nodes, one of which is then chosen at random. In the **rollout step**, a random policy is followed until the molecular generation process is terminated. Finally, the quality of the end state is evaluated and is assigned a scalar reward value. The reward is **backpropagated** through the tree branch to update the value functions of each visited node.

A detailed description of the Markov Decision Process defined for MCTS is given in Section 2.3. The Y6 case does not involve patent-mined fragments, and hence symmetry and reactive position considerations are greatly simplified. However, in the patent-mined fragments case, these considerations need to be incorporated into the generation grammar. Only fragments obtained from the fragment decomposition algorithm (described in Section 2.1.1) are utilized in the generation process, and fragment attachments are made only to reactive position markers. Identically marked positions are attached with fragments of identical identities, while differently marked positions are

attached with fragments of different identities, thereby preserving the connection symmetry as was present in the patent structure.
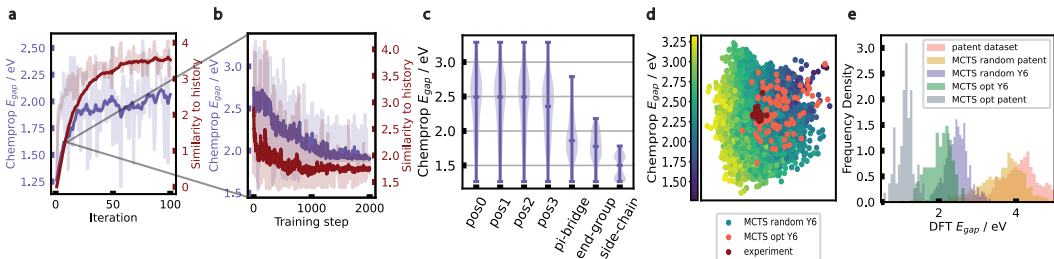

Figure 2: **Features of MCTS training and DFT validation** (a) Bandgap and similarity penalty as a function of iteration. Values shown are the weighted contributions to the reward function. With every iteration, the constrained optimization becomes more challenging resulting in less optimal candidates. (b) Representative reward evolution during training for one MCTS repetition. (c) Pruning of tree during training. The tree is expanded to contain deeper choices such as end-groups and side-chains only if their rewards are promising. This can be seen by shift towards lower predicted bandgaps as we traverse deeper down the tree. (d) PCA of randomly sampled and property-optimized molecules obtained from MCTS, and some popular experimental candidates. PCA is performed on Morgan fingerprints of molecules, and colors of random samples are based on Chemprop-predicted bandgaps. It can be seen that the random samples are diverse chemically and also span a range of property values, while the MCTS optimized candidates are diverse but concentrated at lower bandgap locations of the landscape. a, b, c, and d are plotted on Y6 MDP. (e) Histograms showing shifts in DFT histograms in comparison to training datasets. It can be seen that MCTS-optimized Y6 and patent molecules exhibit a large shift towards lower bandgaps compared to the patent data distribution.

2.2 Generating Chemically Diverse but Optimal Candidates

We aimed to generate a number of chemically diverse, yet also high performing candidates to increase the likelihood of at least a subset of candidates being experimentally synthesizable. In addition, Subramanian et al. [57] have observed empirically in their work that the proxy reward predictor has the tendency to be adversarially attacked during training in reinforcement learning settings because of its potential unreliability in out of domain regions. Proposing chemically diverse molecules that have high predicted rewards can allow for an increased chance that a subset of the proposals is in-domain for the reward predictor. The vanilla formulation of MCTS converges at a policy that maximizes expected cumulative reward [55, 60], and hence does not guarantee chemical diversity. To overcome this limitation, we iteratively re-train MCTS with a time-dependent reward term that penalizes chemical similarity to high-performing molecules from previous iterations. More formally,

$$R_i(x) = C(x) - \max_{y \in Y_{<i}} S(x, y) \qquad (1)$$

where R_i is the reward used for training MCTS at the i_{th} iteration, C is the property reward function, $Y_{<i}$ is the inventory of high-performing molecules collected from iterations $i' < i$, and S is a function computing similarity between molecules x and y (which we choose to be Tanimoto similarity). Figure 2(a) shows the evolution of the bandgap and similarity as a function of iteration. The point displayed at each iteration corresponds to the molecule with the lowest bandgap from the last 100 timesteps of MCTS training. We can see that the constrained optimization task becomes more challenging as more molecules are appended to the inventory $Y_{<i}$, resulting in an upward trend in both blue and red curves. Figure 2(b) zooms into a single iteration and shows the bandgap and similarity curves as a function of training step. In this case, training seems to converge after around 2,000 training steps.

2.3 Markov Decision Process

The Markov Decision Process (MDP) can be defined as a tuple of the state space $\mathcal{S}$, the action space $\mathcal{A}$, the reward R, and the transition probability P. We define MDPs for our two demonstrative tasks below.

5

2.3.1 Y6 derivatives

We formulated this to be a highly focused and relatively small-sized chemical space as a test bed for exploring the capabilities of the approach. Building a molecule involves four hierarchies of actions to be made: 1) modification of heteroatoms and ring sizes in the Y6 **core**, 2) optional addition of **pi-bridges** to extend conjugation in the molecule, 3) addition of terminal **end-groups**, and 4) addition of **alkyl chains**. Many of the fragment choices included here were taken directly from suggestions provided in Yang [67]. We provide a more formal description of the MDP below.

Let $\mathcal{S}_t$ and $\mathcal{A}_t$ denote the state and action spaces respectively at timestep t. Then, $\mathcal{S}_1$ is a singleton set containing the Y6 core marked with modifiable positions shown in Figure 1(a). The first group of actions (with action spaces $\mathcal{A}_{1 \leq t \leq 4}$) are those that modify the structure of the central Y6 **core** by allowing heteroatom and ring-expansion modifications at positions marked with pos0, pos1, pos2, and pos3.

Action spaces $\mathcal{A}_{5,6}$ allow the optional addition of **pi-bridges** to the positions marked by 0 to offer the flexibility to extend conjugation in the molecule, and action space $\mathcal{A}_7$ is the set of allowed **end-groups** that can be attached to 0 (if no pi-bridge was attached) or the newly created reactive position available on the pi-bridge. Finally, $\mathcal{A}_8$ is the set of allowed **alkyl chains** that can be attached to 1, which we restrict to just methyl groups to reduce the expense of conformer generation and optimization during DFT validation later on. This is a reasonable approximation to make given that alkyl chains are typically used to enhance solubility properties of the molecule and have a smaller influence on optoelectronic properties in isolated molecules. However, with this simplification, we disregard the more important role that alkyl chains play in optoelectronic properties when the molecules are in an aggregated state [33]. All fragments used in this MDP are shown in Figure 7. State space $\mathcal{S}_{t>1}$ is the set of possible molecular structures that can be reached by applying transformations in $\mathcal{A}_{t-1}$, to state s_{t-1}.

The reward R is the negative of the property value (bandgap in eV) as predicted by the Chemprop [25] model combined with the similarity penalty term as defined in Equation 1, and is computed once the molecular generation process has terminated. The MDP is deterministic and hence the transition probability is 1 for a particular state transition and 0 for all others at every timestep t, i.e., $P_t = (\delta_{i1}\delta_{i2} \ldots \delta_{in})$ where δ_{ij} is the Kronecker delta function and i is an integer ranging from 1 to n that represents the index of the chosen action. Only element δ_{ii} would be 1, and all other elements of the vector would be 0.

Figure 2(c) and (d) showcase some features of MCTS training on the Y6 MDP. (c) demonstrates the ability of MCTS to prune the tree, with deeper actions being expanded to only if they contribute to lowering the bandgap. (d) is a principal component analysis (PCA) plot showing the partitioning of molecules in the Chemprop embedding space according to Morgan fingerprints of molecules. The MCTS optimized generations are clearly shifted towards the lower bandgap region of the embedding landscape, while also maintaining chemical diversity. They can also be seen to be red-shifted with respect to some popular experimental NFA molecules (see Figure 8 for a list of the chosen experimental molecules). It is to be noted that both (c) and (d) are based on Chemprop predictions and not DFT calculations. Validation with DFT calculations is shown in later sections.

2.3.2 Patent-Extracted Fragment Derivatives

We formulated this to be a more expansive test case with a much larger chemical space, with the aim of discovering highly red-shifted molecules with diverse chemistries. In regimes of this size, exhaustive enumeration of all achievable molecules becomes impractical. Building a molecule in this environment involves three hierarchies of actions: 1) Choosing a **core** fragment, 2) Optional addition of **pi-bridges** to the core, and 3) Addition of terminal **end-groups/alkyl chains**.

In this case, each reaction involves not just the choice of a fragment identity, but also the choice of reactive position to which it has to be attached. The choice of reactive position is included as stochasticity in the environment rather than an action choice in the tree. More details on this choice, and a more formal description of the MDP are provided in the following paragraphs.

In this case, the breadth of the tree is very wide because of the large action space at each level of the tree. This makes it beneficial to trade breadth for depth of the tree. We achieve this by performing k-means clustering [40] of fragments at each hierarchy, so that chemically similar fragments occupy the same cluster. Fragment choices are decomposed into two sequential steps of choosing a cluster,

and choosing a fragment from the cluster. This approach retains the original grammar, but greatly improves efficiency of MCTS convergence. The value of k was chosen arbitrarily as 100, and is a user-defined parameter that can be adjusted. Additionally, we group end-groups and alkyl chains into a single decision here unlike the Y6 case since we do not differentiate between them in the fragment decomposition algorithm described in Section 2.1.1.

$\mathcal{S}_1$ is a singleton set containing an empty molecule, since we build the molecule from scratch in this case. Action space $\mathcal{A}_1$ is the set of cluster indices of central **core** clusters (obtained from k-means clustering). $\mathcal{A}_2$ is the set of fragment choices from the chosen cluster i, i.e., $\mathcal{C} = \{ f \in \mathcal{F} \mid 1 \leq r(f) \leq 4; n(f) \leq 35 \}_i$ where $\mathcal{F}$ is the set of fragments obtained with the fragment decomposition algorithm described in Section 2.1.1, $r(f)$ is the number of reactive positions in fragment f, $n(f)$ is the number of heavy atoms (non-hydrogen atoms) in fragment f, and $\mathcal{X}_i$ denotes the ith cluster obtained from k-means clustering of set $\mathcal{X}$. We placed restrictions on the number of reactive positions and size of fragments to reduce the computational load of DFT validation.

The next group of actions contains the optional choice of **pi-bridges** to allow for extended conjugation in the molecule. This again begins with the choice of a pi-bridge cluster index, followed by the choice of fragment identity from set $P = \{ f \in \mathcal{F} \mid r(f) = 2; n(f) \leq 10 \}_i$. We allow this pi-bridge sequence to be performed at most twice for a molecule (action spaces $\mathcal{A}_{3,5}$ are cluster choices and $\mathcal{A}_{4,6}$ are fragment identity choices). Finally, the last group of actions, with action spaces A_t for $t \in [7, 2k + 7)$, attaches end-groups to all open reactive positions on the molecule, where k is the number of unique reactive positions remaining in the molecule. This attachment is iterated as long as the molecule does not exceed the maximum allowed size of 100 heavy atoms. Note that a maximum of $2k$ steps is required as we alternate between the cluster and fragment choices. The fragment identity choice is given by set $\mathcal{E} = \{ f \in \mathcal{F} \mid r(f) = 1; n(f) \leq 20 \}_i$ where i is the index of the chosen cluster. State space $S_{t>1}$ is the set of possible molecular structures that can be reached by applying transformations in $\mathcal{A}_{t-1}$, to state s_{t-1}.

Similar to the Y6 case, reward R is the negative of Chemprop-predicted bandgap combined with a similarity penalty as described in Equation 1. Unlike the Y6 case however, the MDP is not deterministic here. The transitions that go from a state to a chosen fragment cluster are deterministic, but all remaining actions along the branch have some randomness associated with them. Once an action a_t corresponding to the identity of pi-bridge or end-group is chosen, the reactive position on the current state s_t is sampled from a uniform distribution, and the next state s_{t+1} is reached by reacting the fragment at the sampled reactive position on the current state. The transition probability is therefore defined as

$$P_t = \begin{cases} (\delta_{i1}\, \delta_{i2}\, \cdots\, \delta_{in}) & t \text{ is odd} \\ \left(\frac{1}{k_t}\, \frac{1}{k_t}\, \cdots\, \frac{1}{k_t} \right) & t \text{ is even} \end{cases}$$

k_t is the number of unique reactive positions available on state s_t. We chose to include this stochasticity in reactive position choice for two reasons: 1) The identity and nature of a chemical fragment typically has a higher influence on the optical properties of a molecule, than the reactive position choice does. Hence removing the reactive position choice from the action space allowed us to limit the depth of the tree and improve computational efficiency during training. 2) We wanted MCTS to converge at solutions that identify optimal fragment choices but are relatively robust to choice of reactive positions. While the precursors are determined by the fragments that are present in the molecule, there is more flexibility in the reactive position choice, and therefore during synthetic reaction planning.

2.4 Training Reward Predictor

2.4.1 TD-DFT Calculations

TD-DFT calculations were used as "ground-truth" labels to train our Chemprop reward predictor, as well as to validate the bandgaps of generated candidates. Calculations were performed with the same pipeline that we used in Subramanian et al. [57]. RDKit ETKDG [32] approach was used to generate initial conformations, with at least 1,500 attempts. The conformers were ranked according to their MMFF94 energies, and the 20 with the lowest energies were chosen [50]. An initial geometry optimization was performed with semi-empirical tight-binding density functional theory (GFN2-xTB) [2] in ORCA 4.2.0 [46]. Subsequent geometry optimizations were performed on the lowest
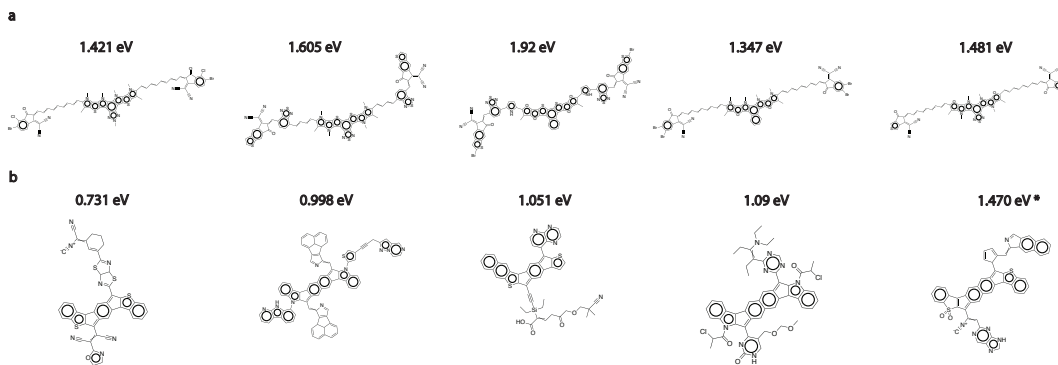
Figure 3: **Final generated candidates and their TD-DFT bandgaps** K-means clustering was performed on final 100 molecules from each category into 5 clusters (based on Morgan fingerprints). The molecules shown are the lowest-bandgap molecules chosen from each cluster. (a) Y6-derivatives MDP, (b) Patent-extracted fragments MDP. Bandgap of the last molecule is denoted with asterisk (*) because the geometry had to be fixed before the TD-DFT calculation was performed. More details are provided in Section A.4 of SI.

energy xTB conformers at the BP86[3]-D3[22]/def2-SVP[62] level of theory. Finally, single-point TD-DFT calculations were performed at the ωB97X-D3[9]/def2-SVPD level of theory using the Tamm–Dancoff approximation (TDA) [26] on the optimized geometries. Bandgaps used for reward training were the lowest-energy singlet vertical excitation energies from the TD-DFT calculations.
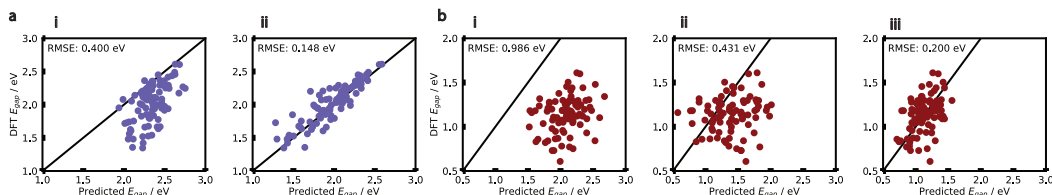

Figure 4: **Active Learning to improve reward prediction** Scatter plots showing improvement in fit with AL iterations for (a) Y6 derivatives (b) Patent-extracted fragment derivatives. The test data in (a) and (b) are the final sets of 100 molecules generated after all AL iterations have been completed with the Y6 and patent-fragment MDPs respectively. In (a) and (b), the plot i corresponds to the model pre-trained on just patent-mined dataset, and plot ii corresponds to model trained on patent dataset + random rollouts from MCTS. Plot iii in (b) corresponds to model trained on patent dataset + random rollouts + diversity & EI acquisition samples. More details are given in Section 2.4.2.

2.4.2 Active Learning for Data Collection

To ensure that the reward predictor is accurate in the generation domain of MCTS, we performed iterations of active learning (AL) with a combination of acquisition functions to collect training data. Figure 4 shows the improvement in reward predictor with each iteration of AL. Chemprop [25] was our architecture of choice for the reward predictor. It was first pre-trained on a dataset containing 5,568 datapoints obtained from patents and labeled with TD-DFT bandgaps which we created in Subramanian et al. [57]. The second round of training was performed on random rollouts generated from MCTS. These two rounds were sufficient to get accurate (we chose the stopping criterion to be an RMSE of 0.2 eV) performance on the focused Y6 chemical space, but the patent-extracted fragment space required an additional round of training. This was done on a combination of datapoints obtained from 1) running the iterative procedure described in Section 2.2 and choosing the best candidate from the last 100 training steps of each iteration, and 2) choosing 100 molecules with the

highest value of expected improvement (EI) from 10,000 randomly sampled candidates from MCTS.

$$EI(x) = (\mu - f(x^\star)) \, \Phi \left(\frac{\mu - f(x^\star)}{\sigma} \right) + \sigma \phi \left(\frac{\mu - f(x^\star)}{\sigma} \right)$$

where μ is the negative of the Chemprop-predicted bandgap, and σ is the standard deviation of predictions from the model ensemble, which is a measure of model uncertainty. Φ and ϕ are the cumulative distribution function (CDF) and probability density function (PDF) of a normal distribution, respectively. $f(x^\star)$ is the highest value of predicted reward from the current batch of candidates. While acquisition method 1 obtains a diverse set of most optimal candidates, method 2 obtains a set of optimal candidates as well as suboptimal but high-uncertainty candidates. Together, these two acquisition methods provide training data in explorative and exploitative regimes.

3   Results and Discussion

We validated the final 100 diverse candidates obtained from each of Y6 and patent MDPs with TD-DFT bandgap calculations. We show histograms of these bandgaps in Figure 2 (e) in comparison to the original patent dataset, and random rollouts sampled from MCTS. It can be seen that the MCTS optimized patent and Y6 candidates are strongly red-shifted with respect to random rollouts as well as the patent distribution.

On the final 100 structures from each MDP, we performed k-means clustering with $k = 5$ based on Morgan fingerprints of molecules to obtain the 5 most chemically diverse clusters of molecules. We then chose the molecule with the lowest TD-DFT bandgap from each cluster. These structures along with their bandgaps are displayed in Figure 3. It can be seen that these molecules display structural features that are to be expected from low-bandgap molecules, such as highly conjugated rings. Furthermore, the predicted bandgaps all fall in the NIR region of the spectrum.

In the case of Y6 derivatives, it is interesting to see a higher tendency of MCTS to generate long polyene bridges between cores and end-groups (see the first, fourth, and fifth molecules in subfigure (a)). Clearly, it has exploited the correlation between degree of conjugation and lowering of bandgap. However, the presence of conical intersections (CI) in long linear polyenes is well known [17, 47] suggesting a possibility for these MCTS-predicted molecules to be photochemically inactive. Linear response TD-DFT has been seen to be unsuitable for accurately predicting the presence of these points, paving way for the development of more sophisticated techniques such as spin-flip TD-DFT (SF-TD-DFT) [34, 45]. This is hence a drawback with the complete reliance of MCTS on the accuracy of TD-DFT predictions, but nevertheless showcases the ability of the algorithm to reliably optimize the chosen reward.

Design strategies such as spatially separating donating and accepting fragments for a "push-pull" effect are also known to have a lowering effect on bandgaps [49]. Among the generated patent-derivatives, this design pattern can be seen in the first and fourth molecules of subfigure (b) which contain electron-rich cores and electron-withdrawing end-groups. We additionally observe the presence of highly reactive sites in some of these structures (such as hydrogen atoms on aromatic amines of second, fourth, and fifth molecules, and chlorine atoms on the fourth molecule). We replace the reactive positions with methyl groups and perform TD-DFT calculations on the "corrected" structures to verify that optical properties do not drastically change even after the substitution. These results can be found in Section A.4 of SI.

Finally, the TD-DFT bandgap calculations that we report work under the assumption of 0 K temperature. Furthermore, a more realistic device-like environment would contain thin-films containing several molecules packed together. We perform coupled MD/TD-DFT calculations on the final 10 molecules to better represent finite temperature and molecular aggregation effects. We report absorption spectra calculated with these methods, and more details about this methodology in section A.3 of the SI.

4   Conclusions

In this paper, we developed an approach based on the Monte Carlo Tree Search algorithm to generate optimal low-bandgap molecules while also maintaining a degree of synthetic accessibility. We achieve this by utilizing structural priors from a domain-focused patent-mined dataset of organic electronics

molecules through a symmetry-aware fragment decomposition algorithm and a fragment-constrained MCTS generator. Our generated candidates retain symmetry constraints from the patent dataset while also exhibiting red-shifted absorption. As future extensions to this work, the optimization objective can be augmented to include solubility (along with the inclusion of alkyl chain exploration into the MDP) and oscillator strengths. Finally, while our approach for diverse candidate generation is iterative, non-iterative objectives as used in GFlowNet [4] can be more efficient alternatives for the future.

References

[1] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[2] C. Bannwarth, S. Ehlert, and S. Grimme. GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation*, 15(3):1652–1671, mar 2019. ISSN 15499626. doi: 10.1021/acs.jctc.8b01176.

[3] Becke and A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098–3100, sep 1988. ISSN 0556-2791. doi: 10.1103/PhysRevA.38.3098. URL http://www.ncbi.nlm.nih.gov/pubmed/9900728https://link.aps.org/doi/10.1103/PhysRevA.38.3098.

[4] Y. Bengio, S. Lahlou, T. Deleu, E. J. Hu, M. Tiwari, and E. Bengio. Gflownet foundations. *The Journal of Machine Learning Research*, 24(1):10006–10060, 2023.

[5] E. J. Bjerrum and R. Threlfall. Molecular generation with recurrent neural networks (RNNs). *arXiv preprint arXiv:1705.04612*, 2017.

[6] J. Bradshaw, B. Paige, M. J. Kusner, M. Segler, and J. M. Hernández-Lobato. A model to search for synthesizable molecules. *Advances in Neural Information Processing Systems*, 32, 2019.

[7] J. Bradshaw, B. Paige, M. J. Kusner, M. Segler, and J. M. Hernández-Lobato. Barking up the right tree: an approach to search over molecule synthesis dags. *Advances in neural information processing systems*, 33:6852–6866, 2020.

[8] C. M. Breneman and K. B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, 11(3):361–373, 1990.

[9] J.-D. Chai and M. Head-Gordon. Long-range corrected double-hybrid density functionals. *The Journal of chemical physics*, 131(17):174105, 2009.

[10] C. W. Coley, L. Rogers, W. H. Green, and K. F. Jensen. SCScore: synthetic complexity learned from a reaction corpus. *Journal of chemical information and modeling*, 58(2):252–261, 2018.

[11] C. W. Coley, D. A. Thomas III, J. A. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao, et al. A robotic platform for flow synthesis of organic compounds informed by ai planning. *Science*, 365(6453):eaax1566, 2019.

[12] R. Coulom. Efficient selectivity and backup operators in monte-carlo tree search. In *International conference on computers and games*, pages 72–83. Springer, 2006.

[13] L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives, and W. L. Jorgensen. Ligpargen web server: an automatic opls-aa parameter generator for organic ligands. *Nucleic acids research*, 45(W1): W331–W336, 2017.

[14] D. C. Elton, Z. Boukouvalas, M. D. Fuge, and P. W. Chung. Deep learning for molecular design—a review of the state of the art. *Molecular Systems Design & Engineering*, 4(4): 828–849, 2019.

[15] P. Ertl and A. Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of cheminformatics*, 1: 1–11, 2009.

[16] D. Flam-Shepherd, T. C. Wu, and A. Aspuru-Guzik. MPGVAE: improved generation of small organic molecules using message passing neural nets. *Machine Learning: Science and Technology*, 2(4):045010, 2021.

[17] W. Fuss, Y. Haas, and S. Zilberg. Twin states and conical intersections in linear polyenes. *Chemical Physics*, 259(2-3):273–295, 2000.

[18] W. Gao and C. W. Coley. The synthesizability of molecules proposed by generative models. *Journal of chemical information and modeling*, 60(12):5714–5723, 2020.

[19] W. Gao, R. Mercado, and C. W. Coley. Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. *arXiv preprint arXiv:2110.06389*, 2021.

[20] S. Genheden, A. Thakkar, V. Chadimová, J.-L. Reymond, O. Engkvist, and E. Bjerrum. Aizynthfinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of cheminformatics*, 12(1):70, 2020.

[21] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2):268–276, 2018.

[22] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, may 2011. ISSN 01928651. doi: 10.1002/jcc.21759. URL http://doi.wiley.com/10.1002/jcc.21759.

[23] M. Hartenfeller, H. Zettl, M. Walter, M. Rupp, F. Reisen, E. Proschak, S. Weggen, H. Stark, and G. Schneider. Dogs: reaction-driven de novo design of bioactive compounds. *PLoS computational biology*, 8(2):e1002380, 2012.

[24] Q. He, P. Ufimkin, F. Aniés, X. Hu, P. Kafourou, M. Rimmele, C. L. Rapley, and B. Ding. Molecular engineering of y-series acceptors for nonfullerene organic solar cells. *SusMat*, 2(5): 591–606, 2022.

[25] E. Heid, K. P. Greenman, Y. Chung, S.-C. Li, D. E. Graff, F. H. Vermeire, H. Wu, W. H. Green, and C. J. McGill. Chemprop: A machine learning package for chemical property prediction. *Journal of Chemical Information and Modeling*, 64(1):9–17, 2023.

[26] S. Hirata and M. Head-Gordon. Time-dependent density functional theory within the tamm–dancoff approximation. *Chemical Physics Letters*, 314(3-4):291–299, 1999.

[27] W. Jin, R. Barzilay, and T. Jaakkola. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, pages 2323–2332. PMLR, 2018.

[28] L. Kocsis and C. Szepesvári. Bandit based monte-carlo planning. In *European conference on machine learning*, pages 282–293. Springer, 2006.

[29] G. Kupgan, X. Chen, and J.-L. Bredas. Molecular packing of non-fullerene acceptors for organic solar cells: distinctive local morphology in y6 vs. itic derivatives. *Materials Today Advances*, 11:100154, 2021.

[30] M. J. Kusner, B. Paige, and J. M. Hernández-Lobato. Grammar variational autoencoder. In *International conference on machine learning*, pages 1945–1954. PMLR, 2017.

[31] H. S. Kwak, Y. An, D. J. Giesen, T. F. Hughes, C. T. Brown, K. Leswing, H. Abroshan, and M. D. Halls. Design of organic electronic materials with a goal-directed generative model powered by deep neural networks and high-throughput molecular simulations. *Frontiers in Chemistry*, 9:800370, 2022.

[32] G. Landrum et al. RDKit: A software suite for cheminformatics, computational chemistry, and predictive modeling. *Greg Landrum*, 2013.

[33] T. Lee, C. E. Song, S. K. Lee, W. S. Shin, and E. Lim. Alkyl-side-chain engineering of nonfused nonfullerene acceptors with simultaneously improved material solubility and device performance for organic solar cells. *ACS omega*, 6(7):4562–4573, 2021.

[34] B. G. Levine, C. Ko, J. Quenneville, and T. J. MartÍnez. Conical intersections and double excitations in time-dependent density functional theory. *Molecular Physics*, 104(5-7):1039–1051, 2006.

[35] C.-H. Li and D. P. Tabor. Generative organic electronic molecular design informed by quantum chemistry. *Chemical Science*, 14(40):11045–11055, 2023.

[36] Y. Li, J. D. Lin, X. Che, Y. Qu, F. Liu, L. S. Liao, and S. R. Forrest. High efficiency near-infrared and semitransparent non-fullerene acceptor organic photovoltaic cells. *Journal of the American Chemical Society*, 139:17114–17119, 11 2017. ISSN 15205126. doi: 10.1021/JACS. 7B11278/SUPPL_FILE/JA7B11278_SI_001.PDF. URL https://pubs.acs.org/doi/abs/ 10.1021/jacs.7b11278.

[37] Y. Li, W. Huang, D. Zhao, L. Wang, Z. Jiao, Q. Huang, P. Wang, M. Sun, and G. Yuan. Recent Progress in Organic Solar Cells: A Review on Materials from Acceptor to Donor. *Molecules*, 27(6):1800, 2022.

[38] F. Liu, Z. Zhou, C. Zhang, J. Zhang, Q. Hu, T. Vergote, F. Liu, T. P. Russell, and X. Zhu. Efficient semitransparent solar cells with high nir responsiveness enabled by a small-bandgap electron acceptor. *Advanced Materials*, 29:1606574, 6 2017. ISSN 1521-4095. doi: 10.1002/ADMA.201606574. URL https://onlinelibrary.wiley.com/doi/full/10. 1002/adma.201606574https://onlinelibrary.wiley.com/doi/abs/10.1002/adma. 201606574https://onlinelibrary.wiley.com/doi/10.1002/adma.201606574.

[39] B. Lu, J. Wang, Z. Zhang, J. Wang, X. Yuan, Y. Ding, Y. Wang, and Y. Yao. Recent progress of Y-series electron acceptors for organic solar cells. *Nano Select*, 2(11):2029–2039, 2021.

[40] J. MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.

[41] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez. Packmol: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164, 2009.

[42] D. Meng, R. Zheng, Y. Zhao, E. Zhang, L. Dou, and Y. Yang. Near-infrared materials: The turning point of organic photovoltaics. *Advanced Materials*, 34: 2107330, 3 2022. ISSN 1521-4095. doi: 10.1002/ADMA.202107330. URL https://onlinelibrary.wiley.com/doi/full/10.1002/adma.202107330https: //onlinelibrary.wiley.com/doi/abs/10.1002/adma.202107330https:// onlinelibrary.wiley.com/doi/10.1002/adma.202107330.

[43] R. Mercado, T. Rastemo, E. Lindelöf, G. Klambauer, O. Engkvist, H. Chen, and E. J. Bjerrum. Graph networks for molecular design. *Machine Learning: Science and Technology*, 2(2): 025023, 2021.

[44] D. Michie. "memo" functions and machine learning. *Nature*, 218(5136):19–22, 1968.

[45] N. Minezawa and M. S. Gordon. Optimizing conical intersections by spin- flip density functional theory: application to ethylene. *The Journal of Physical Chemistry A*, 113(46):12749–12753, 2009.

[46] F. Neese, F. Wennmohs, U. Becker, and C. Riplinger. The ORCA quantum chemistry program package. *The Journal of chemical physics*, 152(22):224108, 2020. ISSN 10897690. doi: 10.1063/5.0004608. URL https://doi.org/10.1063/5.0004608.

[47] A. Nenov and R. de Vivie-Riedle. Conical intersection seams in polyenes derived from their chemical composition. *The Journal of Chemical Physics*, 137(7), 2012.

[48] M. Olivecrona, T. Blaschke, O. Engkvist, and H. Chen. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9(1):1–14, 2017.

[49] G. Pirotte, P. Verstappen, D. Vanderzande, and W. Maes. On the "true" structure of push–pull-type low-bandgap polymers for organic electronics. *Advanced Electronic Materials*, 4(10): 1700481, 2018.

[50] S. Riniker and G. A. Landrum. Better informed distance geometry: using what we know to improve conformation generation. *Journal of chemical information and modeling*, 55(12): 2562–2574, 2015.

[51] M. J. Robertson, J. Tirado-Rives, and W. L. Jorgensen. Improved peptide and protein torsional energetics with the opls-aa force field. *Journal of chemical theory and computation*, 11(7): 3499–3509, 2015.

[52] D. Schwalbe-Koda and R. Gómez-Bombarelli. Generative models for automatic chemical design. In *Machine Learning Meets Quantum Physics*, pages 445–467. Springer, 2020.

[53] M. H. Segler, T. Kogej, C. Tyrchan, and M. P. Waller. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS central science*, 4(1):120–131, 2018.

[54] T. Shan, X. Hou, X. Yin, and X. Guo. Organic photodiodes: device engineering and applications. *Frontiers of Optoelectronics*, 15(1):49, 2022.

[55] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of Go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[56] F. Strieth-Kalthoff, H. Hao, V. Rathore, J. Derasp, T. Gaudin, N. H. Angello, M. Seifrid, E. Trushina, M. Guy, J. Liu, et al. Delocalized, asynchronous, closed-loop discovery of organic laser emitters. *Science*, 384(6697):eadk9227, 2024.

[57] A. Subramanian, K. P. Greenman, A. Gervaix, T. Yang, and R. Gómez-Bombarelli. Automated patent extraction powers generative modeling in focused chemical spaces. *Digital Discovery*, 2 (4):1006–1015, 2023.

[58] L. Sun, K. Fukuda, and T. Someya. Recent progress in solution-processed flexible organic photovoltaics. *npj Flexible Electronics*, 6(1):89, 2022.

[59] K. Swanson, G. Liu, D. B. Catacutan, A. Arnold, J. Zou, and J. M. Stokes. Generative ai for designing and validating easily synthesizable and structurally novel antibiotics. *Nature Machine Intelligence*, 6(3):338–353, 2024.

[60] M. Świechowski, K. Godlewski, B. Sawicki, and J. Mańdziuk. Monte carlo tree search: A review of recent modifications and applications. *Artificial Intelligence Review*, 56(3):2497–2562, 2023.

[61] H. M. Vinkers, M. R. de Jonge, F. F. Daeyaert, J. Heeres, L. M. Koymans, J. H. van Lenthe, P. J. Lewi, H. Timmerman, K. Van Aken, and P. A. Janssen. Synopsis: synthesize and optimize system in silico. *Journal of medicinal chemistry*, 46(13):2765–2773, 2003.

[62] F. Weigend and R. Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, sep 2005. ISSN 14639076. doi: 10.1039/ b508541a. URL http://xlink.rsc.org/?DOI=b508541a.

[63] J. Westermayr, J. Gilkes, R. Barrett, and R. J. Maurer. High-throughput property-driven generative design of functional organic molecules. *Nature Computational Science*, pages 1–10, 2023.

[64] H. Xu, J. Liu, J. Zhang, G. Zhou, N. Luo, and N. Zhao. Flexible organic/inorganic hybrid near-infrared photoplethysmogram sensor for cardiovascular monitoring. *Advanced Materials*, 29:1700975, 8 2017. ISSN 1521-4095. doi: 10.1002/ADMA.201700975. URL `https://onlinelibrary.wiley.com/doi/full/10.1002/adma.201700975https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201700975https://onlinelibrary.wiley.com/doi/10.1002/adma.201700975`.

[65] X. Xu, M. Davanco, X. Qi, and S. R. Forrest. Direct transfer patterning on three dimensionally deformed surfaces at micrometer resolutions and its application to hemispherical focal plane detector arrays. *Organic Electronics*, 9:1122–1127, 12 2008. ISSN 1566-1199. doi: 10.1016/J.ORGEL.2008.07.011.

[66] C. Yan, S. Barlow, Z. Wang, H. Yan, A. K.-Y. Jen, S. R. Marder, and X. Zhan. Non-fullerene acceptors for organic solar cells. *Nature Reviews Materials*, 3(3):1–19, 2018.

[67] Y. Yang. The original design principles of the y-series nonfullerene acceptors, from y1 to y6. *ACS nano*, 15(12):18679–18682, 2021.

A   Appendix / supplemental material

A.1   Preprocessing of Patent-Mined Fragments

After fragment decomposition was performed, we performed some minimal preprocessing steps to ensure that the fragments are representative of the chemical space we want to explore. We first performed an RDKit filter, to ensure that only SMILES strings representing valid molecules are retained, followed by an element filter, which retained only fragments containing the following elements: C, O, N, H, Cl, Br, S, F, I, and Si. We finally removed fragments that do not contain any aromatic atoms.

A.2   Correlation Between DFT and Experimental Bandgaps

We performed TD-DFT calculations on a subset of experimental molecules from the list shown in Figure 7 for which we had access to experimentally measured bandgaps. We see R^2 value of 0.57, and a Spearman rank correlation coefficient (SRCC) of 0.74. While the magnitudes are not well calibrated, the trends of experiments are well-captured by TD-DFT, making minimization a suitable objective for MCTS as opposed to optimization for targeted bandgap values. As a future extension of this work, it could be possible to fit a calibration function between TD-DFT and experimental bandgaps and use the calibrated rewards for training. While it could be somewhat reasonable to extrapolate such calibrations for the Y6 MDP, it might be more challenging to obtain representative experimental examples for the patents MDP on which a calibration fit can be performed.

The mismatch in magnitudes between TD-DFT and experiments can be caused by several factors including (but not limited to) 1) finite temperature effects, 2) molecular aggregation effects from experimental values being measured on thin films, 3) inherent error in TD-DFT functional.

A.3   Finite Temperature and Aggregation Effects with Coupled MD/TD-DFT Calculations

A.3.1   Computational Methods

We combined MD simulations with TD-DFT calculations to approximate the effects of finite temperature and molecular aggregation on absorption spectra. Following Kupgan et al. [29], we simulated an amorphous morphology for the molecules. Starting with the optimized geometries presented in the main text, we used PACKMOL [41] to pack 200 structures in a cubic box at a low density ($\sim$ 0.1 g/cm^3). The OPLS-AA force field [51] was used via the LigParGen server [13]. The system was equilibrated for 30 ns at 650 K, cooled to room temperature (300 K) at 10 K/ns, and subjected to a 30 ns production run at 300 K. All MD simulations were performed under the NPT ensemble at 1 atm, using GROMACS 2023 [1]. From the last 20 ns of the production run, we randomly selected a molecule and included all neighboring molecules within a 5 Å radius. CHELPG charges [8] were computed for each neighboring molecule, and TD-DFT calculations were performed for the selected molecule with the neighboring point charges. For electronic structure calculations, we used the
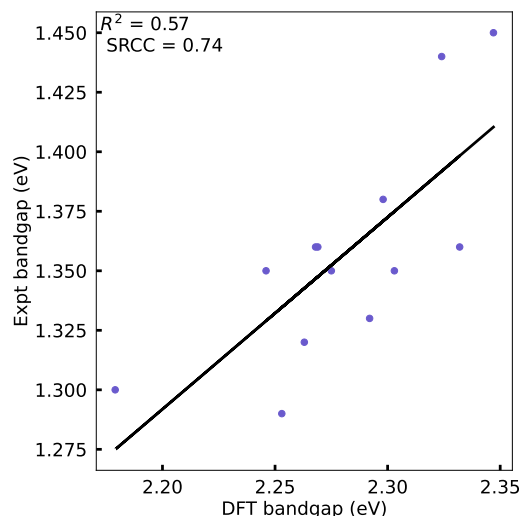
Figure 5: **Correlation between TD-DFT and experimental bandgaps** R^2 score and Spearman rank correlation coefficient are shown on top left.

same settings as in the main text. The absorption spectra were averaged over 10 clusters to obtain a statistical estimate.

Two of the patent-derived molecules (the first and fifth molecules in Figure 3) contain an isocyanide moiety, which could not be parametrized correctly using the LigParGen server. Therefore, for these two molecules, we substituted the isocyanide (–NC) functional group with cyanide (–CN) for this analysis. This substitution does not impact the band gap calculations significantly, with differences in bandgaps for the optimized geometries of only 2 and 6 meV, respectively, indicating negligible effects on the result analysis.

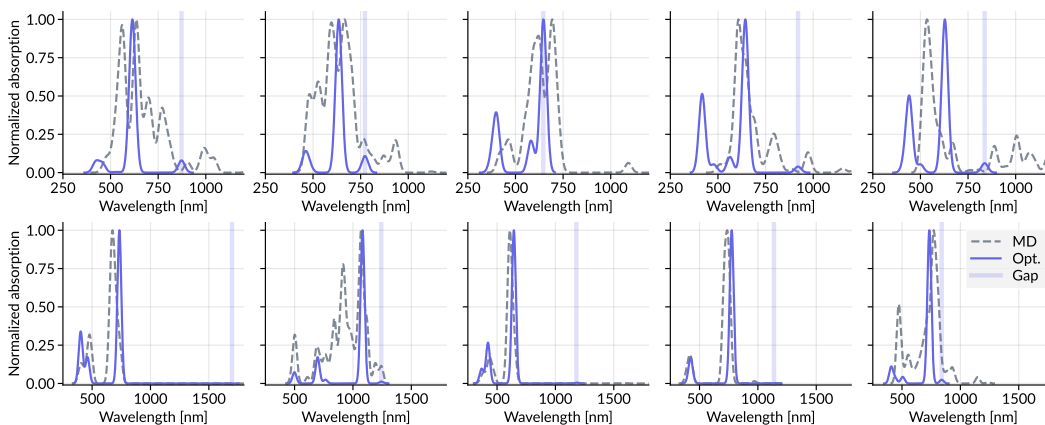### A.3.2 Absorption Spectra of Final Candidates

Figure 6: **Computed absorption spectra for final candidate molecules.** The spectra for molecules are ordered as shown in Figure 3 in the main text, with Y6 derivatives in the upper row and patent-extracted fragment designs in the lower row. Solid lines represent single-point TD-DFT calculations at the optimized geometry, while dotted lines represent statistically averaged TD-DFT spectra from molecular clusters sampled from MD simulations. The translucent vertical lines indicate the band gaps from the single-point calculations. The spectra are normalized so that the maximum absorption corresponds to 1.0.

The absorption spectra for the designed molecules are presented in Figure 6, with the Y6 derivatives in the upper row and patent derivatives in the lower row. The vertical lines indicate the TD-DFT band gap for the optimized geometries, as reported in Figure 3 of the main text. The MD/TD-DFT pipeline results for the primary absorption peak locations are consistent with those from static (optimized) TD-DFT calculations, though the MD/TD-DFT spectra exhibit broadening, likely due to thermally accessible conformers and neighboring molecules. Note that the reddest absorption peak from TD-DFT corresponding to the band gap might not accurately reflect the absorption spectra when its oscillator strength is very low. This limitation, inherent to the design objective focused solely on the band gap, indicates the need to consider transition probability in future designs. Nonetheless, the band gap serves as a robust design objective, and this result demonstrates that our MCTS design pipeline effectively produces low-bandgap molecules.

A.4 Structural Modifications

For the fifth molecule shown in Figure 3, we identified that errors in atomic connectivity were introduced during the xTB optimization step. We fixed the geometry by bypassing the xTB optimization step altogether. We instead directly performed BP86 DFT optimization on the five lowest energy conformers obtained from RDKit ETKDG algorithm, and performed the TD-DFT calculation on the lowest energy BP86 conformer.

As mentioned in Section 3, we performed TD-DFT calculations on methyl-substituted versions of reactive molecules to confirm that the bandgap does not change significantly after substitution. Bandgaps of second, fourth, and fifth molecules changed very minimally from 0.998, 1.09, and 1.470, to 0.978, 1.049, and 1.470 eV respectively.

A.5 Reactive Positions in Practice

While we explain the algorithm in Section 2.1.1 using IDs belonging to $[0, N)$, in the implementation, we use Helium isotopes with IDs ranging from $[100, 100 + N)$ where N is the number of fragments in the vocabulary, to preserve reactive position identities in the RDKit `Mol` object. The range was chosen arbitrarily but to be large enough so that there are no clashes in atomic mass with other elements in the dataset.

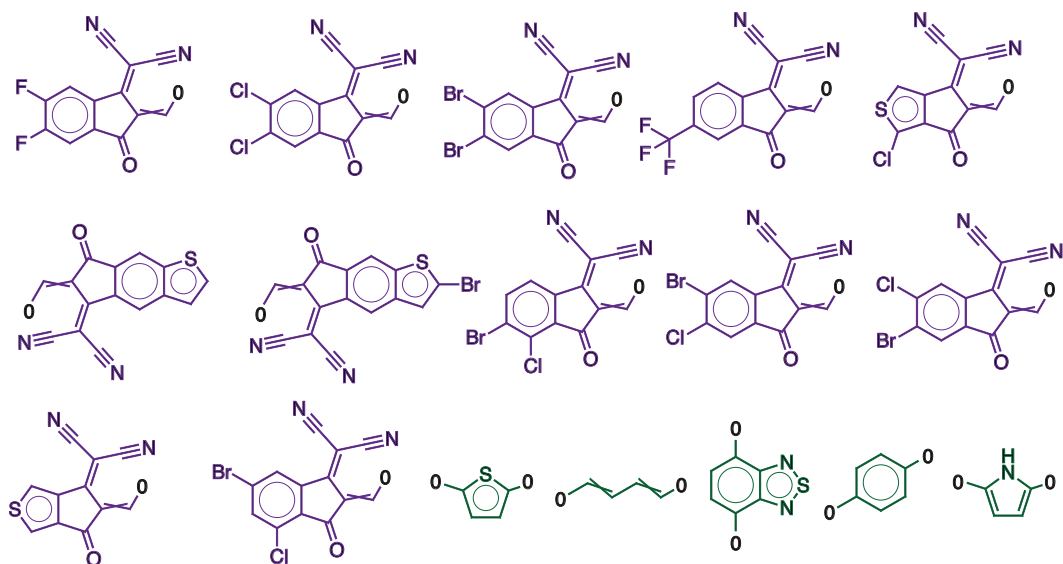## A.6 Fragments in Y6 MDP and Experimental Comparisons

Figure 7: **Fragments used in Y6 MDP** end-groups (purple) and pi-bridges (green). Core-modification actions are implemented as string replacements, the list of which can be found in the GitHub repository. We also omit illustrating alkyl chain actions here since methyl group is the only choice explored.
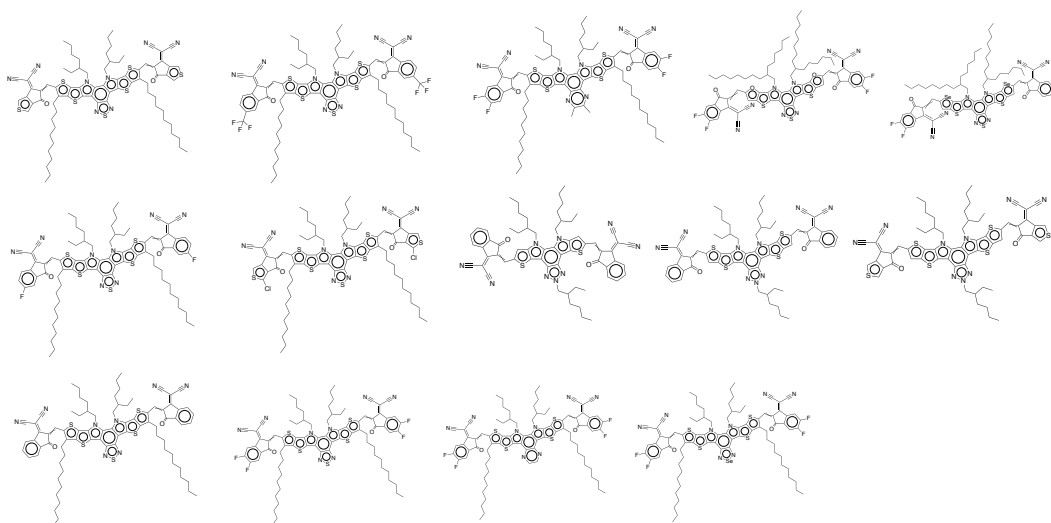

Figure 8: **Some popular experimentally used acceptor molecules**. These were chosen from Lu et al. [39].