Unified Fragment Growing, Linking, and Merging with Energy-Aware Pretraining and Reward-Ranked Alignment

Xuefeng Liu[1,2,†,*], **Songhao Jiang**[2*], **Tinson Xu**[3*], **Qinan Huang**[4], **Rick Stevens**[2,5]
[1]College of Medicine and AI for Health Institute, University of Florida
[2]Department of Computer Science, University of Chicago
[3]Laboratory Schools, University of Chicago
[4]Pritzker School of Molecular Engineering, University of Chicago
[5]Argonne National Laboratory

Abstract

Fragment-based drug discovery optimizes small molecular hits through three recurring transformations: fragment growing, linking, and merging. Existing generative methods, however, usually specialize in a single transformation and have focused primarily on linker design. We introduce FRAGMENTGPT, a unified chemical language model that casts all three operations as conditional molecular generation. To build this capability, we construct a chemically-aware pre-training corpus using energy-guided bond cleavage, together with overlap-based examples that teach the model to connect separated fragments and consolidate shared substructures. We further introduce Reward-Ranked Alignment with Expert Exploration (RAE), which combines expert-guided candidate generation, Pareto- and composite-reward-based selection, and supervised fine-tuning for multi-objective molecular optimization. Conditioned on one or more fragments, FRAGMENTGPT can expand a scaffold, connect distinct subunits, or merge overlapping structures into a complete molecule. Experiments on RTCB and SARS-CoV-2 main protease (3CLPro), together with ablation studies, show high chemical validity and strong aggregate performance across multiple drug-design objectives.

1 Introduction

Fragment-based drug discovery (FBDD) provides an efficient route for exploring chemical space by starting from small, low-molecular-weight compounds that can probe difficult binding surfaces, including protein–protein interfaces and shallow or cryptic pockets (Xu and Kang, 2025; Murray and Rees, 2009). In contrast to high-throughput screening of much larger libraries of more complex compounds (Leach and Hann, 2011), fragment libraries can be screened with sensitive biophysical techniques such as NMR, X-ray crystallography, SPR, and thermal-shift assays. FBDD has contributed to eight FDA-approved drugs and more than 50 clinical candidates, including vemurafenib, venetoclax, and sotorasib (Bollag et al., 2012; Fairbrother et al., 2019; Lanman et al., 2020; Xu and Kang, 2025). Because initial fragment hits often bind only in the micromolar–millimolar range, subsequent optimization is required to improve potency while retaining ligand efficiency and chemical tractability (Wasko et al., 2015).

A central step in this optimization is deciding how additional chemical matter should be introduced around the fragment hit. Medicinal chemists commonly use three complementary strategies (Murray and Rees, 2009; Exertier et al., 2024). **Fragment growing** expands a single fragment toward adjacent regions of a binding site. For example, MRTX1719 was advanced from a 4-(aminomethyl)phthalazin-1(2H)-one fragment identified by SPR and characterized by X-ray crystallography (Smith et al., 2022); related growing campaigns produced focal adhesion kinase inhibitors and the Akt inhibitor capivasertib (Heinrich et al., 2013; Wyatt et al., 2008; Addie et al., 2013). **Fragment linking** connects two distinct fragments that occupy neighboring regions through a chemically suitable linker, which must satisfy geometric constraints while limiting entropic penalties. **Fragment merging** combines overlapping fragments by consolidating shared substructures or pharmacophoric features. Successful examples include trypanothione reductase inhibitors (Exertier et al., 2024) and HSP90 inhibitors (Ren et al., 2014). Linking and merging can be especially difficult to automate because they require compatible attachment geometry or correct identification of shared structural motifs.

Despite their different chemical interpretations, growing, linking, and merging admit a common

*Equal Contribution. †: Part of this work was completed at the University of Chicago. Correspondence to: Xuefeng Liu <xuefeng.liu@medicine.ufl.edu>.

Findings of the Association for Computational Linguistics: EMNLP 2026, pages 32439–32452
October 24-29, 2026 ©2026 Association for Computational Linguistics

computational view: each can be expressed as conditional generation over SMILES. Growing maps one fragment to an expanded molecule; linking maps separated fragments to a molecule containing a connecting structure; and merging maps overlapping fragments to a molecule in which the shared region is identified and consolidated. Most existing generative systems focus on linker design and do not jointly model these three input–target relationships. A unified formulation can instead share chemical syntax and molecular priors while learning the transformation associated with each fragment configuration.

This shared generative capability is only part of the problem. Drug design is inherently multi-objective, requiring trade-offs among predicted binding, drug-likeness, synthesizability, solubility, and structural similarity. Maximum-likelihood pretraining does not directly optimize these properties. At the same time, pretrained molecular GPT models can provide useful expert policies (Liu et al., 2022, 2023b), while Pareto ranking and composite rewards offer complementary mechanisms for selecting candidates under multiple objectives. This motivates the following question:

Can a single molecular language model support fragment growing, linking, and merging while incorporating chemical priors and expert-guided, multi-objective alignment?

We introduce FRAGMENTGPT, a chemical language model that addresses these tasks within one framework. Our main contributions are:

A unified, energy-aware pre-training corpus. We cast growing, linking, and merging in a common conditional-completion format. Energy-based bond cleavage preferentially breaks weaker non-aromatic bonds while preserving aromatic rings, and cross-molecule maximum-common-substructure pairing adds merging examples in which two molecules contribute overlapping fragments to a new target. The resulting corpus teaches both linker construction and shared-substructure recognition without requiring an external fragment catalogue.

Reward-Ranked Alignment with Expert Exploration (RAE). RAE augments learner-generated candidates with samples from pretrained molecular expert policies, ranks the combined pool using Pareto and composite criteria, and uses the selected molecules for supervised fine-tuning. The procedure aligns a shared fragment generator with multiple pharmaceutical objectives while retaining exploration from expert models.

Evaluation across two molecular targets. We evaluate FRAGMENTGPT on RTCB and SARS-CoV-2 main protease (3CLPro). On RTCB, we further report five-seed variability, quantitative growing results, and an overlap-aware analysis of merging. Across both targets, FRAGMENTGPT improves validity and Average Normalized Reward over the input-matched Link-INVENT baseline for linking and merging. The ablation study further separates the roles of the two training stages: corpus pre-training provides most of the overlap-aware merging capability, whereas the full RAE procedure has its clearest effect on linking.

2 Preliminaries

LM. Let $F = [f_1, f_2, \cdots, f_n]$ denote an input sequence of tokens (fragment-based prompt), where each f_i belongs to a vocabulary $\mathcal{V}$. The output sequence is defined as $M = [m_1, m_2, \cdots, m_T]$ with $m_i \in \mathcal{M}$, where $\mathcal{V}$ and $\mathcal{M}$ may be distinct vocabularies. For convenience, we write $\mathbf{m}{<}\, t = [m_1, \cdots, mt-1]$ and $\mathbf{m}_T := M$, and use T to denote sequence length. Each training sample begins with a start token [BOS], followed by a sequence $\mathbf{m}$ with $y_i \in \mathcal{V}$, and terminates with an end token [EOS]. Molecules are represented as token sequences $\mathbf{m}$ that assemble into SMILES strings, covering both partial and complete structures. We denote string concatenation by $\circ$ and the Kleene closure of $\mathcal{V}$ by $\mathcal{V}^*$. The training corpus C is thus: $C := \{[\text{BOS}] \circ \mathbf{v} \circ [\text{EOS}] \mid \mathbf{v} \in \mathcal{V}^*\}$. The generator policy π_θ, parameterized by a deep neural network with weights θ, defines a distribution over sequences as $\pi_\theta(\mathbf{m}|\mathbf{f}) = \prod_{t=1}^{|\mathbf{m}|} \pi_\theta(m_t|\mathbf{f}, \mathbf{m}_{<t})$, where $\pi_\theta(m_t|\mathbf{f}, \mathbf{m}{<}\, t) = P(m_t|\mathbf{m}{<}\, t, X)$ denotes the conditional probability of token m_t. Decoding aims to select the most likely sequence from the hypothesis space: $\mathbf{m}^\star = \arg\max_{\mathbf{m}\in\mathcal{Y}_T} \log \pi_\theta(\mathbf{m}|\mathbf{f})$. Causal language modeling (CLM) reformulates the training objective as $\max_\theta \sum_{i=1}^{n} \log P(f_i|F_{<i};\theta)$, where $P(f_i|F_{<i};\theta)$, where $P(f_i|F_{<i};\theta)$ is the conditional probability of token f_i given its history. In this work, we employ a GPT policy π_θ trained on dataset $\mathcal{B}$ to capture prior chemical knowledge and generate chemically valid molecules. Molecules are represented using SMILES strings (Weininger, 1988). The tokenizer is first pre-trained with the Byte Pair Encoding (BPE) algorithm (Gage, 1994;

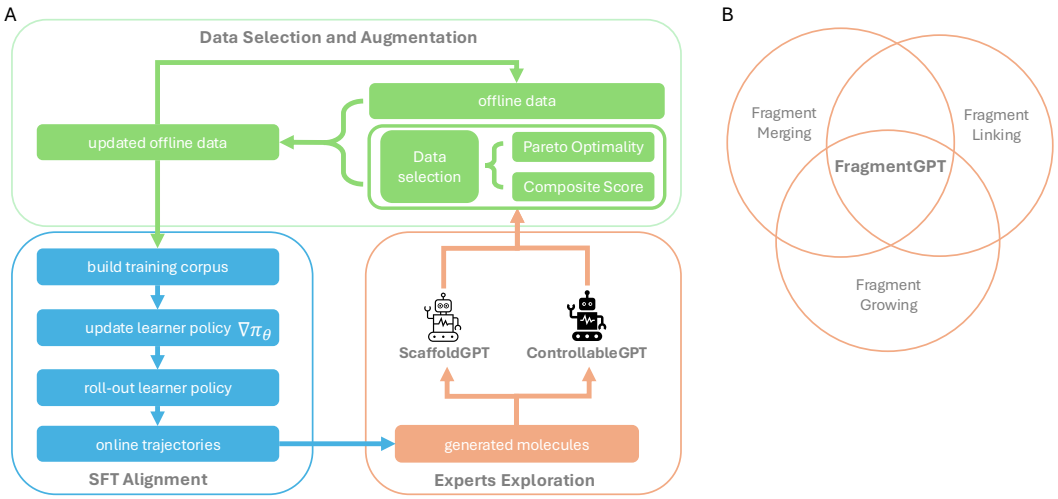


Figure 1: (A) The framework of Reward Ranked Alignment with Expert Exploration (RAE). (B) Functional overview of FRAGMENTGPT, which supports fragment merging, linking, and growing.

Sennrich et al., 2015). Building on this, we adopt an incremental pre-training strategy inspired by ScaffoldGPT (Liu et al., 2025b), which significantly enhances the model's ability to reconstruct valid molecules from fragments.

Fragment-based Drug Discovery. Given an initial prompt (fragment candidates) $\{F_{i=1}^n\} \sim \mathcal{X}$ and a drug discovery policy π_θ, the goal in fragment-based drug discovery is to find the optimal policy π_{θ^*} that maximizes the following objective:

$$\pi_{\theta^*} = \arg\max_{\pi_\theta} \mathbb{E}_{\{F\}\sim d_0}[R(M)|\theta, \{F_{i=1}^n\}], \quad (1)$$

where $M = \pi_\theta(\cdot \mid \{F_{i=1}^n\}), M_{1:T} = (m_1, \ldots, m_t, \ldots, m_T), m_t \in \mathcal{V}$. Regards to $\{F_{i=1}^n\}$, For $\{F_{i=1}^n\}$, F_1 corresponds to fragment growing, F_1^2 corresponds to fragment linking and merging, and F_1^n with $n > 2$ corresponds to the linking and merging of multiple fragments.

3 The FRAGMENTGPT Algorithm

In this work, we bridge the gap among fragment growing, linking, and merging by introducing FRAGMENTGPT (Fig. 1). FRAGMENTGPT directly trains a GPT model on SMILES representations, enabling it to generate complete molecules from fragment-based prompts. Furthermore, to address multi-objective molecular optimization, we propose RAE (Reward-Ranked Alignment with Expert Exploration)—a framework that combines supervised fine-tuning (SFT) retraining with expert-guided exploration, data selection, and augmentation.

Stage 1. Chemically-Aware Multi-Task Pre-training

Pre-Training Corpus Construction. To enable the model to perform three key operations—linking, merging, and growing—we explicitly construct pretraining corpus that reflects these tasks. Our approach begins by converting a database of molecules from their SMILES strings into graph representations, where atoms serve as nodes and bonds as edges. The core of our methodology is a chemically-aware, energy-based bond cleavage strategy. To ensure the generation of stable fragments, we annotate each bond with its dissociation energy and explicitly prevent the cleavage of bonds within aromatic rings. The algorithm preferentially targets weaker bonds for disconnection, guided by two key parameters: a maximum of three fragments per molecule to ensure substructures are significant, and a bond energy cutoff of 90 kcal/mol (377 kJ/mol) to selectively cleave common single bonds while preserving stronger ones. By applying this graph-based fragmentation algorithm across a large-scale molecular database, we have successfully generated a comprehensive corpus. The resulting corpus consists of chemically meaningful fragment–molecule pairs.

Linking: After we constructed a large corpus of chemically meaningful fragments for linker and merger tasks, we generate a fragment-pair dataset. As illustrated in Figure 2, given a molecule M, we partition it into three components: A, B, C. This yields two classes of fragment pairs. First, the pair (A, C) where the task is to generate a linker bridging the two fragments. Each resulting fragment

Algorithm 1 Reward Ranked Alignment with Expert Exploration (RAE)

Require: GPT-based generator policy π_θ^G; offline dataset $\mathcal{D}$; replay buffer $\mathcal{B}$; number of experts E.

$\mathcal{D} \to \mathcal{B}$. Initialize π_θ^G with GPT2-like Transformer with random weight θ.

for $n = 1, \ldots, N$ **do**

▷ /* Stage 1: SFT retraining */

$\{M\}_{i=1}^N \sim \rho_0$, where $\rho_0 \in \Delta(\mathcal{D})$. $\mathcal{B} \leftarrow \emptyset$.

Build the prompt $\mathcal{X}$ and training corpus C_{growing}, C_{linking}, C_{merging} (2) by using $\{M\}_{i=1}^N$.

Pre-train π_θ on C_{growing}, C_{linking} and C_{merging} via CLM objective (3).

Generate $M_{i=1}^N = \{(m_t, \ldots, m_T)\}_{i=1}^N \sim \pi_\theta(\cdot|\{F_{i=1}^n\}), \{F_{i=1}^n\} \sim \mathcal{X}$.

$\mathcal{B} \leftarrow \mathcal{B} \cup M_{i=1}^N$.

▷ /* Stage 2: Expert Exploration */

$\{M^{\text{EXPERT}}\}_{i=1}^N \sim \pi_j^{\text{EXPERT}}(M_i), i \in [N], j \in [E]$.

$\mathcal{B} \leftarrow \mathcal{B} \cup \{M^{\text{EXPERT}}\}_{i=1}^N$.

▷ /* Stage 3: Data Selection and Augmentation */

$\mathcal{D} \leftarrow 0.8 \cdot \textsc{Priority}(\mathcal{B}, R^{\text{P}}, R^{\text{Comp}}) \cup 0.2 \cdot \textsc{Random}(\mathcal{D})$. ▷ R^{Pareto} and $R^{\text{Composite}}$ are abbreviated.

pair, together with the original molecule M, forms a training instance for linking or merging.

Merging: Second, the pairs (AB, BC) where the task is to merge fragments based on their shared substructure B. To further enrich the corpus and encourage the model to learn molecule merging, we construct additional examples by randomly selecting two molecules and merging them via their maximum common substructure (MCS). As shown in Figure 2, given two molecules MOL_A and MOL_B, we first identify their MCS (highlighted in red), then align and merge them to form a new molecule MOL_M. The corresponding training corpus is represented as <p1>SMILES$_A$<p2>SMILES$_B$<L>SMILES$_M$ where the model learns to merge from two random molecules into a larger molecule.

Growing: Fragment growing is trained as an independent capability. Given a fragment A and a molecule M containing A as a substructure, the model learns to expand A into M. The training instance is represented as <p1>SMILES$_A$<L>SMILES$_M$. Growing, linking, and merging are trained as separate conditional-generation tasks and do not require a fixed execution order.

The corpus C and prompts $\mathcal{X}$ are constructed as follows:

$$\left\{[BOS], \overbrace{\langle p_1\rangle, \underbrace{f_{11}, \cdots, f_{1T}}_{\text{fragment } F_1}, \langle p_2\rangle, \underbrace{f_{21}, \cdots, f_{2T}}_{\text{fragment } F_2}, \langle L\rangle}^{\text{PROMPT}}, \underbrace{m_1, \cdots, m_T}_{\text{target molecule M}}, [EOS]\right\}. \tag{2}$$

Consequently, the model is able to generate the corresponding complete molecule when provided with a pair of fragments as the input prompt.

Pretrain a GPT generator. We denote the GPT generator policy by π_θ, which models the probability p of producing the t^{th} token in a target molecule Y. This probability is conditioned on all previously generated tokens $\mathbf{y}_{<t} = [y_1, \ldots, y_{t-1}]$ in the sequence, as well as on the given pair of fragment compounds $\{F_{i=1}^n\}$, formally expressed as $\pi_\theta(y_t | \mathbf{y}_{<t}, \{F_{i=1}^n\}) = p(y_t | \mathbf{y}_{<t}, \{F_{i=1}^n\})$. The parameters θ of the generator policy π_θ are optimized by minimizing the negative log-likelihood (NLL) over the training corpus:

$$\begin{aligned} \text{NLL} &= -\log P(Y \mid \{F_{i=1}^n\}) \\ &= -\sum_{t=1}^{T} \log P(y_t \mid y_{t-1}, \ldots, y_1, \{F_{i=1}^n\}) \\ &= -\sum_{t=1}^{T} \log \pi_\theta(y_t \mid y_{1:t-1}, \{F_{i=1}^n\}). \end{aligned} \tag{3}$$

Here, T denotes the total number of tokens in the target molecule Y. The negative log-likelihood (NLL) thus measures the likelihood of reconstructing a designated target molecule from the given fragment pair $\{F_{i=1}^n\}$. We adopt an incremental molecular pre-training strategy following Liu et al. (2025b). In the first phase, the GPT model learns the distribution and syntax of valid molecular SMILES from a large molecular corpus. In the second phase, training is extended to the conditional fragment–molecule corpora described above, enabling the same autoregressive model to perform fragment growing, linking, and merging.

Stage 2. Expert Exploration. We begin by sampling a batch of prompts $\mathcal{X}_t = \{\{F_{i=1}^n\}_1^t, \cdots, \{F_{i=1}^n\}_b^t\} \subseteq \mathcal{X}$, and for each $\{F_{i=1}^n\}_i^t \in \mathcal{X}_t$, we generate molecules

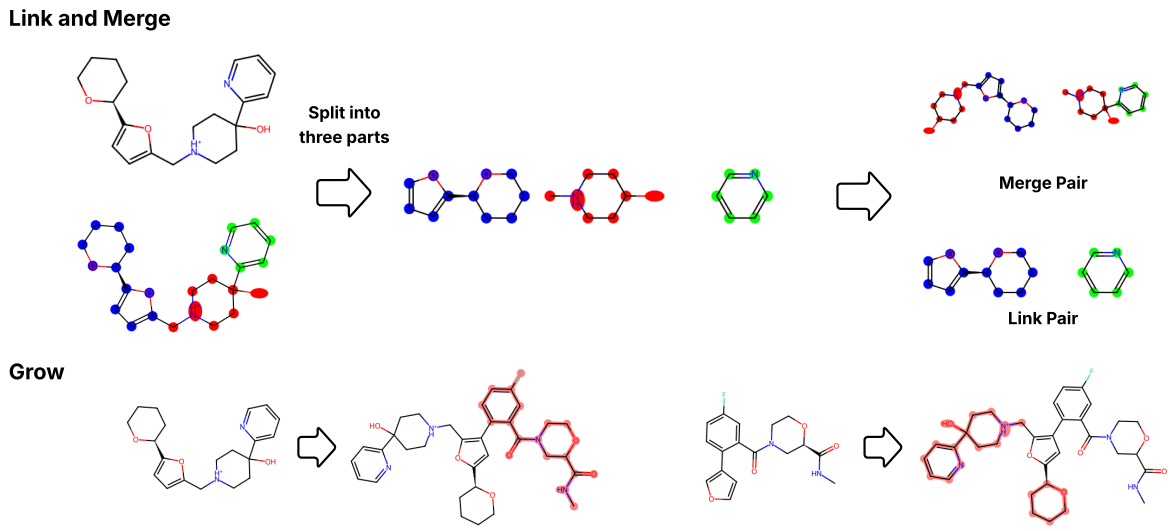


Figure 2: Overview of training corpus construction for linking, merging and growing. For linking and merging, given a SMILES string, we divide it into three parts, which we denote as fragments A, B, and C. For the construction of the *linker training corpus*, we use fragments A and C as inputs, and represent each training instance as <p1>SMILES$_A$<p2>SMILES$_C$<L>SMILES$_{ABC}$. Similarly, for the *merging training corpus*, we take the concatenated fragments AB and BC as inputs, resulting in the representation <p1>SMILES$_{AB}$<p2>SMILES$_{BC}$<L>SMILES$_{ABC}$. For growing given fragment A and a molecule M which contains A as substructure, the model will take <p1>SMILES$_A$<L>SMILES$_M$

$M_1, \cdots, M_K \sim \pi_\theta\left(\cdot|\{F_{i=1}^n\}_i^t\right)$. Next, we leverage the expert policy π^{E} to produce more diverse molecular candidates: $M_1^{\textsc{Expert}}, \cdots, M_K^{\textsc{Expert}} \sim \pi^{\textsc{Expert}}\left(\cdot|M_i\right), i \in [K]$. Specifically, we employ ScaffoldGPT (Liu et al., 2025b) and ControllableGPT (Liu et al., 2025c,d) to promote structural diversity. Finally, we combine the molecules generated by both the learner policy and the expert policy into the replay buffer $\mathcal{B}$, $\mathcal{B} \leftarrow \{M\} \cup \left\{M^{\textsc{Expert}}\right\}$.

Stage 3. Data Selection and Augmentation. In this step, we employ the Pareto optimality reward model $R^{\textsc{Pareto}}(M)$ from Abeer et al. (2024) and the composite score reward model $R^{\textsc{Composite}}(M)$ from Liu et al. (2025b). In this work, we regarded each pharmaceutical property as a critic C and considered an ensemble critics $\mathbf{C}$ as follows:

$$[C^{\text{Druglikeness}}, C^{\text{Solubility}}, C^{\text{Synthesizability}}, C^{\text{Docking}}, C^{\text{Tanimoto}}],$$

where each critic $C : M \to \mathbb{R}$ acts as a distinct evaluator for a specific pharmaceutical attribute. We built the composite score reward function as follow:

$$R^{\textsc{composite}}(M) := \sum_{i=1}^{|\mathbf{C}|} \text{Norm}\left(C_i(M)\right),$$

where Norm is employed to standardize diverse attributes to a consistent scale. For Pareto optimality, we identify the Pareto front by iteratively collected fronts $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_n$ until the replay buffer $\mathcal{B}$ is empty. These fronts partitions $\mathcal{B}$ with $\mathcal{P}_1 \cap \mathcal{P}_2 = \emptyset$ for $i \neq j$ and $\mathcal{B} = \cup_i \mathcal{P}_i$. The ranking method assign each point $M \in \mathcal{P}_j$ the rank same as the reward as

$$\begin{aligned} R^{\textsc{Pareto}}(M) &= \textsc{Rank}_{\mathcal{B}}(M) \\ &= \sum_i |\mathcal{P}_i|, \forall M \in \mathcal{P}_i. \end{aligned} \tag{4}$$

Both models are used to evaluate each collected molecule in the replay buffer $\mathcal{B}$. We then reconstruct our offline dataset $\mathcal{D}$

$$\begin{aligned} \mathcal{D} \leftarrow 80\% \cdot \textsc{Priority}\left(\mathcal{B}, R^{\textsc{Pareto}}, R^{\textsc{Composite}}\right) \\ \cup\, 20\% \cdot \textsc{Random}\left(\mathcal{D}\right) \end{aligned} \tag{5}$$

by randomly sampling 20% from the previous $\mathcal{D}$ and 80% from the buffer, with sampling prioritized according to their ranking (Liu et al., 2025e). We then proceed to Stage 1, where we retrain the learner policy on the dataset aligned with Pareto optimality and composite score.

4 Experiments

4.1 Experimental Setup

Language model. We use GPT-2-like Transformers for causal language modeling. Training consists of two pre-training phases followed by RAE fine-tuning, as detailed in Section 3 and Algorithm 1.

Baselines. We compare against Link-INVENT (Guo et al., 2023), an encoder–decoder model based on recurrent neural networks that receives two molecular fragments (*warheads*) and generates a linker between them. We additionally report 3DLinker (Huang et al., 2022) and DiffLinker (Igashov et al., 2024). 3DLinker predicts anchor atoms and jointly generates linker graphs and 3D structures with an E(3)-equivariant graph variational autoencoder, whereas DiffLinker uses an E(3)-equivariant conditional diffusion model with 3D fragment information.

Link-INVENT is our primary input-matched baseline because it receives the same two-fragment SMILES condition as FRAGMENTGPT. The released 3DLinker and DiffLinker pipelines instead use their intended geometry-aware inputs: reference-conformer fragment coordinates for both models and SDF/anchor information for DiffLinker. Their results therefore provide contextual rather than like-for-like comparisons. Neither model natively supports merging; for rows marked *adapted*, overlapping fragments are passed through the corresponding linking formulation. SyntaLinker (Yang et al., 2020) is a comparable sequence model, but its published results were not reproducible under our evaluation protocol. None of the evaluated baselines accepts single-fragment input, so growing is reported as an absolute capability evaluation.

Datasets. We use one million compounds from ZINC15 docked against the human cancer protein RTCB (PDB ID: 4DWQ), following the Cancer dataset used in prior work (Liu et al., 2023a, 2025a). We also evaluate SARS-CoV-2 main protease (3CLPro; PDB ID: 7BQY) using the corresponding resource described by Xuefeng et al. (2024). The two targets differ in protein family, pocket geometry, and disease context.

Metrics. We evaluate generated molecules with RDKit (Landrum et al.). **Validity** measures whether a generated SMILES string is syntactically valid. **Drug-likeness** is measured by QED. **Solubility** is approximated from the water–octanol partition coefficient (LogP). **Synthesizability** uses the synthetic accessibility (SA) score, where lower values indicate easier synthesis (Ertl and Schuffenhauer, 2009). **Docking** measures predicted binding potential using the same SMILES-based surrogate for all methods (see Appendix A.5). **Similarity** is the Tanimoto similarity between the generated molecule and the source molecule. **Average Top-10% Normalized Reward** is the mean normalized reward among the top 10% of generated molecules, and **Average Normalized Reward**, our primary metric, averages the five normalized critics—docking, QED, SA, solubility, and similarity—over all valid molecules, with equal weights. Each evaluation uses 1,280 molecules.

4.2 Main Results

Table 1 summarizes the RTCB results. Link-INVENT provides the main input-matched comparison, while 3DLinker and DiffLinker are shown separately as geometry-aware contextual baselines.

On RTCB, finetuned FRAGMENTGPT improves Average Normalized Reward over Link-INVENT by 0.101 for linking and 0.208 for merging. Because both methods receive the same SMILES inputs, these are the primary like-for-like comparisons, and the margins are large relative to the five-seed variation. For overlapping merge prompts, Link-INVENT can connect rather than collapse the shared substructure, creating a duplicated region and a more complex molecule. The 3DLinker and DiffLinker rows should instead be interpreted as contextual comparisons because those methods receive geometry-aware inputs, and their merge results are adapted rather than native capabilities.

Growing is evaluated quantitatively in Table 1 and qualitatively in Figure 3 and Appendix A.1. FRAGMENTGPT produces valid molecules for every evaluated growing prompt and reaches a top-10% normalized reward of 0.733, close to its linking result of 0.734. Its Average Normalized Reward for growing (0.531) is lower than for linking (0.616), consistent with the weaker conditioning provided by a single input fragment.

Fine-tuning has a task-dependent effect. Average Normalized Reward increases from 0.563 to 0.616 for linking, but only from 0.630 to 0.651 for merging. This pattern suggests that the energy-aware corpus supplies most of the overlap-aware merging capability, whereas RAE is more influential when the model has greater freedom to optimize a linker.

The main RTCB pattern extends to 3CLPro (Table 2). FRAGMENTGPT maintains 1.000 validity, compared with 0.520 and 0.430 for Link-INVENT

Task	Method	Valid. ↑	Avg. Reward ↑	Top-10% ↑	Dock. ↓	QED ↑	SA ↓	Sol. ↑	Sim. ↑
Source	Original molecules	–	0.540	0.768	-8.505	0.720	2.981	2.364	–
Grow	FRAGMENTGPT	1.000	0.531	0.733	-6.814	0.563	2.933	3.463	0.623
Link	Link-INVENT (Guo et al., 2023)	0.517	0.515	0.609	-7.379	0.540	3.059	**4.003**	0.521
Link	3DLinker† (Huang et al., 2022)	0.762	0.403	0.636	-7.448	0.571	4.220	2.379	0.428
Link	DiffLinker† (Igashov et al., 2024)	0.891	0.428	0.655	-7.868	0.606	4.004	2.559	0.451
Link	FRAGMENTGPT (pretrained)	0.979	0.563	**0.738**	-8.206	0.701	2.815	2.920	**0.590**
Link	FRAGMENTGPT (finetuned)	**0.991**	**0.616**	0.734	**-8.679**	**0.730**	**2.418**	3.684	0.564
Merge	Link-INVENT (Guo et al., 2023)	0.409	0.443	0.510	-2.529	0.229	3.607	**5.257**	0.720
Merge	3DLinker†‡ (Huang et al., 2022)	0.708	0.321	0.515	-4.652	0.358	4.672	3.397	0.475
Merge	DiffLinker†‡ (Igashov et al., 2024)	0.858	0.345	0.550	-4.972	0.365	4.574	3.557	0.528
Merge	FRAGMENTGPT (pretrained)	0.947	0.630	0.788	-8.235	0.692	2.859	2.829	0.953
Merge	FRAGMENTGPT (finetuned)	**0.993**	**0.651**	**0.791**	**-8.370**	**0.727**	**2.699**	2.815	**0.965**

Table 1: RTCB (PDB ID: 4DWQ) results. Bold and underline denote the best and second-best results, respectively, within each task where the comparison is directly meaningful. †Geometry-aware baseline with additional 3D input; ‡adapted linking formulation used for merging. Growing is reported as an absolute capability evaluation because the evaluated baselines do not accept single-fragment input.

Task	Method	Valid. ↑	Avg. Reward ↑	Top-10% ↑	Dock. ↓	QED ↑	SA ↓	Sol. ↑	Sim. ↑
Source	Original molecules	1.000	0.527	0.718	-8.462	0.694	3.105	2.482	–
Link	Link-INVENT	0.520	0.534	0.616	**-8.436**	0.498	3.063	**4.320**	0.543
	FRAGMENTGPT (pretrained)	1.000	0.538	0.718	-8.073	**0.671**	2.790	2.897	0.596
	FRAGMENTGPT (finetuned)	**1.000**	**0.547**	**0.727**	-7.996	0.664	**2.766**	2.887	**0.598**
Merge	Link-INVENT	0.430	0.450	0.524	-6.476	0.204	3.805	**5.377**	0.747
	FRAGMENTGPT (pretrained)	1.000	0.508	0.718	**-7.754**	**0.628**	**2.768**	2.806	**0.896**
	FRAGMENTGPT (finetuned)	**1.000**	**0.583**	**0.774**	-7.750	0.625	2.789	2.816	0.895

Table 2: Results on SARS-CoV-2 main protease (3CLPro; PDB ID: 7BQY). Bold denotes the best generated-model result within each task and metric.

Set	N	Valid	Exact	Mean Sim.
Overlap prompts	1,280	1,273	0.647	0.881

Table 3: Overlap-aware merging on RTCB. Exact reconstruction is the fraction of prompts for which the designated merged product is recovered; source similarity is averaged over non-exact generations.

on linking and merging, and obtains the highest Average Normalized Reward for both tasks. The improvement is not uniform across individual objectives: Link-INVENT obtains the better linking docking score (-8.436 vs. -7.996), and RAE yields only a modest linking reward gain over pretraining (0.547 vs. 0.538). For merging, RAE increases the aggregate reward from 0.508 to 0.583 without improving every individual critic.

Because the within-molecule merge task is reconstruction-oriented, Table 3 measures overlap identification rather than novelty. Given AB and BC, exact recovery of ABC requires identifying and collapsing the shared region B. FRAGMENTGPT exactly reconstructs the designated product for 64.7% of 1,280 overlap prompts, with 0.995 validity; non-exact generations retain a mean source similarity of 0.881. The cross-molecule MCS corpus and the COVID Moonshot case study in Appendix A.1.2, which uses a separate set of human-designed molecules, provide complementary evidence.

4.3 Ablation Study

Table 4 summarizes the ablation results on RTCB. The ablation pattern differs substantially by task. For linking, removing PFRank, all experts, ScaffoldGPT, or ControllableGPT reduces Average Normalized Reward from 0.616 to 0.563–0.569, close to the pretrained value of 0.563. Each ablation therefore removes most of the RAE fine-tuning gain. For merging, pretraining already reaches 0.630, the complete model reaches 0.651, and the ablations remain at 0.627–0.630. The maximum

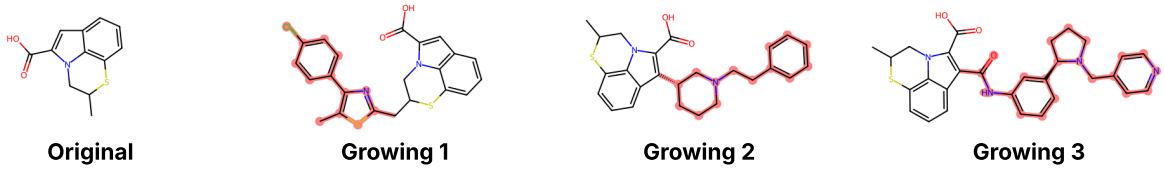


Figure 3: Fragment-growing examples produced by FRAGMENTGPT. Highlighted substructures indicate the newly generated regions.

Variant	Link ↑	Merge ↑
FRAGMENTGPT (finetuned)	**0.616**	**0.651**
FRAGMENTGPT (pretrained)	0.563	0.630
No PFRank	0.565	0.629
No experts	0.569	0.630
No ScaffoldGPT	0.563	0.627
No ControllableGPT	0.569	0.630

Table 4: Ablation results on RTCB using Average Normalized Reward, our primary metric. RAE has a substantially larger effect on fragment linking than on fragment merging. Full per-metric ablation results are reported in the Appendix.

common substructure largely fixes the merged core, leaving less freedom for reward-guided optimization. These results support the corpus as the main source of merging capability and the complete RAE loop as most consequential for linking. Full results for all evaluation metrics are provided in Table 5 in the Appendix.

4.4 Case Studies

4.4.1 Fragment Growing

Figure 3 shows fragment-growing examples based on an original molecule from an NMR-based screening study (G. Ferreira and D. Andricopulo, 2017). The examples illustrate how FRAGMENTGPT expands a simple input fragment into chemically plausible, drug-like structures. Additional case studies are provided in Appendix A.1.

5 CONCLUSION

We introduced FRAGMENTGPT, a unified molecular language model for fragment growing, linking, and merging. The central component is a chemically-aware pre-training corpus that casts the three fragment transformations into a shared conditional-generation format, using energy-guided bond cleavage and overlap-aware merging examples. We further introduced Reward-Ranked Alignment with Expert Exploration (RAE) to align the pretrained generator with multiple drug-design objectives through expert-guided exploration, Pareto/composite-reward selection, and supervised fine-tuning.

Across RTCB and 3CLPro, FRAGMENTGPT achieves high validity and improves Average Normalized Reward over the input-matched Link-INVENT baseline for both linking and merging, although the advantage does not hold for every individual metric. Our analysis further shows that corpus pre-training provides most of the overlap-aware merging capability, whereas RAE has its clearest effect on fragment linking. These results support unified conditional pre-training as a practical approach for learning complementary fragment-design operations within a single molecular language model.

6 LIMITATIONS

Although FRAGMENTGPT demonstrates strong performance on fragment growing, linking, and merging, the current framework is limited to SMILES-based molecular representations and does not explicitly model 3D molecular geometry or protein–pocket interactions. Our evaluation covers RTCB and 3CLPro, but further experiments across more diverse targets and protein families are needed to establish broader generalizability. In addition, docking scores are estimated using a learned surrogate for computational efficiency, which may not fully capture physics-based docking or experimental binding affinity. The within-molecule merging benchmark is also partly reconstruction-oriented and therefore does not directly measure molecular novelty or diversity. FRAGMENTGPT may pose dual-use risks for harmful compound design; generated molecules therefore require expert review and validation. Future work will incorporate structural representations, protein–pocket

conditioning, broader target evaluation, and more comprehensive prospective validation.

Acknowledgements

We thank Ian Foster, Mengdi Wang, and Hening Lin for their helpful discussions. This work is supported by the college of medicine and UF AI for health institute, the RadBio-AI project (DE-AC02-06CH11357), U.S. Department of Energy Office of Science, Office of Biological and Environment Research, the Improve project under contract (75N91019F00134, 75N91019D00024, 89233218CNA000001, DE-AC05-00OR22725), the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

References

ANM Nafiz Abeer, Nathan M Urban, M Ryan Weil, Francis J Alexander, and Byung-Jun Yoon. 2024. Multi-objective latent space optimization of generative molecular design models. *Patterns*, 5(10).

Matt Addie, Peter Ballard, David Buttar, Claire Crafter, Gordon Currie, Barry R. Davies, Judit Debreczeni, Hannah Dry, Philippa Dudley, Ryan Greenwood, Paul D. Johnson, Jason G. Kettle, Clare Lane, Gillian Lamont, Andrew Leach, Richard W. A. Luke, Jeff Morris, Donald Ogilvie, Ken Page, and 3 others. 2013. Discovery of 4-amino-n-[(1s)-1-(4-chlorophenyl)-3-hydroxypropyl]-1-(7h-pyrrolo[2,3-d]pyrimidin-4-yl)piperidine-4-carboxamide (azd5363), an orally bioavailable, potent inhibitor of akt kinases. *Journal of Medicinal Chemistry*, 56(5):2059–2073. PMID: 23394218.

Melissa L Boby, Daren Fearon, Matteo Ferla, Mihajlo Filep, Lizbé Koekemoer, Matthew C Robinson, COVID Moonshot Consortium‡, John D Chodera, Alpha A Lee, Nir London, and 1 others. 2023. Open science discovery of potent noncovalent sars-cov-2 main protease inhibitors. *Science*, 382(6671):eabo7201.

Gideon Bollag, Jack Tsai, Jian Zhang, Christina Zhang, Pamela Ibrahim, Keith Nolop, and Paul Hirth. 2012. Vemurafenib: the first drug approved for braf-mutant cancer. *Nature Reviews Drug Discovery*, 11(11):873–886. Epub 2012 Oct 12. PMID: 23060265.

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

Cécile Exertier, Alessandra Salerno, Lorenzo Antonelli, Annarita Fiorillo, Riccardo Ocello, Francesca Seghetti, Jessica Caciolla, Elisa Uliassi, Matteo Masetti, Eleonora Fiorentino, Stefania Orsini, Trentina Di Muccio, Andrea Ilari, and Maria Laura Bolognesi. 2024. Fragment merging, growing, and linking identify new trypanothione reductase inhibitors for leishmaniasis. *Journal of Medicinal Chemistry*, 67(1):402–419. PMID: 38164929.

W. J. Fairbrother, J. D. Leverson, D. Sampath, and A. J. Souers. 2019. Discovery and development of venetoclax, a selective antagonist of bcl-2. In J. Fischer, C. Klein, and W. E. Childers, editors, *Successful Drug Discovery*, volume 4, pages 225–245. Wiley-VCH, Weinheim. [Google Scholar].

Leonardo G. Ferreira and Adriano D. Andricopulo. 2017. From protein structure to small-molecules: recent advances and applications to fragment-based drug discovery. *Current topics in medicinal chemistry*, 17(20):2260–2270.

Philip Gage. 1994. A new algorithm for data compression. *The C Users Journal*, 12(2):23–38.

Jeff Guo, Franziska Knuth, Christian Margreitter, Jon Paul Janet, Kostas Papadopoulos, Ola Engkvist, and Atanas Patronov. 2023. Link-invent: generative linker design with reinforcement learning. *Digital Discovery*, 2(2):392–408.

Timo Heinrich, Jeyaprakashnarayanan Seenisamy, Lourdusamy Emmanuvel, Santosh Kulkarni, Jorg Bomke, Felix Rohdich, Hartmut Greiner, Christina Esdar, Mireille Krier, Ulrich Gradler, and 1 others. 2013. Fragment-based discovery of new highly substituted 1 h-pyrrolo [2, 3-b]-and 3 h-imidazolo [4, 5-b]-pyridines as focal adhesion kinase inhibitors. *Journal of medicinal chemistry*, 56(3):1160–1170.

Yinan Huang, Xingang Peng, Jianzhu Ma, and Muhan Zhang. 2022. 3DLinker: An e(3) equivariant variational autoencoder for molecular linker design. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 9280–9294. PMLR.

Ilia Igashov, Hannes Stärk, Clément Vignac, Arne Schneuing, Victor Garcia Satorras, Pascal Frossard, Max Welling, Michael Bronstein, and Bruno Correia. 2024. Equivariant 3d-conditional diffusion model for molecular linker design. *Nature Machine Intelligence*, 6(4):417–427.

Greg Landrum and 1 others. RDkit: Open-source cheminformatics software. `https://www.rdkit.org`. Accessed Oct 2023.

Brian A. Lanman, Jennifer R. Allen, John G. Allen, Albert K. Amegadzie, Kate S. Ashton, Shon K. Booker, Jian Jeffrey Chen, Ning Chen, Michael J. Frohn, Guy Goodman, David J. Kopecky, Longbin Liu, Patricia Lopez, Jonathan D. Low, Vu Ma, Ana E. Minatti, Thomas T. Nguyen, Nobuko Nishimura, Alexander J. Pickrell, and 25 others. 2020. Discovery of a covalent inhibitor of krasg12c (amg 510) for the treatment of solid tumors. *Journal of Medicinal Chemistry*, 63(1):52–65. PMID: 31820981.

Andrew R Leach and Michael M Hann. 2011. Molecular complexity and fragment-based drug discovery: ten years on. *Current Opinion in Chemical Biology*, 15(4):489–496. Next Generation Therapeutics.

Xuefeng Liu, Songhao Jiang, Siyu Chen, Zhuoran Yang, Yuxin Chen, Ian Foster, and Rick Stevens. 2025a. Drugimprovergpt: A large language model for drug optimization with fine-tuning via structured policy optimization. *arXiv preprint arXiv:2502.07237*.

Xuefeng Liu, Songhao Jiang, Ian Foster, Jinbo Xu, and Rick Stevens. 2025b. Scaffoldgpt: A scaffold-based gpt model for drug optimization. *arXiv preprint arXiv:2502.06891*.

Xuefeng Liu, Songhao Jiang, Ian Foster, Jinbo Xu, and Rick L Stevens. Scaffold-driven gpt model for drug optimization. In *ICML 2025 Generative AI and Biology (GenBio) Workshop*.

Xuefeng Liu, Songhao Jiang, Bo Li, and Rick Stevens. 2025c. Controllablegpt: A ground-up designed controllable gpt for molecule optimization. *arXiv preprint arXiv:2502.10631*.

Xuefeng Liu, Songhao Jiang, Bo Li, and Rick L Stevens. 2025d. A ground-up designed controllable gpt for molecule optimization. In *ICML 2025 Generative AI and Biology (GenBio) Workshop*.

Xuefeng Liu, Songhao Jiang, Archit Vasan, Alexander Brace, Ozan Gokdemir, Thomas Brettin, Fangfang Xia, Ian Foster, and Rick Stevens. 2023a. Drugimprover: Utilizing reinforcement learning for multi-objective alignment in drug optimization. In *NeurIPS 2023 Workshop on New Frontiers of AI for Drug Discovery and Development*.

Xuefeng Liu, Hung TC Le, Siyu Chen, Rick Stevens, Zhuoran Yang, Matthew R Walter, and Yuxin Chen. 2025e. Active advantage-aligned online reinforcement learning with offline data. *arXiv preprint arXiv:2502.07937*.

Xuefeng Liu, Fangfang Xia, Rick L Stevens, and Yuxin Chen. 2022. Cost-effective online contextual model selection. *arXiv preprint arXiv:2207.06030*.

Xuefeng Liu, Takuma Yoneda, Rick L Stevens, Matthew R Walter, and Yuxin Chen. 2023b. Blending imitation and reinforcement learning for robust policy improvement. *arXiv preprint arXiv:2310.01737*.

Christopher W. Murray and David C. Rees. 2009. The rise of fragment-based drug discovery. *Nature Chemistry*, 1(3):187–192.

Jie Ren, Jie Li, Yu Wang, Wei Chen, An Shen, Honglin Liu, Dan Chen, Dongmei Cao, Ying Li, Nan Zhang, Yili Xu, Meiyu Geng, Jian He, Bing Xiong, and Jianping Shen. 2014. Identification of a new series of potent diphenol hsp90 inhibitors by fragment merging and structure-based optimization. *Bioorganic & Medicinal Chemistry Letters*, 24(11):2525–2529. Epub 2014 Apr 8.

Rico Sennrich, Barry Haddow, and Alexandra Birch. 2015. Neural machine translation of rare words with subword units. *arXiv preprint arXiv:1508.07909*.

Christopher R Smith, Ruth Aranda, Thomas P Bobinski, David M Briere, Aaron C Burns, James G Christensen, Jeffery Clarine, Lars D Engstrom, Robin J Gunn, Anthony Ivetac, and 1 others. 2022. Fragment-based discovery of mrtx1719, a synthetic lethal inhibitor of the prmt5• mta complex for the treatment of mtap-deleted cancers. *Journal of Medicinal Chemistry*, 65(3):1749–1766.

Archit Vasan, Rick Stevens, Arvind Ramanathan, and Vishwanath Venkatram. 2023. Benchmarking language-based docking models.

Matthew J. Wasko, Kelly A. Pellegrene, Jeffry D. Madura, and Christopher K. Surratt. 2015. A role for fragment-based drug design in developing novel lead compounds for central nervous system targets. *Frontiers in Neurology*, 6:197.

David Weininger. 1988. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of chemical information and computer sciences*, 28(1):31–36.

Georg E Winter, Andreas Mayer, Dennis L Buckley, Michael A Erb, Justine E Roderick, Sarah Vittori, Jaime M Reyes, Julia di Iulio, Amanda Souza, Christopher J Ott, and 1 others. 2017. Bet bromodomain proteins function as master transcription elongation factors independent of cdk9 recruitment. *Molecular cell*, 67(1):5–18.

Paul G Wyatt, Andrew J Woodhead, Valerio Berdini, John A Boulstridge, Maria G Carr, David M Cross, Deborah J Davis, Lindsay A Devine, Theresa R Early, Ruth E Feltell, and 1 others. 2008. Identification of n-(4-piperidinyl)-4-(2, 6-dichlorobenzoylamino)-1 h-pyrazole-3-carboxamide (at7519), a novel cyclin dependent kinase inhibitor using fragment-based x-ray crystallography and structure based drug design. *Journal of medicinal chemistry*, 51(16):4986–4999.

Weijun Xu and Congbao Kang. 2025. Fragment-based drug design: From then until now, and toward the future. *Journal of Medicinal Chemistry*, 68(5):5000–5004. PMID: 39992814.

Liu Xuefeng, Tien Chih-Chan, Ding Peng, Jiang Songhao, and Stevens Rick. 2024. Entropy-reinforced planning with large language models for de novo drug discovery. *arXiv 2024*.

Yuyao Yang, Shuangjia Zheng, Shimin Su, Chao Zhao, Jun Xu, and Hongming Chen. 2020. Syntalinker: automatic fragment linking with deep conditional transformer neural networks. *Chem. Sci.*, 11:8312–8322.

A Appendix

A.1 Case Studies

A.1.1 Fragment Linking

dBET6, as visualized in figure 4, is a PROTAC molecule with high structural selectivity that bridges the E3 ubiquitin ligase CRBN and the target protein BRD4 (Winter et al., 2017). We use this example to investigate whether our model can identify reasonable linkage sites and propose a suitable linker, and results are represented in figure 5.

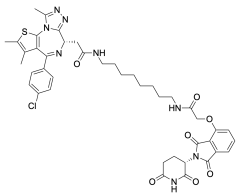


Figure 4: dBET6 Chemical Structure

The molecular generation process predominantly yielded structures where the two fragments are connected via a linker, with only a few exceptions showing direct linking. All generated linkers were determined to be chemically stable. Notably, several of these linkers incorporate an imidazolinone scaffold. This structural motif is highly significant in medicinal chemistry, widely utilized in drug molecules for its chemical stability and biocompatibility. Furthermore, a particularly noteworthy achievement is that our model successfully identified the acyl linking site on one of the fragments, despite not being provided with any explicit information about this site in the inference stage. This demonstrates the model's robust capability for implicit learning and recognition of key chemical features.

A.1.2 Fragment Merging

To assess our model's ability to merge different fragments, we utilized the COVID Moonshot dataset (Boby et al., 2023)—a fully open science, crowd-sourced, and structure-enabled drug discovery campaign against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (Mpro). This dataset provides an ideal testbed for evaluating fragment-based drug design approaches due to its comprehensive collection of crystallographic fragment hits and subsequent rational design submissions. The submission TRY-UNI-714a760b represents a class of rationally designed molecules that were constructed based on the strategic combination of five distinct fragments identified through crystallographic screening. These fragments were carefully selected and merged by human medicinal chemists who recognized their spatial complementarity and potential for synergistic binding interactions within the Mpro active site. To evaluate our model's fragment merging capabilities, we challenged it to reproduce this human-designed approach by providing the same five fragments as input and tasking it with generating novel molecular combinations. Remarkably, our model successfully generated four molecules among the twenty four molecules present in the submission series, as shown in figure 6. Our model can efficiently identify a variety of drug-related functional groups, including pyridine rings, amide bonds, and others. This achievement demonstrates our model's sophisticated ability to identify mergeable sites and recognize the structural compatibility between different fragment scaffolds.

The recovery of four human-designed molecules from the submission series highlights several strengths of our approach: (1) the model's capacity to recognize pharmacophoric features that are essential for maintaining binding affinity, (2) its ability to identify chemically feasible linking strategies between fragments. This performance validates our model's potential as a valuable tool for fragment-based drug design, particularly in its ability to automate and accelerate the traditionally labor-intensive process of fragment merging and optimization.

A.2 Full Ablation Results

Table 5 reports the complete per-metric ablation results on RTCB. The task-dependent pattern is consistent with the main-text analysis: RAE has its clearest effect on linking, whereas the energy-aware pretraining corpus supplies most of the merging capability.

A.3 Pre-training and finetuning dataset

For pretraining, we utilized the ZINC dataset, selecting only molecules categorized as Standard, In-Stock, and Drug-Like, which resulted in approximately 11 million compounds. Preprocessing involved two straightforward steps: (i) canonicalizing SMILES strings using `Chem.MolToSmiles(Chem.MolFromSmiles(mol), True)`, and (ii) discarding molecules with empty scaffold SMILES. These procedures are also incorporated into the HuggingFace data repository.

For finetuning, we employed 1 million compounds from the ZINC15 dataset, each docked to the RTCB protein (PDB ID: 4DWQ) implicated in

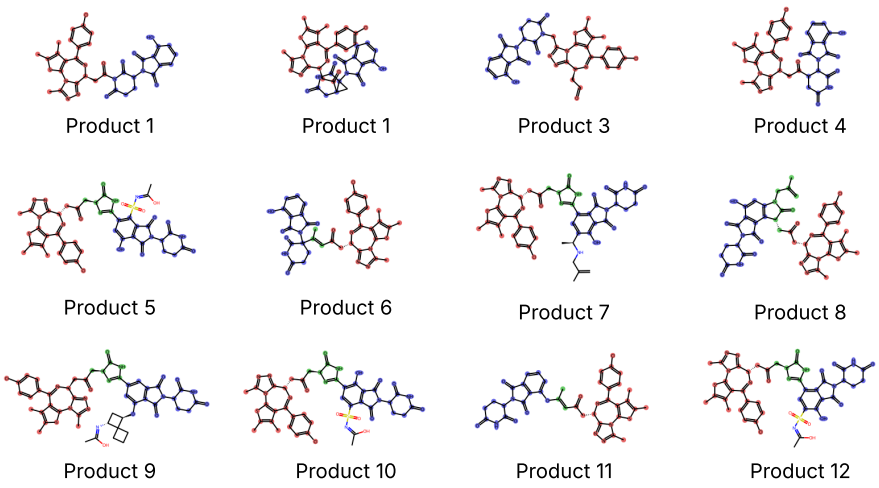


Figure 5: Chemical structure of a bifunctional PROTAC degrader. The molecule is designed to induce proximity between an E3 ubiquitin ligase and a Protein of Interest (POI). The E3-recruiting fragment is colored red, the POI-binding warhead is blue, and the interconnecting linker is highlighted in green.

Variant	Valid. ↑	Avg. Reward ↑	Top-10% ↑	Dock. ↓	QED ↑	SA ↓	Sol. ↑	Sim. ↑
Fragment linking								
FRAGMENTGPT (finetuned)	0.991	**0.616**	0.734	-8.679	0.730	2.418	3.684	0.564
FRAGMENTGPT (pretrained)	0.979	0.563	0.738	-8.206	0.701	2.815	2.920	0.590
No PFRank	0.975	0.565	0.740	-8.152	0.694	2.771	3.039	0.583
No experts	0.977	0.569	0.742	-8.250	0.710	2.753	2.922	0.588
No ScaffoldGPT	0.969	0.563	0.736	-8.122	0.694	2.795	3.033	0.578
No ControllableGPT	0.982	0.569	0.741	-8.234	0.701	2.779	2.987	0.594
Fragment merging								
FRAGMENTGPT (finetuned)	0.993	**0.651**	0.791	-8.370	0.727	2.699	2.815	0.965
FRAGMENTGPT (pretrained)	0.947	0.630	0.788	-8.235	0.692	2.859	2.829	0.953
No PFRank	0.935	0.629	0.788	-8.206	0.688	2.849	2.877	0.946
No experts	0.949	0.630	0.788	-8.262	0.691	2.860	2.841	0.951
No ScaffoldGPT	0.935	0.627	0.788	-8.142	0.680	2.863	2.919	0.943
No ControllableGPT	0.934	0.630	0.788	-8.193	0.688	2.848	2.838	0.952

Table 5: Full ablation results on RTCB for fragment linking and merging. Average Normalized Reward is the primary metric discussed in the main text.

human cancer, as provided by the Cancer dataset of Liu et al. (2023a), and applied them consistently across all baselines. In each finetuning epoch, a random subset of the desired number of molecules was sampled.

For evaluation and experimentation, we used 1,280 molecules from the Cancer dataset that did not overlap with the finetuning set.

A.4 Pretraining

Incremental pre-training. In the first phase, we train a GPT model exclusively on molecular data using Causal Language Modeling (CLM). In this autoregressive setup, the model learns to predict the next token in a sequence conditioned only on the preceding tokens. The Phase 1 training corpus is constructed as follows:

$$C_{\text{Phase 1}} = \left\{ [BOS], \langle L \rangle, \underbrace{y_1, \cdots, y_T}_{\text{target molecule Y}}, [EOS] \right\},$$

$\langle L \rangle$ is the token for ligand. (6)

In the second phase, we build upon the GPT model developed in Phase 1, which has already demonstrated strong performance in molecular generation. Here, the training objective is extended to focus on fragment–molecule relationships, where the model is trained on pairs of fragments and their corresponding complete molecules using Causal Language Modeling (CLM). The Phase 2 training

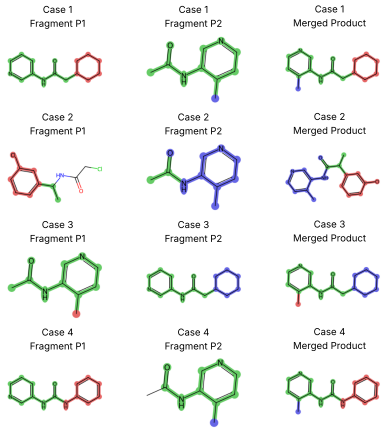


Figure 6: Four merging cases were generated, with all generated molecules contained within the 24 molecules of submission TRY-UNI-714a760b. Different fragments are marked in red and blue, while the merged overlapping regions are highlighted in green.

corpus is constructed as follows:

$$C_{\text{Phase 2}} = \left\{ [BOS], \langle p_1 \rangle, \underbrace{f_{11}, \cdots, f_{1T}}_{\text{fragment } F_1}, \langle p_2 \rangle, \underbrace{f_{21}, \cdots, f_{2T}}_{\text{fragment } F_2}, \langle L \rangle, \underbrace{y_1, \cdots, y_T}_{\text{target molecule Y}}, [EOS] \right\} \tag{7}$$

where $\langle F_1, F_2 \rangle$: fragments. Consequently, the model is able to generate the corresponding complete molecule when provided with a pair of fragments as the input prompt.

A.5 Surrogate model

The surrogate model (Vasan et al., 2023) adopts a simplified BERT-style transformer architecture, widely used in natural language processing. In this framework, tokenized SMILES strings are embedded with positional encodings and passed through a stack of five transformer blocks. Each block consists of a multi-head attention mechanism with 21 heads, followed by dropout, layer normalization with residual connections, and a feedforward subnetwork. The feedforward subnetwork is composed of two dense layers, again followed by dropout and layer normalization with residual connections. After the transformer stack, a final feedforward layer produces the predicted docking score. The validation performance achieved r^2 values of 0.73 on the RTCB dataset.

A.6 Computing infrastructure and wall-time comparison

We trained our docking surrogate models using 4 nodes where each node contains CPUs (64 cores) and 4 A100 GPU nodes The training time for each model was approximately 3 hours. We conducted other RL experiments on a cluster that includes CPU nodes (approximately 280 cores) and GPU nodes (approximately 110 Nvidia GPUs, ranging from Titan X to A6000, set up mostly in 4- and 8-GPU configurations). Based on the computing infrastructure, we obtained the wall-time comparison in Table 6 as follows.

Methods	Total Run Time
Pretraining	36h
Finetuning	30h

Table 6: Wall-time comparison between different methods.

A.7 Hyperparameters and architectures

Table 7 provides a list of hyperparameter settings we used for our experiments.

For experimentation, we use a set of 1,280 molecules from the RTCB dataset, selecting those with docking scores in the range of $[-14, -6]$, following the range reported in (Xuefeng et al., 2024).

When computing the average normalized reward of the original molecules—without incorporating similarity constraints—we assign equal weights of 0.25 to each of the four objectives: docking, drug-likeness, synthesizability, and solubility.

Parameter	**Value**
Pretraining	
Learning rate	$5 \times e^{-5}$
Batch size	24
Optimizer	Adam
# of Epochs for Training First Phase	10
# of Epochs for Training Second Phase	10
Model # of Params	$124M$
Finetuning	
# of molecules generated K	4000
# of Epochs N	50
Batch Size	16
Learning rate	$5 \times e^{-5}$
Generation	
TopK	$[10, 15, 20]$
TopP	$[0.85, 0.9, 0.95]$
Shared	
Tamimoto Similarity Fingerprint Size	1024

Table 7: Hyperparameters for pretraining and generation.