HybridLinker: Topology-Guided Posterior Sampling for Enhanced Diversity and Validity in 3D Molecular Linker Generation

Minyeong Hwang [1] **Ziseok Lee** [2] **Kwang-soo Kim** [3 4] **Kyungsu Kim** [2] **Eunho Yang** [1]

Abstract

Linker generation is critical in drug discovery applications such as lead optimization and PROTAC design, where molecular fragments are assembled into diverse drug candidates via molecular linker. Existing methods fall into point cloud-free and point cloud-aware categories based on their use of fragments' 3D poses alongside their topologies in sampling the linker's topology. Point cloud-free models prioritize sample diversity but suffer from lower validity due to overlooking fragments' spatial constraints, while point cloud-aware models ensure higher validity but restrict diversity by enforcing strict spatial constraints. To overcome these trade-offs without additional training, we propose HybridLinker, a framework that enhances point cloud-aware inference by providing diverse bonding topologies from a pretrained point cloud-free model as guidance. At its core, we propose LinkerDPS, the first diffusion posterior sampling (DPS) method operating across point cloud-free and point cloud-aware spaces, bridging molecular topology with 3D point clouds via an energy-inspired function. By transferring the diverse sampling distribution of point cloud-free models into the point cloud-aware distribution, HybridLinker significantly surpasses baselines, improving both validity and diversity in foundational molecular design and applied drug optimization tasks, establishing a new DPS framework in the molecular domains beyond imaging.

[1]Kim Jaechul Graduate School of AI, KAIST [2]School of Transdisciplinary Innovations, Interdisciplinary Program in Bioengineering, and Interdisciplinary Program in Artificial Intelligence, Seoul National University [3]Harvard Medical School [4]McLean Hospital. Correspondence to: Kyungsu Kim <kyskim@snu.ac.kr >, Eunho Yang <eunhoy@kaist.ac.kr>.

Proceedings of the Workshop on Generative AI for Biology at the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

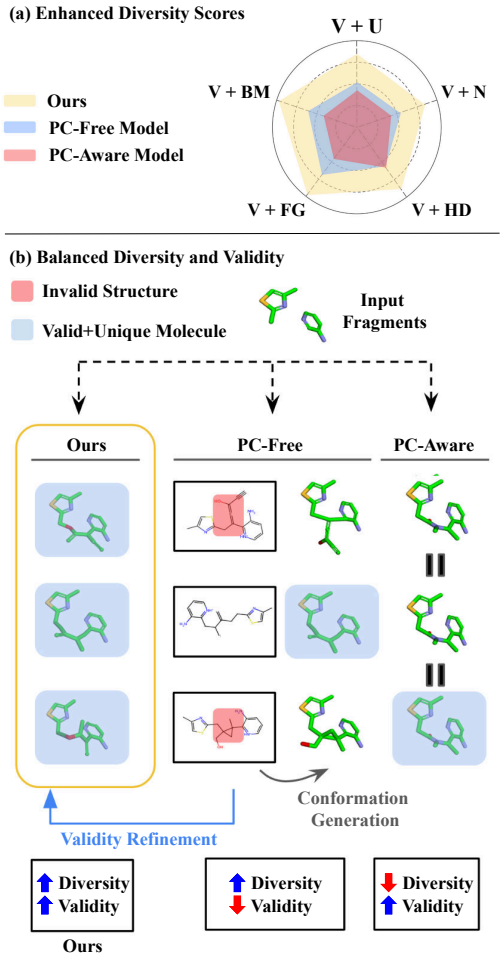

Figure 1. (a) Qualitative comparison of our method and existing pipelines on diversity of valid molecule samples measured on five metrics. (b) Trade-off between diversity and validity in baseline models and our hybrid approach overcoming these limitations.

1. Introduction

Fragment-based drug discovery strategically assembles molecular fragments to create stable and effective drug candidates. A key challenge is linker generation, which involves designing molecular linkers that connect fragments of fixed 3D poses while ensuring their validity. In this work, we categorize existing linker generation models into Point Cloud-Free (PC-Free) and Point Cloud-Aware (PC-Aware) models, distinguished by whether they account for the 3D

conformation (3D poses) of fragments when determining the bonding topology of the complete molecule. Furthermore, through a fair comparison, we highlight that achieving a balance between diversity (exploring novel topological spaces) and validity (ensuring spatial consistency between fragments and the linker) remains a significant challenge in current linker design methodologies.

Point Cloud-Free Models (↑ Diversity ↓ Validity). These models, such as FFLOM(Jin et al., 2023) and DeLinker(Imrie et al., 2020), generate diverse bonding topologies based on fragment connectivity while ignoring the 3D conformations of the given fragments. Subsequently, they predict 3D conformations that reflect the predefined fragment geometries and the sampled topology using conformation generators(Riniker & Landrum, 2015; Xu et al., 2022; 2021). By excluding 3D fragment conformations during topology generation, these methods increase entropy in the sampling distribution, thereby enhancing diversity. However, the sampled topology often fails to produce a ligand conformation that aligns well with the predefined fragment geometry, resulting in invalid, high-energy molecules

Point Cloud-Aware Models (↓ Diversity ↑ Validity). DiffLinker(Igashov et al., 2024) and 3DLinker(Huang et al., 2022) achieve high validity by incorporating fragments' 3D poses when sampling the linker's bonding topology, ensuring spatial alignment between fragments and generated linkers. They achieve this by performing generation in point cloud space, where the 3D poses of fragments are explicitly encoded as atomic coordinates, and the linker's topology is derived by from the linker's point cloud generated to align with the fragments' atomic coordinates. However, their strict spatial constraints lower sampling entropy, limiting the exploration of diverse topologies. This constrained search space increases the risk of overfitting, making it challenging to generate topologically diverse drug candidates.

To overcome the diversity-validity trade-off, we propose **HybridLinker**, a framework that integrates the strengths of PC-Free and PC-Aware models. By guiding a PC-Aware model with bonding topologies sampled from a PC-Free model, HybridLinker enhances diversity while preserving validity. At its core, we introduce LinkerDPS, the first diffusion posterior sampling (DPS) method that bridges molecular topology and point cloud space via an energy-inspired function. This approach transfers the highly diverse samples of PC-Free models into the validity-focused distribution of PC-Aware models, achieving a balanced trade-off between diversity and validity. Section 1 illustrates how HybridLinker surpasses existing methods.

We evaluate HybridLinker on the ZINC(Irwin et al., 2020) test dataset, a standard benchmark in drug discovery, demonstrating its ability to generate diverse and valid molecules from fragment inputs. By leveraging zero-shot cooperation between PC-Free and PC-Aware models, HybridLinker surpasses existing methods in diversity of valid molecules and its enhanced diversity also drives superior performance in drug-likeness optimization task, highlighting its potential as a foundational model. Moreover, HybridLinker's success validates LinkerDPS, showcasing its versatility across various domains.

Table 1. Trade-off between diversity and validity in PC-Free and PC-Aware models. The **F/A** column indicates whether the model belongs to PC-Free (F) or PC-Aware (A). Uniqueness is a metric that measures the diversity of samples, while V + U represents uniqueness counting only valid molecules.

Method	F / A	Uniquess (↑)	Validity (↑)	V + U (↑)
FFLOM	F	**0.840** (High)	0.370 (Low)	0.313
DeLinker	F	0.638 (High)	0.575 (Low)	**0.386**
DiffLinker	A	0.349 (Low)	**0.711** (High)	0.269
3DLinker	A	0.443 (Low)	0.654 (High)	0.326

Our contributions can be summarized as follows:

- We are the first to highlight the trade-off between diversity and validity in a fair comparison of Point Cloud-Free and Point Cloud-Aware linker generation models.
- We present HybridLinker, a simple yet effective framework that integrates pretrained Point Cloud-Free and Point Cloud-Aware linker generation models within a two-step generation pipeline, enabling zero-shot inference inheriting their strengths.
- We introduce LinkerDPS, the first DPS method beyond the image domain, operating across molecular topology and point cloud spaces. It bridges these domains through an energy-inspired cross-domain function, enabling effective topology-guided molecular point cloud generation. Its cross-domain guidance overcomes challenges in point cloud space by leveraging topology space as an intermediary, offering a wide range of potential applications.
- We evaluate HybridLinker across both the fundamental task of diverse drug-candidates sampling and the application-driven task of drug-likeness optimization, demonstrating its potential as a foundation model. Furthermore, validated perfomance of LinkerDPS in linker generation show its potential for broader applications, particularly in challenging point cloud tasks benefiting from topological guidance.

2. Background

In the appendix, we provide a table of notations (Appendix A), a summary of relevant tasks in molecule generation (Appendix J), and an extended definition of molecular

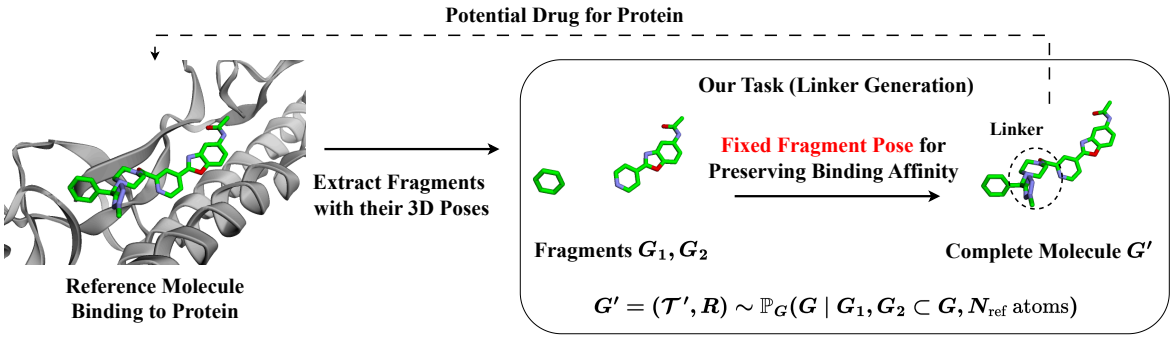

Figure 2. Problem setup for linker generation with its context in drug discovery, illustrating the fragments G_1 and G_2 embedded in a candidate molecule G'. N_{ref} is the size of reference molecule.

validity that incorporates 3D conformation alongside topological rules (Appendix C).

2.1. Molecular Representations

We provide a detailed description of molecular representations in Appendix B. Briefly, we represent a molecule as a 3D graph $G = (V, E, R)$, where $V \in \{0,1\}^{N \times A}$ denotes the one-hot encoded atom types for N atoms and A atom types, $E \in \{0,1\}^{N \times N \times B}$ represents the presence of B bond types between atom pairs, and $R \in \mathbb{R}^{N \times 3}$ specifies the 3D coordinates (i.e., 3D pose) of the atoms. We further define two key sub-representations: the **bonding topology** $\mathcal{T} = (V, E)$, which captures atomic connectivity, and the **point cloud representation** (V, R), which captures the spatial configuration of atoms. A 3D molecular graph can thus be equivalently expressed as $G = (\mathcal{T}, R)$, combining both topological and geometric information. We denote by $\mathbb{P}_G = \mathbb{P}_{\mathcal{T},R}$ the distribution (i.e., dataset) of valid 3D molecular graphs, viewed as the joint distribution over bonding topologies and 3D conformations. In prior works on molecular modeling (Hoogeboom et al., 2022; Igashov et al., 2024; Huang et al., 2022; Jin et al., 2023), the choice of molecular representation varies depending on task objectives and modeling efficiency.

Point Cloud Representation for 3D Diffusion Models. In recent 3D molecule diffusion models (Hoogeboom et al., 2022; Igashov et al., 2024; Corso et al., 2023b; Schneuing et al., 2024; Corso et al., 2023a), the one-hot encoded atom types $V \in \{0,1\}^{N \times A}$ and atomic coordinates $R \in \mathbb{R}^{N \times 3}$ are flattened and embedded into a continuous space $\mathbb{R}^d$, where $d = N \times (A + 3)$, combining both atom identity and spatial information. The diffusion process operates in this continuous space. After sampling, discrete atom types are recovered via `argmax` over V, and chemical bonds E are predicted post hoc using a bond predictor $\mathcal{E}$, resulting in the final molecular graph $G = (V, E, R)$.

2.2. Problem Setup:Linker Generation

Linker generation (Shuker et al., 1996) is a fundamental approach for high-affinity drug design, starting from fragments of a reference drug that interact with the target protein. The task is defined as generating submolecules, referred to as linkers, which connect the two fragments into a complete molecule. The 3D poses of the fragments are preserved as in the reference drug to maintain their spatial compatibility with the binding pocket, thereby contributing to the high affinity of the generated molecules.

Input. Suppose we are given two molecule fragments $G_1 = (\mathcal{T}_1, R_1)$ and $G_2 = (\mathcal{T}_2, R_2)$, where G_1 and G_2 are subgraphs of a reference molecule $G_{\text{ref}} = (\mathcal{T}_{\text{ref}}, R_{\text{ref}})$. The reference molecule G_{ref}, which contains N_{ref} atoms, is chemically valid, and belongs to a molecular graph dataset. Further, we set $\mathcal{T}_{\text{cond}} = \mathcal{T}_1 \cup \mathcal{T}_2$, and $R_{\text{cond}} = R_1 \cup R_2$.

Output. Our task is to generate new complete 3D molecular graphs $G' = (\mathcal{T}', R')$, by sampling from the following conditional distribution[1] (Figure 2):

$$G' \sim \mathbb{P}_G(G \mid \underbrace{G_1, G_2 \subset G, N_{\text{ref}} \text{ atoms}}_{\mathcal{F} = \mathcal{T}_{\text{cond}} \cap R_{\text{cond}}\text{: fragment conditions}}) \tag{1}$$

For simplicity, we use the shorthand $\mathbb{P}_G(G \mid \mathcal{F})$ to denote this joint distribution which captures the topological ($\mathcal{T}_{\text{cond}}$) and geometric ($R_{\text{cond}}$) constraints imposed by a fragment condition $\mathcal{F}$. We analogously write the marginal distributions of $\mathcal{T}, R$ as $\mathbb{P}_{\mathcal{T}}$ and $\mathbb{P}_R$, respectively.

2.3. PC-Free and PC-Aware Approaches

Existing works in linker generation divide into two approaches: Point Cloud Free (PC-Free) and Point Cloud Aware (PC-Aware), based on the utilization of fragments' 3D poses R_{cond} when sampling the topology of the output molecule $\mathcal{T}$. To clarify, PC-Free models use the following

[1]We write $G_1 \subset G'$ to denote that G' contains G_1 as its subgraph, meaning that $V_1 \subset V', E_1 \subset E', R_1 \subset R'$. We analogously write $\mathcal{T}_1 \subset \mathcal{T}$ to mean $V_1 \subset V', E_1 \subset E'$.

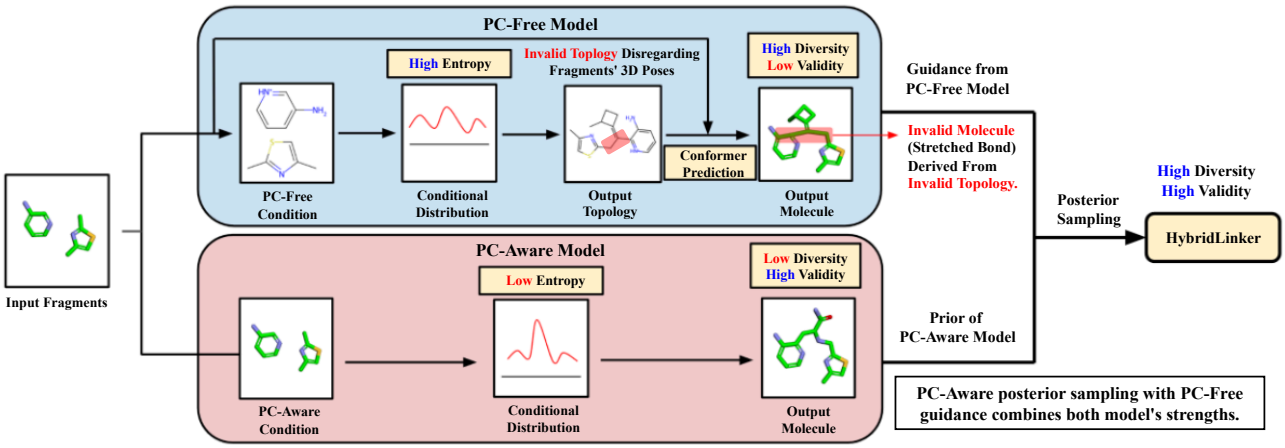

Figure 3. Comparison of generation pipelines for PC-Free and PC-Aware models. HybridLinker is designed to leverage the strengths of both approaches, inheriting high diversity from the PC-Free model and high validity from the PC-Aware model.

approximation:

$$\mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \underbrace{\mathcal{T}_{\text{cond}}, R_{\text{cond}}}_{\mathcal{F}}) \approx \mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \mathcal{T}_{\text{cond}}), \tag{2}$$

and the probability distribution of $\mathcal{T}$ in two approaches' samples are distinguished as following:

$$\begin{aligned} &\text{PC-Free} : \mathbb{P}_{\mathcal{T}|\mathcal{T}_{\text{cond}}}(\mathcal{T} \mid \mathcal{T}_{\text{cond}}) \\ &\text{PC-Aware} : \mathbb{P}_{\mathcal{T}|\mathcal{T}_{\text{cond}}, R_{\text{cond}}}(\mathcal{T} \mid \mathcal{T}_{\text{cond}}, \boxed{R_{\text{cond}}}) \end{aligned} \tag{3}$$

In the following, we explain each approach in detail with related models.

Point Cloud Free (PC-Free) Approach. PC-Free approaches (Jin et al., 2023; Imrie et al., 2020) model the conditional distribution as in (4), sequentially sampling the bonding topology $\mathcal{T}$ conditioned on **only the fragments' bonding topology** $\mathcal{T}_{\text{cond}}$ and then the 3D conformation R.

$$\begin{aligned} \mathbb{P}_G(G \mid \mathcal{F}) &\overset{\text{Bayes}}{=} \mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \boxed{\mathcal{F}}) \cdot \mathbb{P}_{R|\mathcal{T}}(R \mid \mathcal{T}, \mathcal{F}) \\ &\overset{(2)}{\approx} \mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \boxed{\mathcal{T}_{\text{cond}}}) \cdot \mathbb{P}_{R|\mathcal{T}}(R \mid \mathcal{T}, \mathcal{F}) \\ &=: \mathbb{P}_{\text{2D}}(G \mid \mathcal{F}) \end{aligned} \tag{4}$$

However, the approximation in (2), which assumes that the linker's topology is independent of R_{cond}, is often inaccurate. In reality, the topology of the linker generally depends on R_{cond}, as its topology-driven conformation must align with R_{cond} to ensure energetic stability. To emphasize the distinction between the true distribution $\mathbb{P}_G$ and its approximation, we denote the latter as $\mathbb{P}_{\text{2D}}$, referring to it as the **surrogate distribution**.

PC-Free models sample from (4) by using a neural network to approximate $\mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \mathcal{T}_{\text{cond}})$ and an off-the-shelf conformation generator $\mathcal{R}$, i.e., RDkit (Landrum, 2022), which produces the 3D conformation of the molecule including condition of R_1 and R_2 as $R' = \mathcal{R}(\mathcal{T}', R_1, R_2) \in \mathbb{R}^{N \times 3}$.

Point Cloud Aware (PC-Aware) Approach. In contrast, PC-Aware models like DiffLinker (Igashov et al., 2024) and 3DLinker (Huang et al., 2022) learn the distribution of bonding topology conditioned on both 3D poses and bonding topology of fragments, as in (3). In particular, DiffLinker and 3DLinker incorporate atom-wise distances between linker atoms and fragments to account for R_{cond} in determining both topology and conformation of linker atoms, thereby resulting in a non-approximated sampling framework (5).

$$\mathbb{P}_G(G \mid \mathcal{F}) = \underbrace{\mathbb{P}_{V,R}(V, R \mid \mathcal{F})}_{\text{Generative Model}} \cdot \underbrace{\mathbb{P}_{E|V,R}(E \mid \mathcal{F}, V, R)}_{E=\mathcal{E}(\mathcal{F},V,R)} \tag{5}$$

This formulation allows generative models to focus on sampling atomic positions while leveraging external bond inference model $\mathcal{E}$, i.e., Open Babel (O'Boyle et al., 2011).

2.4. Trade-Off in Diversity and Validity

In generating drug candidates, attaining both high validity and diversity is crucial. However, due to different sampling strategy, PC-Free and PC-Aware models exhibit an inverse relationship between diversity and validity. Specifically, PC-Free models demonstrate **high diversity but low validity**, whereas PC-Aware models achieve **high validity but low diversity**.

Diversity PC-Free models generate $\mathcal{T}$ without conditioning on R_{cond}, leading to greater diversity compared to PC-Aware models, which incorporate R_{cond} in sampling $\mathcal{T}$. This follows from the general principle that conditioning reduces entropy,

$$H(\mathcal{T} \mid \mathcal{T}_{\text{cond}}, R_{\text{cond}}) \leq H(\mathcal{T} \mid \mathcal{T}_{\text{cond}}). \tag{6}$$

Validity PC-Free models face validity issues due to the inaccurate assumptions (2) in (4). As described in (3), since

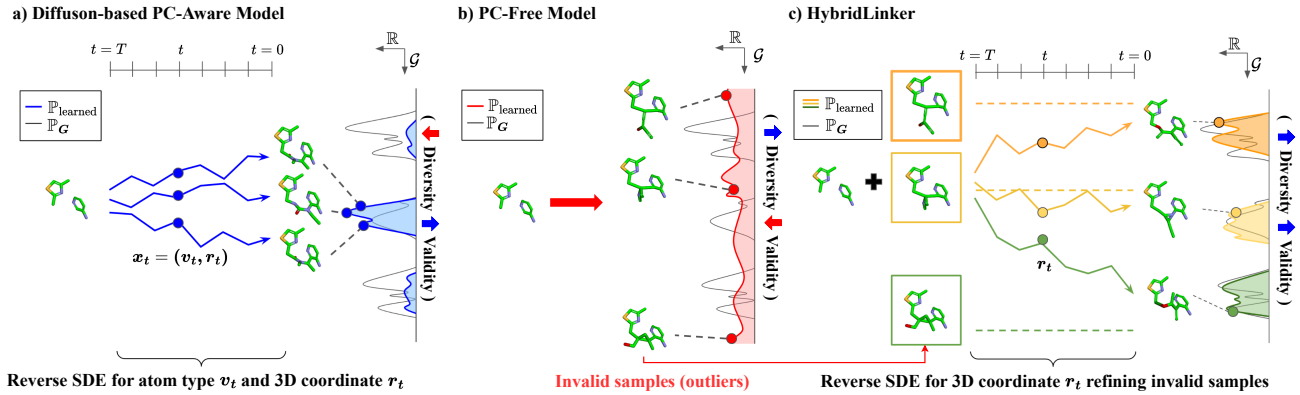

Figure 4. Conceptual comparison of sampling distributions in Diffusion-based PC-Aware, PC-Free, and HybridLinker models. (a) Diffusion-based PC-Aware models focus on validity but suffer from low diversity due to spatial constraints. (b) PC-Free models explore a broad molecular space but often generate invalid molecules. (c) HybridLinker leverages LinkerDPS to balance diversity and validity, enhancing exploration while maintaining correctness.

the 3D conformations of fragments, R_1 and R_2, are ignored when determining the bonding topology $\mathcal{T}'$, even a model that perfectly learns the target distribution $\mathbb{P}_{\text{2D}}$ will still deviate from the true molecular distribution $\mathbb{P}_G$. In contrast, PC-Aware models inherently overcome this limitation by explicitly training on the linker generation task in (5).

2.5. Diffusion-based PC-Aware Model

Following their success in a wide range of molecular generation tasks (Hoogeboom et al., 2022; Igashov et al., 2024; Corso et al., 2023b), diffusion-based models have recently become the baseline for PC-Aware linker generation (Igashov et al., 2024). It leverages the reverse SDE to iteratively denoise a randomly sampled $(v_T, r_T) \sim \mathcal{N}(0, I)$ into a valid point cloud $(v_0, r_0) = (V, R)$, conditioned on fragment input $\mathcal{F}$, over T timesteps. Specifically, the process is formulated as

$$\begin{cases} (v_T, r_T) \sim \mathcal{N}(0, I) \\ d(v_t, r_t) = \left[f(v_t, r_t, t) - g^2(t)\nabla_{v_t, r_t} \log p_t(v_t, r_t \mid \mathcal{F})\right] dt \\ \qquad\qquad + g(t) d\bar{W}_t \end{cases} \tag{7}$$

where $\bar{W}_t$ denotes the reverse Wiener process. The drift coefficient is defined as $f(v_t, r_t, t) = -\frac{1}{2}\beta_t \cdot (v_t, r_t)$, and the diffusion coefficient as $g(t) = \sqrt{\beta_t}$, with β_t being a predefined noise schedule. The term $\nabla_{v_t, r_t} \log p_t(v_t, r_t)$, referred to as the score, is estimated using a trained network s_{θ^*}. The final denoised graph x_0 is converted into a one-hot encoded molecular point cloud using `argmax`, followed by a post-hoc bond predictor, as described in Section 2.1.

3. Method

Rather than directly sampling $G' \sim \mathbb{P}_G$ in a single step, we propose a hybrid two-step approach that leverages both PC-Free and PC-Aware linker generation models.

3.1. Hybrid Approach

We develop a straightforward and intuitive pipeline that integrates two pretrained models, leveraging a novel theoretical methodology to execute the process. To adopt two types of models, we rewrite our sampling as two-step pipeline given a fragment condition $\mathcal{F} = \mathcal{T}_{\text{cond}} \cap R_{\text{cond}}$ as $\mathbb{P}_{G,\tilde{G}}(G, \tilde{G} \mid \mathcal{F}) = \mathbb{P}_{\tilde{G}}(\tilde{G} \mid \mathcal{F}) \cdot \mathbb{P}_{G|\tilde{G}}(G \mid \tilde{G}, \mathcal{F})$, where we first sample surrogate molecule $\tilde{G}$ to support sampling of G.

Now, we plug-in the PC-Free and PC-Aware models to implement the two steps pipeline. In the first stage, we set $\mathbb{P}_{\tilde{G}}$ as the PC-Free distribution $\mathbb{P}_{\text{2D}}$, which captures the topological diversity of $\mathbb{P}_G$ but also includes invalid samples. Correspondingly, $\mathbb{P}_{G|\tilde{G}}$ in the second stage becomes sampling of valid molecule equal or similar to $\tilde{G}$. To clarify, if $\tilde{G}$ is invalid, we run a molecular refinement process to obtain a similar but valid molecule; otherwise, we set G to be $\tilde{G}$. To implement the molecular refinement process, we perform posterior sampling using the validity-focused prior distribution of G, learned by the PC-Aware model, and our likelihood of $\tilde{G}$, which favors similarity with G. To summarize, we implement the pipeline as (8).

$$\mathbb{P}_{G,\text{2D}}(G, \tilde{G} \mid \mathcal{F}) = \underbrace{\mathbb{P}_{\text{2D}}(\tilde{G} \mid \mathcal{F})}_{\text{Step 1: PC-Free}} \cdot \underbrace{\mathbb{P}_{G|\text{2D}}(G \mid \tilde{G}, \mathcal{F})}_{\text{Step 2: PC-Aware}} \tag{8}$$

At a high level, our implementation can be viewed as transferring the high-entropy of the PC-Free distribution into the validity-focused PC-Aware distribution, leading to a coordinated distribution that ensures both high diversity and validity of G. Specifically, our process produces a tuple comprising a valid 3D molecule sample $G \sim \mathbb{P}_G$ and a surrogate $\tilde{G} \sim \mathbb{P}_{\text{2D}}$. We describe the algorithmic difference of HybridLinker to PC-Aware model and PC-Free model in Appendix E. Similary, Figure 4 visualizes and compares

the sampling distribution of PC-Aware, PC-Free, and HybridLinker.

3.2. Motivation for a Hybrid Approach

Experimental Comparison To validate this trade-off, we evaluate PC-Free and PC-Aware models on the ZINC dataset using standard metrics: **Validity** for validity, **Uniqueness** for diversity and **V + U** for diversity of valid molecules. The experimental setup follows Section 4.1. As shown in Table 1, PC-Free models excel in diversity, while PC-Aware models perform best in validity, confirming our hypothesis.

As observed in V + U score, the trade-off between diversity and validity significantly limits the practical utility of existing linker generation models. To overcome this limitation, we propose a **hybrid approach** that integrates the strengths of both PC-Free and PC-Aware models without requiring any training. Our hybrid pipeline leverages pretrained PC-Free and PC-Aware models cooperatively to achieve both high diversity and validity. Figure 3 illustrates the generation processes of PC-Free and PC-Aware models, along with goal of our proposed hybrid strategy of inheriting strength of two models.

3.3. Posterior Sampling From $\mathbb{P}_{G|\mathbf{2D}}$ via LinkerDPS

The challenge in sampling $\mathbb{P}_{G|2\mathrm{D}}$ lies in refining an invalid surrogate molecule into a valid one, which corresponds to the inverse problem(Gutha et al., 2024; Yang et al., 2024; Chung et al., 2024a) in the molecular domain. Since this area remains unexplored, we introduce LinkerDPS, the first DPS(Chung et al., 2024a) method designed for the molecular domain. By adopting the reverse process of the diffusion-based PC-Aware model, LinkerDPS samples a refined point cloud (V, R) favored by the pretrained prior distribution of valid point clouds.

$$p(V, R \mid \tilde{G}, \mathcal{F}) = \frac{\overbrace{p(V, R \mid \mathcal{F})}^{\text{prior}} \cdot \overbrace{p(\tilde{G} \mid V, R, \mathcal{F})}^{\text{similarity}}}{p(\tilde{G} \mid \mathcal{F})} \tag{9}$$

As shown in (9), leveraging Bayes' rule, we decompose $p(V, R \mid \tilde{G}, \mathcal{F})$ into the prior distribution $p(V, R \mid \mathcal{F})$ and the likelihood $p(\tilde{G} \mid V, R, \mathcal{F})$, which respectively ensure the validity of (V, R) and its similarity with $\tilde{G}$. Beyond the prior learned by the pretrained PC-Aware model, we design the likelihood of $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{R})$ as follows. We disregard the condition $\mathcal{F}$ in likelihood since $\tilde{G}$ always satisfies the condition. Additionally, we assume mutual independence of $\tilde{V}, \tilde{E}, \tilde{R}$ given V and R.

Likelihood of Atom. To ensure consistency between the surrogate and generated molecules, we enforce that the atoms of both molecules remain the same. Formally, the likelihood of $\tilde{V}$ is defined as

$$\mathbb{P}_{\tilde{V}|V}(\tilde{V} \mid V) := \mathbb{I}(\tilde{V} = V) \tag{10}$$

where $\mathbb{I}$ denotes the indicator function.

Likelihood of Bond and Conformation To account for the cross-domain nature of bond likelihood given a point cloud, we introduce a molecular energy-inspired function U, defined as

$$U(E^*, R^*) := \sum_{1 \le i,j \le N} \mathbb{1}_{E^*_{i,j} \neq \mathbf{0}} \|R^*_i - R^*_j\|, \tag{11}$$

which quantifies the energy of a bond-conformation system forming a molecule. We define the cross-domain similarity of $\tilde{E}$ with a given R and its likelihood as the probability that a molecule encompassing R adopts the bond $\tilde{E}$. Utilizing the Boltzmann distribution, this quantity is expressed as being proportional to $\exp(-U(\tilde{E}, R))$. Similarly, to model the likelihood of $\tilde{R}$, we employ a Gaussian kernel κ to quantify the similarity between $\tilde{R}$ and R as $\kappa(\tilde{R}, R) = e^{-\|\tilde{R}-R\|^2}$, and we define the likelihood of $\tilde{R}$ to be proportional to the quantity. Combining these components, we formulate the joint likelihood of $(\tilde{E}, \tilde{R})$ as

$$p(\tilde{E}, \tilde{R} \mid R) := \frac{1}{Z} \kappa(\tilde{R}, R) e^{-U(\tilde{E}, R)}, \tag{12}$$

where Z is the normalization constant, assumed to be independent of R.

Accordingly, the likelihood of $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{R})$ in (9) is defined as

$$p(\tilde{G} \mid V, R, \mathcal{F}) = \mathbb{I}(\tilde{V} = V) \cdot \frac{1}{Z} \kappa(\tilde{R}, R) e^{-U(\tilde{E}, R)} \tag{13}$$

LinkerDPS Formulation By (13), sampling refined point cloud (V, R) via (9) reduces to sampling R while fixing $V = \tilde{V}$, formulated as (14) where $\tilde{V} = V$ gives $p(\tilde{G} \mid R, V, \mathcal{F}) = p(\tilde{E}, \tilde{R} \mid R)$.

$$p(R \mid \tilde{G}, V, \mathcal{F}) \overset{\text{Bayes}}{\propto} p(R \mid V, \mathcal{F}) \cdot \underbrace{p(\tilde{G} \mid R, V, \mathcal{F})}_{=p(\tilde{E}, \tilde{R} \mid R) = (12)}, \tag{14}$$

where $\mathbb{I}(\tilde{V} = V) = 1$ gives $p(\tilde{G} \mid R, V, \mathcal{F}) = p(\tilde{E}, \tilde{R} \mid R)$ by (13).

To sample R from (14), we derive the reverse SDE in (15) and solve it using the pretrained score network of the PC-Aware model.

$$\begin{aligned} dr_t &= \Big[f(r_t, t) - g^2(t) \nabla_{r_t} \log p_t(r_t \mid \tilde{G}, V, \mathcal{F}) \Big] dt + g(t) d\bar{W}_t \\ &\overset{\text{Bayes}}{=} \Big[f(r_t, t) - g^2(t) (\underbrace{\nabla_{r_t} \log p_t(r_t \mid V, \mathcal{F})}_{\text{(a)}} \\ &\quad + \underbrace{\nabla_{r_t} \log p_t(\tilde{G} \mid r_t, V, \mathcal{F})}_{\text{(b)}}) \Big] dt + g(t) d\bar{W}_t \end{aligned} \tag{15}$$

Table 2. Comparison of linker generation models across diversity and validity metrics. The F/A column indicates whether the model belongs to a PC-Free (F), PC-Aware (A), or hybrid (F + A) category. The "Diversity w/ Validity" columns quantify the diversity of molecules that meet validity criteria and underscore HybridLinker's ability to sample a broad range of valid drug candidates. We also report the supporting metrics that independently measure validity and diversity without considering validity (i.e., only satisfying topological valence rules) in the columns named "Validity" and "Diversity w/o Validity," respectively.

	F/A	• Diversity w/ Validity					Validity (%)	Diversity w/o Validity	
		V+U (%)	V+N (%)	V+HD	V+FG	V+BM		Uniqueness (%)	Novelty (%)
FFLOM	F	31.25	31.32	6.38	16.37	14.48	37.07	**84.04**	**58.29**
DeLinker	F	38.57	38.53	7.32	16.20	17.33	57.44	63.78	31.63
DiffLinker	A	26.90	26.90	5.08	11.37	10.91	71.08	34.94	16.32
3DLinker	A	32.61	32.64	6.14	15.40	14.23	65.31	44.31	21.83
HybridLinker (FFLOM)	F+A	**55.10**	55.09	**10.81**	**23.92**	**24.59**	69.03	68.39	44.67
HybridLinker (DeLinker)	F+A	55.02	**55.28**	10.21	21.82	24.23	**77.27**	68.09	35.70

In (15), $f(r_t, t) = -\frac{1}{2}\beta_t r_t$ and $g(t) = \sqrt{\beta_t}$ align with the reverse SDE of diffusion-based PC-Aware models in Equation (7) and $r_0 = R$. Thus, two r_t in (15) and Equation (7) are equivalent.

We have (a) and (b) in Equation (15), not trivial to compute. To calculate (a), we apply inpainting trick (Lugmayr et al., 2022) on pretrained score network of PC-Aware model $s_{\theta^\star}(t, v_t, r_t \mid \mathcal{F}) \approx \nabla_{v_t, r_t} \log p_t(v_t, r_t|\mathcal{F})$ in Section 2.5 to implement score network additionally conditioned on V without additional training:

$$s_{\theta^\star}(t, r_t \mid V, \mathcal{F}) \approx \nabla_{r_t} \log p_t(r_t|V, \mathcal{F}) \tag{16}$$

We describe the detailed operation of this estimator in Appendix F.5. In the case of (b), since $p_t(\tilde{G} \mid r_t, V, \mathcal{F}) = \mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[p_t(\tilde{G} \mid r_0, V, \mathcal{F})\right]$, which equals to $\mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[p_t(\tilde{E}, \tilde{R} \mid r_0)\right]$ as in Equation (14), we adapt the DPS approximation(Chung et al., 2024a) for our likelihood in (12), obtaining:

$$\mathbb{E}_{r_0|r_t,V,\mathcal{F}}[p_t(\tilde{E}, \tilde{R} \mid r_0)] \approx p(\tilde{E}, \tilde{R} \mid \hat{r}), \tag{17}$$

where $\hat{r} = \mathbb{E}_{r_0|r_t,V,\mathcal{F}}[r_0]$ is tractable. Now, (b) is computable through simple additional steps. We describe the adaptation of the DPS approximation in Appendix F.3 and detailed computation of (15) using (16) and (17) in Appendix F.4.

Finally, we apply **ancestral sampling** (Ho et al., 2020a) to sample from the reverse SDE. The sampled conformation is concatenated with V to construct the point cloud, which is subsequently converted into a complete molecular structure using the post-hoc bond predictor $\mathcal{E}$, following the PC-Aware model. The full algorithm for LinkerDPS is provided in Appendix F.6.

4. Experiments

4.1. Experiment Setup

We evaluate linker generation algorithms on for 400 fragment-linker pairs from ZINC-250K(Gómez-Bombarelli et al., 2018), which are the test data in prior linker generation work (Imrie et al., 2020). Further, we run 50 times sampling for each fragment and use 20,000 samples in total for our experiments.To evaluate the validity of generated samples, we use the metric **Validity**, based on the extended definition that considers the coherence between bonding topology and 3D conformation alongside traditional topological rules, as described in Appendix C. To score the diversity of samples, we use the followings to capture diversity in multiple perspective: **Uniqueness**, **Novelty**, **V+U**, **V+N**, **V+HD**, **V+FG**, and **V+BM**. Uniqueness and Novelty accounts for all samples satisfying topological valence rule, but the others measures the diversity of valid molecules by counting only the valid samples. More details of evaluation metrics in Appendix D.3. Furthermore, we present the results of additional experiments on ablation study and descriptor optimization, in Appendix H and Appendix G.

For baseline comparisons, we use all pretrained PC-Free and PC-Aware models trained on the ZINC dataset: FFLOM(Jin et al., 2023) and DeLinker(Imrie et al., 2020) for PC-Free models, and 3DLinker(Huang et al., 2022) and DiffLinker(Igashov et al., 2024) for PC-Aware models. We evaluate two implementations of HybridLinker, HybridLinker(FFLOM) and HybridLinker(DeLinker), both utilizing DiffLinker as the diffusion-based PC-Aware model while incorporating FFLOM and DeLinker as their respective PC-Free models. For all approaches, we use ETKDGv3(Riniker & Landrum, 2015) algorithm provided by RDkit(Landrum, 2022) as conformation predictor $\mathcal{R}$, and Obabel(O'Boyle et al., 2011) as post-hoc bond predictor $\mathcal{E}$. For further implementation details, see Appendix D.2.

4.2. Results

Advanced Diversity of Valid Molecules. We run quantitative comparison of each linker generation algorithm in Table 2, demonstrating that HybridLinker successfully balance both high validity and diversity while all the baselines fail. The high scores in all diversity counting only valid moleucles (V+U, V+N, V+HD, V+FG, V+BM), indicate that its diversity is driven by meaningful substructural

Table 3. Comparison on Drug-Likeness Optimization of Reference Molecule. We report the percentage of input fragments in ZINC test data for which the model discovers the molecules surpassing the reference in drug-likeness metrics (lower QED, higher SA, lower PLogP, and their combinations).

Model	QED (%)	SA (%)	PLogP (%)	Q&S (%)	S&P (%)	Q&P (%)	Q&S&P (%)
DiffLinker	64.7	46.2	77.5	29.5	36.0	38.7	17.0
FFLOM	60.7	42.7	83.5	21.5	35.5	38.7	11.7
HybridLinker (FFLOM)	74.0 (+21.8%)	51.0 (+19.3%)	85.7 (+2.6%)	30.5 (+41.8%)	41.7 (+17.6%)	48.5 (+25.1%)	17.7 (+51.0%)
DeLinker	74.2	59.7	83.2	41.5	49.2	49.7	24.7
HybridLinker (DeLinker)	**76.7 (+3.3%)**	**61.7 (+3.3%)**	**84.7 (+1.8%)**	**43.2 (+4.2%)**	**52.2 (+6.0%)**	**54.0 (+8.5%)**	**27.0 (+9.0%)**

variations. Notably, the high V+N score highlights HybridLinker's ability to discover non-trivial linkers for given fragments. When using DeLinker as the surrogate generator, HybridLinker achieves the highest validity, while utilizing surrogates from FFLOM boosts diversity while maintaining competitive validity. Importantly, HybridLinker also maintains high Uniqueness and Novelty scores, showing that its superior diversity of valid molecules stems from a balanced enhancement of both validity and diversity.

Application: Drug-Likeness Optimization. Drug-likeness optimization, which aims to identify molecules with better drug-likeness than reference design is a direct application benefitted by sample diversity. In Table 3, we demonstrate that the high diversity of HybridLinker enhances drug-likeness optimization results. In this experiment, we evaluate each linker generation algorithm regard to the success rate of drug-likeness optimization over the fragments-reference drug pairs in ZINC test dataset. To clarify, for each pair, the algorithm "success" in optimization if the generated samples contain molecule with drug-likeness score better than the reference molecule. We utilize three widely recognized drug-likeness scores—QED, SA, and PLogP—where higher values for QED and PLogP, and lower values for SA indicate better drug-likeness, and their combinations for stricter evaluation. The result highlights HybridLinker's strong performance, specifically, HybridLinker (DeLinker) outperfoms baselines in all scores while HybridLinker (FFLOM) significantly enhanced the performance of FFLOM.

4.3. Discussion

HybridLinker as a Foundational Model As shown in Table 2, HybridLinker excels in linker generation by simultaneously achieving high diversity and validity—two fundamental pillars of drug discovery. Beyond its strong performance in core metrics, it also demonstrates exceptional results in application-driven tasks, as highlighted in Table 3. Moreover, as a zero-shot framework that seamlessly integrates pretrained PC-Free and PC-Aware models, HybridLinker remains highly adaptable to future advancements in these models, further enhancing its capabilities. These strengths establish HybridLinker as a foundational framework for linker generation, offering both versatility and long-term scalability in molecular design.

Versatility of LinkerDPS At the core of HybridLinker is LinkerDPS, the first DPS approach to enable cross-domain guidance from molecular topology to molecular point clouds. By decomposing molecule generation into two sequential tasks, LinkerDPS facilitates the seamless integration of specialized models for each domain. We anticipate that its capability will extend to a broader range of applications, particularly in tackling complex challenges at the intersection of topology and point cloud modeling by leveraging domain-specific expertise. Notably, large-scale molecule generation—one of the most pressing challenges in modern drug discovery due to its complexity—could become a new frontier for LinkerDPS, offering a scalable solution to this critical problem.

Impact of Balancing Diversity and Validity in Drug Discovery Balancing diversity and validity is a fundamental challenge in drug discovery, as it directly determines the success of downstream applications. Our experimental results further reinforce this: HybridLinker's strength in Table 2 translates to superior performance in Table 3. The findings reveal that models excelling in only one aspect—either generating diverse but invalid molecules or producing valid but overly constrained structures—lack practical utility in drug design. This highlights the necessity of striking a balance between diversity and validity for practical applications, establishing HybridLinker as a crucial framework for advancing molecular generation in drug discovery.

We provide discussion on the impact of surrogate quality on HybridLinker's performance in Appendix I.

5. Conclusion

We introduce HybridLinker, a zero-shot framework that integrates pretrained PC-Free and PC-Aware models to balance diversity and validity without additional training. At its core, LinkerDPS enables cross-domain guidance from molecular topology to point clouds, decomposing molecular sampling into sequential subtasks. Our results validate effectiveness of LinkerDPS as well as HybridLinker, and we anticipate its application in scalable large-molecule generation by breaking complex problems into manageable steps.

Acknowledgements

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) [RS-2019-II190075, Artificial Intelligence Graduate School Program (KAIST)] grant funded by the Korean government (MSIT). Additionally, this work was supported by a grant from the Institute of Information & Communications Technology Planning & Evaluation (IITP) [RS-2025-02305581], a grant from the IITP [RS2021-II211343, Artificial Intelligence Graduate School Program, Seoul National University], a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : RS-2025-02307233), and the National Research Foundation of Korea (NRF) grant (RS-2023-00242443, RS-2023-00282907), both funded by the Korean government (MSIT). This research was conducted as part of the "AI Media and Cultural Enjoyment Expansion" Project, supported by the Ministry of Science and ICT and the National IT Industry Promotion Agency (NIPA) in 2025. Kyungsu Kim is affiliated with the School of Transdisciplinary Innovations, Department of Biomedical Science, Medical Research Center, Interdisciplinary Program in Bioengineering, and Interdisciplinary Program in Artificial Intelligence at Seoul National University, Seoul, Republic of Korea.

Impact Statement

This work aims to advance drug discovery by leveraging the remarkable computational power of modern artificial intelligence. While we acknowledge both the significant benefits and potential risks associated with our approach, such as accelerating drug development for emerging viral threats like COVID-19 or inadvertently facilitating the discovery of unintended compounds, we emphasize the importance of responsible application. We hope this technology will be utilized ethically to drive positive advancements in healthcare and pharmaceutical research.

References

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., and Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nat Chem*, 2012(4):90–98, 1 2012. doi: 10.1038/nchem.1243. 2) . Published.

Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C. Diffusion posterior sampling for general noisy inverse problems, 2024a. URL https://arxiv.org/abs/2209.14687.

Chung, H., Kim, J., Park, G. Y., Nam, H., and Ye, J. C. Cfg++: Manifold-constrained classifier free guidance for diffusion models, 2024b. URL https://arxiv.org/abs/2406.08070.

Corso, G., Stärk, H., Jing, B., Barzilay, R., and Jaakkola, T. Diffdock: Diffusion steps, twists, and turns for molecular docking, 2023a. URL https://arxiv.org/abs/2210.01776.

Corso, G., Stärk, H., Jing, B., Barzilay, R., and Jaakkola, T. Diffdock: Diffusion steps, twists, and turns for molecular docking, 2023b. URL https://arxiv.org/abs/2210.01776.

Dhariwal, P. and Nichol, A. Diffusion models beat gans on image synthesis, 2021. URL https://arxiv.org/abs/2105.05233.

Ertl, P. and Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Cheminform*, 1: 8, 2009. doi: 10.1186/1758-2946-1-8.

Gao, X., Sitharam, M., and Roitberg, A. E. Bounds on the jensen gap, and implications for mean-concentrated distributions. *arXiv preprint arXiv:1712.05267*, 2017.

Geng, Z., Xie, S., Xia, Y., Wu, L., Qin, T., Wang, J., Zhang, Y., Wu, F., and Liu, T.-Y. De novo molecular generation via connection-aware motif mining, 2023. URL https://arxiv.org/abs/2302.01129.

Ghorbani, M., Gendelev, L., Beroza, P., and Keiser, M. J. Autoregressive fragment-based diffusion for pocket-aware ligand design, 2023. URL https://arxiv.org/abs/2401.05370.

Guan, J., Qian, W. W., Peng, X., Su, Y., Peng, J., and Ma, J. 3d equivariant diffusion for target-aware molecule generation and affinity prediction, 2023. URL https://arxiv.org/abs/2303.03543.

Guan, J., Peng, X., Jiang, P., Luo, Y., Peng, J., and Ma, J. Linkernet: fragment poses and linker co-design with 3d equivariant diffusion. *Advances in Neural Information Processing Systems*, 36, 2024.

Gutha, S. B. C., Vinuesa, R., and Azizpour, H. Inverse problems with diffusion models: A map estimation perspective, 2024. URL https://arxiv.org/abs/2407.20784.

Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., and Aspuru-Guzik, A. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, January 2018. ISSN 2374-7951. doi: 10.1021/acscentsci.

7b00572. URL http://dx.doi.org/10.1021/acscentsci.7b00572.

Halgren, T. A. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of Computational Chemistry*, 17(5-6):490–519, 1996. doi: https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-987X%28199604%2917%3A5/6%3C490%3A%3AAID-JCC1%3E3.0.CO%3B2-P.

Ho, J. and Salimans, T. Classifier-free diffusion guidance, 2022. URL https://arxiv.org/abs/2207.12598.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models, 2020a. URL https://arxiv.org/abs/2006.11239.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models, 2020b. URL https://arxiv.org/abs/2006.11239.

Ho, J., Salimans, T., Gritsenko, A., Chan, W., Norouzi, M., and Fleet, D. J. Video diffusion models, 2022. URL https://arxiv.org/abs/2204.03458.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3d, 2022. URL https://arxiv.org/abs/2203.17003.

Hu, X., Liu, G., Yao, Q., Zhao, Y., and Zhang, H. Hamiltonian diversity: effectively measuring molecular diversity by shortest hamiltonian circuits. *Journal of Cheminformatics*, 16(1):94, Aug 2024. ISSN 1758-2946. doi: 10.1186/s13321-024-00883-4. URL https://doi.org/10.1186/s13321-024-00883-4.

Huang, H., Sun, L., Du, B., and Lv, W. Learning joint 2-d and 3-d graph diffusion models for complete molecule generation. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

Huang, Y., Peng, X., Ma, J., and Zhang, M. 3dlinker: an e (3) equivariant variational autoencoder for molecular linker design. *arXiv preprint arXiv:2205.07309*, 2022.

Hussain, J. and Rea, C. Computationally efficient algorithm to identify matched molecular pairs (mmps) in large data sets. *Journal of chemical information and modeling*, 50 (3)::339–348, 2010.

Igashov, I., Stärk, H., Vignac, C., Schneuing, A., Satorras, V. G., Frossard, P., Welling, M., Bronstein, M., and Correia, B. Equivariant 3d-conditional diffusion model for molecular linker design. *Nature Machine Intelligence*, pp. 1–11, 2024.

Imrie, F., Bradley, A. R., van der Schaar, M., and Deane, C. M. Deep generative models for 3d linker design. *Journal of chemical information and modeling*, 60(4):1983–1995, 2020.

Irwin, J. J., Tang, K. G., Young, J., Dandarchuluun, C., Wong, B. R., Khurelbaatar, M., Moroz, Y. S., Mayfield, J., and Sayle, R. A. Zinc20—a free ultralarge-scale chemical database for ligand discovery. *Journal of chemical information and modeling*, 60(12):6065–6073, 2020.

Ji, C., Zheng, Y., Wang, R., Cai, Y., and Wu, H. Graph polish: A novel graph generation paradigm for molecular optimization, 2020. URL https://arxiv.org/abs/2008.06246.

Jin, J., Wang, D., Shi, G., Bao, J., Wang, J., Zhang, H., Pan, P., Li, D., Yao, X., Liu, H., et al. Fflom: A flow-based autoregressive model for fragment-to-lead optimization. *Journal of Medicinal Chemistry*, 66(15):10808–10823, 2023.

Jing, B., Corso, G., Chang, J., Barzilay, R., and Jaakkola, T. Torsional diffusion for molecular conformer generation, 2023. URL https://arxiv.org/abs/2206.01729.

Jo, J., Kim, D., and Hwang, S. J. Graph generation with diffusion mixture. *arXiv:2302.03596*, 2024. URL https://arxiv.org/abs/2302.03596.

Kao, C.-T., Lin, C.-T., Chou, C.-L., and Lin, C.-C. Fragment linker prediction using the deep encoder-decoder network for protacs drug design. *Journal of chemical information and modeling*, 63(10):2918–2927, 2023.

Kim, T., Seo, H., Ahn, S., and Yang, E. Rebind: Enhancing ground-state molecular conformation via force-based graph rewiring, 2024. URL https://arxiv.org/abs/2410.14696.

Kirchmair, J., Wolber, G., Laggner, C., and Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *Journal of Chemical Information and Modeling*, 46(4):1848–1861, Jul 2006. ISSN 1549-9596. doi: 10.1021/ci060084g. URL https://doi.org/10.1021/ci060084g.

Landrum, G. *Rdkit*. // www, http, 2022. rdkit.org.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, CA, Morgan Kaufmann, 2000. Stanford.

Liu, X. K. T. Graphpiece: Efficiently generating high-quality molecular graph with substructures. *arXiv preprint arXiv:2106.15098*, 2021.

Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., and Van Gool, L. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11461–11471, 2022.

Madhawa, K., Ishiguro, K., Nakago, K., and Abe, M. Graphnvp: An invertible flow model for generating molecular graphs. arXiv preprint, 2019.

Maziarz, K., Jackson-Flux, H., Cameron, P., Sirockin, F., Schneider, N., Stiefl, N., Segler, M., and Brockschmidt, M. Learning to extend molecular scaffolds with structural motifs, 2024. URL https://arxiv.org/abs/2103.03864.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. Open babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(1):33, Oct 2011. ISSN 1758-2946. doi: 10.1186/1758-2946-3-33. URL https://doi.org/10.1186/1758-2946-3-33.

Peach, M. L., Cachau, R. E., and Nicklaus, M. C. Conformational energy range of ligands in protein crystal structures: The difficult quest for accurate understanding. *Journal of Molecular Recognition*, 30, 2017. URL https://api.semanticscholar.org/CorpusID:4615425.

Peng, X., Luo, S., Guan, J., Xie, Q., Peng, J., and Ma, J. Pocket2mol: Efficient molecular sampling based on 3d protein pockets, 2022. URL https://arxiv.org/abs/2205.07249.

Peng, X., Guan, J., Liu, Q., and Ma, J. Moldiff: Addressing the atom-bond inconsistency problem in 3d molecule diffusion generation, 2023. URL https://arxiv.org/abs/2305.07508.

Riniker, S. and Landrum, G. A. Better informed distance geometry: Using what we know to improve conformation generation. *J Chem Inf Model*, 55(12):2562–74, 12 2015. doi: 10.1021/acs.jcim.5b00654. PMID: 26575315. Epub 2015 Nov 30.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models, 2022. URL https://arxiv.org/abs/2112.10752.

Schneuing, A., Harris, C., Du, Y., Didi, K., Jamasb, A., Igashov, I., Du, W., Gomes, C., Blundell, T., Lio, P., Welling, M., Bronstein, M., and Correia, B. Structure-based drug design with equivariant diffusion models, 2024. URL https://arxiv.org/abs/2210.13695.

Shi, C., Xu, M., Zhu, Z., Zhang, W., Zhang, M., and Tang, J. Graphaf: a flow-based autoregressive model for molecular graph generation, 2020. URL https://arxiv.org/abs/2001.09382.

Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. Discovering high-affinity ligands for proteins: Sar by nmr. *Science*, 274(5292):1531–1534, 1996.

Sitzmann, M., Weidlich, I. E., Filippov, I. V., Liao, C., Peach, M. L., Ihlenfeldt, W.-D., Karki, R. G., Borodina, Y. V., Cachau, R. E., and Nicklaus, M. C. Pdb ligand conformational energies calculated quantum-mechanically. *Journal of Chemical Information and Modeling*, 52(3): 739–756, Mar 2012. ISSN 1549-9596. doi: 10.1021/ci200595n. URL https://doi.org/10.1021/ci200595n.

Song, J., Meng, C., and Ermon, S. Denoising diffusion implicit models, 2022. URL https://arxiv.org/abs/2010.02502.

Torge, J., Harris, C., Mathis, S. V., and Lio, P. Diffhopp: A graph diffusion model for novel drug design via scaffold hopping, 2023. URL https://arxiv.org/abs/2308.07416.

Vignac, C., Osman, N., Toni, L., and Frossard, P. Midi: Mixed graph and 3d denoising diffusion for molecule generation, 2023. URL https://arxiv.org/abs/2302.09048.

Wang, C., Ong, H. H., Chiba, S., and Rajapakse, J. C. De novo molecule generation with graph latent diffusion model. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2121–2125, 2024a. doi: 10.1109/ICASSP48485.2024.10447480.

Wang, C., Ong, H. H., Chiba, S., and Rajapakse, J. C. De novo molecule generation with graph latent diffusion model. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2121–2125, 2024b. doi: 10.1109/ICASSP48485.2024.10447480.

Xu, M., Wang, W., Luo, S., Shi, C., Bengio, Y., Gomez-Bombarelli, R., and Tang, J. An end-to-end framework for molecular conformation generation via bilevel programming, 2021. URL https://arxiv.org/abs/2105.07246.

Xu, M., Yu, L., Song, Y., Shi, C., Ermon, S., and Tang, J. Geodiff: a geometric diffusion model for molecular conformation generation, 2022. URL https://arxiv.org/abs/2203.02923.

Xu, M., Powers, A., Dror, R., Ermon, S., and Leskovec, J. Geometric latent diffusion models for 3d molecule generation, 2023. URL https://arxiv.org/abs/2305.01140.

Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., and Yang, M.-H. Diffusion models: A comprehensive survey of methods and applications, 2024. URL https://arxiv.org/abs/2209.00796.

You, J., Liu, B., Ying, Z., Pande, V., and Leskovec, J. Graph convolutional policy network for goal-directed molecular graph generation. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/file/d60678e8f2ba9c540798ebbde31177e8-Paper.pdf.

Zhang, H., Huang, J., Xie, J., Huang, W., Yang, Y., Xu, M., Lei, J., and Chen, H. Grelinker: A graph-based generative model for molecular linker design with reinforcement and curriculum learning. *Journal of Chemical Information and Modeling*, 64(3):666–676, 2024.

Zhou, X., Cheng, X., Yang, Y., Bao, Y., Wang, L., and Gu, Q. Decompopt: Controllable and decomposed diffusion models for structure-based molecular optimization, 2024. URL https://arxiv.org/abs/2403.13829.

A. Notation Table

Table 4. Notation table for molecular graph generation and related concepts.

Symbol	Description
G	A complete 3D molecular graph $G = (V, E, R) = (\mathcal{T}, R)$, where the number of atoms, N, is usually fixed to a reference value N_{ref}.
$V \in \{0,1\}^{N \times A}$	A list of atoms in G. A is the number of atom types, and N is the number of atoms.
$E \in \{0,1\}^{N \times N \times B}$	An adjacency tensor representing chemical bonds in G.
$R \in \mathbb{R}^{N \times 3}$	A 3D coordinate matrix specifying the coordinates of each atom.
$\mathcal{T}$	The bonding topology (2D representation) $\mathcal{T} = (V, E)$ which omits spatial information.
PC	The point cloud representation (V, R), which omits bond information.
$\mathbb{P}_G = \mathbb{P}_{\mathcal{T},R}$	Distribution of valid 3D molecular graphs, or the joint distribution over bonding topology and 3D conformation.
$\mathcal{E}(\mathcal{F}, V, R)$	Post-hoc bond predictor conditioned on the fragment condition $\mathcal{F}$.
$\mathcal{E}(\emptyset, V, R)$	Unconditioned post-hoc bond predictor.
$\mathcal{R}(\mathcal{T}, R_1, R_2) \in \mathbb{R}^{N \times 3}$	Off-the-shelf conformation generator given the topology $\mathcal{T}$ and fragment conformations R_1, R_2.
$\mathcal{T}_{\text{cond}}, R_{\text{cond}}$	Given fragments $G_1 = (\mathcal{T}_1, R_1)$ and $G_2 = (\mathcal{T}_2, R_2)$: $\mathcal{T}_{\text{cond}} = \{G = (\mathcal{T}, R) : \mathcal{T}_1 \cup \mathcal{T}_2 \subset \mathcal{T}\}$, $R_{\text{cond}} = \{G = (\mathcal{T}, R) : R_1 \cup R_2 \subset R\}$.
$\mathcal{F}$	Fragment condition: $\mathcal{F} = \mathcal{T}_{\text{cond}} \cap R_{\text{cond}} = \{G = (\mathcal{T}, R) : \mathcal{T}_1 \cup \mathcal{T}_2 \subset \mathcal{T},\ R_1 \cup R_2 \subset R\}$.
$\mathbb{P}_G(G \mid \mathcal{F})$	$\mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \mathcal{F}) \cdot \mathbb{P}_{R\mid\mathcal{T}}(R \mid \mathcal{T}, \mathcal{F})$: true distribution used in PC-aware linker generation with spatial information.
$\mathbb{P}_{\text{2D}}(G \mid \mathcal{F})$	$\mathbb{P}_{\mathcal{T}}(\mathcal{T} \mid \mathcal{T}_{\text{cond}}) \cdot \mathbb{P}_{R\mid\mathcal{T}}(R \mid \mathcal{T}, \mathcal{F})$: surrogate distribution used in PC-free linker generation without spatial information.
$\mathcal{U}$	Molecular potential energy.
s_{θ^*}	Score estimator with pretrained parameter θ^*.
T	A transformation in $\mathrm{SE}(3)$.

B. Detailed Descriptions for Molecular Representations

A molecule is commonly represented as a **molecular graph**, where nodes correspond to atoms, and edges represent covalent bonds(Huang et al., 2024). We formalize this below.

Definition B.1 (3D Molecular Graph)**.** A **3D molecular graph** of N atoms is defined as a triplet $G = (V, E, R)$, where:

- $V \in \{0, 1\}^{N \times A}$ is the one-hot encoded list of **atoms** in the molecule where A is the number of atom types.
- $E \in \{0, 1\}^{N \times N \times B}$ is the set of **chemical bonds**, represented as an adjacency tensor where B is the number of bond types.
- $R \in \mathbb{R}^{N \times 3}$ denotes the **3D conformation**, specifying the spatial coordinates of each atom.

We also introduce two key sub-representations:

- The **bonding topology**, denoted as $\mathcal{T} = (V, E)$, which captures the connectivity of atoms without considering spatial information.
- The **point cloud representation**, given by (V, R), which encodes atom types and 3D positions but omits bond information.

Thus, a 3D molecular graph can be equivalently expressed as $G = (\mathcal{T}, R)$, combining both connectivity and geometry. We write $\mathbb{P}_G = \mathbb{P}_{\mathcal{T},R}$ to denote the distribution (dataset) of valid 3D molecular graphs which can be seen as the joint distribution over the bonding topology and 3D conformation.

Point Cloud Representation for 3D Molecule Diffusion Models Recent molecular diffusion models(Hoogeboom et al., 2022; Igashov et al., 2024; Corso et al., 2023b; Schneuing et al., 2024; Corso et al., 2023a) represent molecules as point clouds (V, R), disregarding explicit chemical bonds (E) during the denoising steps. In these methods, the discrete point cloud representation is embedded in a continuous space $\mathbb{R}^d$, where $d = N(A + 3)$ accounts for both atom types (one-hot encoded) and spatial coordinates. Diffusion denoising steps operate in $\mathbb{R}^d$, gradually refining the continuous representation. After sampling, the discrete atom representation is recovered using an `argmax` operation on V. Chemical bonds are inferred via a post-hoc bond predictor $\mathcal{E}$ as $E = \mathcal{E}(\mathcal{F}, V, R)$ or $\mathcal{E}(\emptyset, V, R)$ depending on fragment conditions $\mathcal{F}$, producing $G = (V, E, R)$.

C. Extended Definition of Validity

In this work, we extend the previous definition of molecule validity, focusing on each topology representation, to reflect its point cloud representations, which is a crucial consideration to make the the validity indicates its stability in 3D space where they serve a drug. We accounts for not only the valence rule in molecular topology but also the coherence of molecular point cloud with the topology. Molecular potential energy $\mathcal{U}$ is adopted to score the pair of topology and point cloud. We determine the validity of given molecule based on its strein energy, the gap of its energy the its optimal energy attained in its most stable conformation without external constraints as fragments condition.

To clarify, higher conformation energy means the molecule is unstable, also can be interpreted by Boltzmann distribution. **In our setup of linker generation, the output molecule can be unstable (high conformation energy) if the generated linker is not appropriate for the given fragments' 3D pose condition, i.e., some bonds should be stretched to connect fragments.** We choose the widely used cut-off 25(kcal / mol)(Sitzmann et al., 2012; Kirchmair et al., 2006; Peach et al., 2017) of practically allowable conformation energy as our threshold τ^{val}, and label the molecule valid if its conformation energy is lower than the τ^{val}. We use RDKit(Landrum, 2022) to compute $\mathcal{U}$ and molecule's optimal conformation.

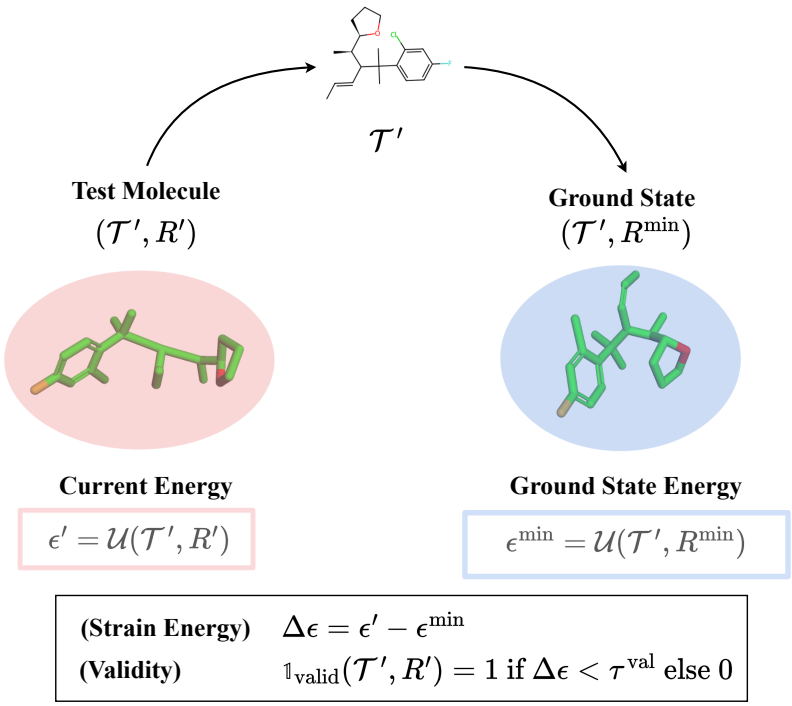

Figure 5. Extended definition of molecular validity, accounting for its point cloud feature as well as topological valence rule via utilizing strain energy.

D. Experimetal Details

D.1. Dataset Construction

Following DeLinker(Imrie et al., 2020), we select a subset of 250,000 molecules from the ZINC dataset. The energetically stable conformation of each molecule is determined using MMFF force field optimization(Halgren, 1996) implemented in RDKit(Landrum, 2022). As in prior work(Hussain & Rea, 2010), molecular fragmentation is performed by applying double bond cuts on acyclic single bonds that are not part of functional groups. Following the preprocessing steps in DeLinker(Imrie et al., 2020), we construct 418,797 fragment-linker pairs and further select 400 pairs as our test set, aligning with the test set used in DeLinker(Imrie et al., 2020). In this study, we exclusively use the test set to evaluate HybridLinker, as our method does not require additional training but instead operates in a zero-shot manner using pretrained models.

D.2. Implmentation of Baselines and HybridLinker

We compare HybridLinker with recent baselines for linker generation, falls into either PC-Free models or PC-Aware models. FFLOM(Jin et al., 2023) and DeLinker(Imrie et al., 2020) are included to represents PC-Free models. For the PC-Free models, we adopt the strong and easy-to-use conformation generation algorithm, ETKDGv3(Riniker & Landrum, 2015), to predict conformation of molecule topology they generate. For PC-Aware models, we include 3DLinker(Huang et al., 2022) and DiffLinker(Igashov et al., 2024) as baseline models.

HybridLinker is implemented utilizing both pretrained PC-Free models and PC-Aware models, respectively. As far we know, DiffLinker is the only strong PC-Aware baseline, and we adopt it in HybridLinker's second stage. As for the PC-Free model in the first step, we utilze FFLOM and DeLinker, as the two are only strong PC-Free baselines. In the experiment, we refer the PC-Free model in HybridLinker's first stage as surrogate generator. Note that we do not conduct additional training of baselines. Further, we found that if surrogate molecule $\tilde{G}$ from the first stage is already valid, skipping second stage to sample G but directly sample $G = \tilde{G}$ is effective to achieve high performance. We follow this scheme in our experiments.

Further, we use ETKDGv3(Riniker & Landrum, 2015) algorithm provided by RDkit(Landrum, 2022) as the off-the-shelf conformation predictor $\mathcal{R}$, and Obabel(O'Boyle et al., 2011) as post-hoc bond predictor $\mathcal{E}$.

D.3. Evaluation Metrics

In Section 4.2, we first evaluate the samples generated by each algorithm based on their validity and diversity. Notably, we extend the definition of **Validity** from prior works to account for both valence rules in bonding topology and energetic stability in the 3D graph. To evaluate sample diversity, we incorporate standard metrics such as **Uniqueness** and **Novelty**. However, unlike prior works that define Novelty as the fraction of novel molecules, we compute it as the fraction of molecules that are both unique and novel. Additionally, we introduce three more diversity metrics introduced in (Hu et al., 2024)—**HamDiv, FG, BM**—to capture diversity from different perspectives, specifically, molecule fingerprint, number of unique functional group, and number of unique molecular scaffolds, respectively. We follow the implementation in the repository (https://github.com/HXYfighter/HamDiv) to caculate them. We also heavily focus on calculating each diversity score counting only the valid molecules (diversity metrics for valid molecules).In the main paper, we denote them as V+U (Uniqueness), V+N (Novelty), V+HD (HamDiv), V+FG (FG), and V+BM (BM). For instance, the calculation of V+U for a given fragment pair differs from that of Uniqueness as follows:

$$\begin{aligned} \text{Uniqueness} &= \frac{|\mathcal{S}|}{N}, \\ \text{V + U} &= \frac{|\{\mathcal{T} \mid \mathcal{T} \in \mathcal{S},\ \mathcal{T} \text{ is valid}\}|}{N}, \end{aligned} \tag{18}$$

where $\mathcal{S} = \{\mathcal{T}_i\}_{i=1}^{N}$ represents the set of topologies for the N generated molecules corresponding to the fragment pair. The other diversity metrics for valid molecules are computed in a similar manner. All metrics above are first calculated for 50 samples for each fragments, and averaged over the test fragments.

We also evaluate the drug-likeness and chemical properties of the generated samples. In Section 4.2, we compare the drug-likeness of samples from each algorithm. For drug-likeness scoring, we use **QED**(Bickerton et al., 2012), **SA**(Ertl & Schuffenhauer, 2009), and **PLogP**(You et al., 2018), which measure general drug quality, synthetic accessibility, and the octanol-water partition coefficient penalized by the synthetic accessibility score, respectively.

E. Algorithmic Comparison on PC-Aware model, PC-Free model, and HybridLinker

Here, we illustrate the two-step generation pipelin of HybridLinker (Figure 6) and algorithmic distinction of HybridLinker compared to PC-Aware and PC-Free models(Algorithm 1, Algorithm 2, Algorithm 3). While PC-Aware and PC-Free models perform either PC-Aware or PC-Free inference, specializing in validity or diversity respectively, HybridLinker integrates both inference types in two-step pipeline, effectively leveraging the strengths of both approaches.

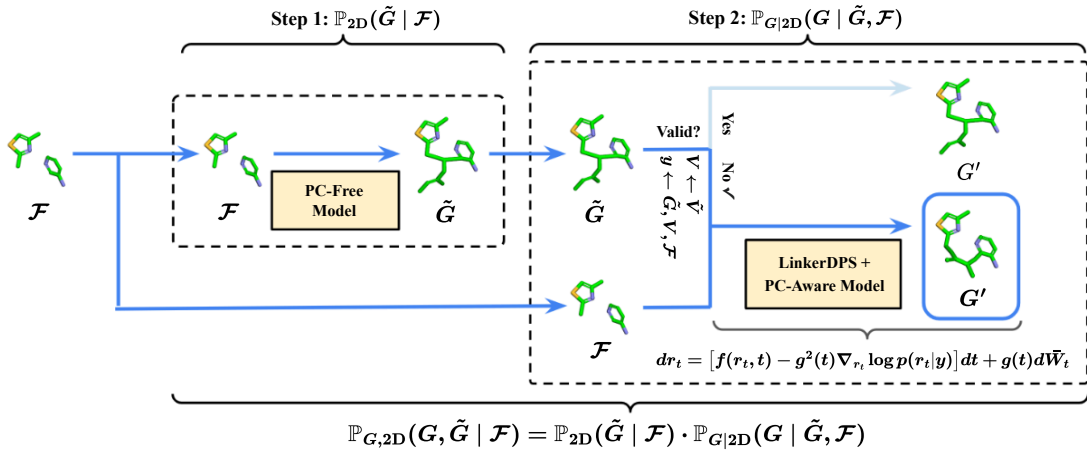

Figure 6. Two-step generation pipeline of HybridLinker. (Step 1.) diverse molecular samples are generated using pretrained PC-Free models. (Step 2.) posterior sampling is performed for each molecule from the previous step. If a molecule is invalid, LinkerDPS is applied to a diffusion-based PC-Aware model to refine it into a valid structure while maintaining similarity to the guiding molecule. Valid molecules from the first step are directly used as the final output.

Algorithm 1 PC-Aware Linker Generation Model (Diffusion-based)

INPUT fragments $G_1 = (\mathcal{T}, R_1) = (V_1, E_1, R_1)$, $G_2 = (\mathcal{T}_2, R_2) = (V_2, E_2, R_2)$, PC-Aware condition $\mathcal{F} = G_1 \cap G_2$, pretrained score estimator s_{θ^*}, bond predictor $\mathcal{E}$, variance level $\{\bar{\beta}_t = 1 - \exp(-\int_0^t \beta_u du)\}_{t=1}^T$
$x_T \sim \mathcal{N}(0, I)$ ◁ run reverse process for atom generation
for $t = T, \ldots, 1$ **do**
  $\hat{s} \leftarrow s_{\theta^*}(t, x_t \mid \mathcal{F})$ ◁ estimate score $\nabla_{x_t} \log p_t(x_t)$
  $z \sim \mathcal{N}(0, I)$
  $\hat{x} \leftarrow \frac{1}{\sqrt{1-\bar{\beta}_t}}(x_t + \bar{\beta}_t \hat{s})$ ◁ compute $\hat{x} \approx \mathbb{E}_{x_0|x_t)}[x_0]$
  $x_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{x} + \sqrt{\bar{\beta}_t} z$ ◁ update x_{t-1}
end for
$(v', R') \leftarrow x_0$
$V' \xleftarrow{\text{argmax}} v'$
$E' \leftarrow \mathcal{E}(V', R')$
$G' \leftarrow (V', E', R')$
return G'

Algorithm 2 PC-Free Linker Generation Model

INPUT fragments $G_1 = (\mathcal{T}, R_1) = (V_1, E_1, R_1)$, $G_2 = (\mathcal{T}_2, R_2) = (V_2, E_2, R_2)$, PC-Free condition $\mathcal{T}_{\text{cond}} = \mathcal{T}_1 \cap \mathcal{T}_2$, PC-Aware condition $\mathcal{F} = G_1 \cap G_2$, parameterized function for PC-Free model f_{θ^F}, conformation predictor $\mathcal{R}$
$\mathcal{T}' \leftarrow f_{\theta^F}(\mathcal{T}_{\text{cond}})$ ◁ sample topology from topology condition: $\mathcal{T}' \sim \mathbb{P}_{\mathcal{T}|\mathcal{T}_{\text{cond}}}$
$R' \leftarrow \mathcal{R}(\mathcal{T}', \mathcal{F})$ ◁ conformation prediction $R' \sim \mathbb{P}_{R|\mathcal{T}}(R', \mathcal{T}')$
$G' \leftarrow (\mathcal{T}' \mid R', \mathcal{F})$
return G'

Algorithm 3 HybridLinker

INPUT fragments $G_1 = (\mathcal{T}, R_1) = (V_1, E_1, R_1)$, $G_2 = (\mathcal{T}_2, R_2) = (V_2, E_2, R_2)$, PC-Free condition $\mathcal{T}_{\text{cond}} = \mathcal{T}_1 \cap \mathcal{T}_2$, PC-Aware condition $\mathcal{F} = G_1 \cap G_2$, pretrained score estimator for diffusion-based PC-Aware model s_{θ^*}, parameterized function for PC-Free model f_{θ^F}, conformation predictor $\mathcal{R}$, bond predictor $\mathcal{E}$
$\tilde{\mathcal{T}} \leftarrow f_{\theta^F}(\mathcal{T}_{\text{cond}})$
$\tilde{R} \leftarrow \mathcal{R}(\mathcal{T}', \mathcal{F})$
$\tilde{G} \leftarrow (\tilde{\mathcal{T}}, \tilde{R})$ ◁ 1st Stage (PC-Free Generation): $\tilde{G} \sim \mathbb{P}_{\text{2D}}(\tilde{G} \mid \mathcal{F})$
if $\tilde{G}$ is invalid **then**
  $(V', R') \leftarrow$ **LinkerDPS**$(s_{\theta^*}, \mathcal{F}, \tilde{G})$ ◁ For invalid $\tilde{G}$, run LinkerDPS with PC-Aware model to enhance validity
  $E \leftarrow \mathcal{E}(V', R')$
  $G' \leftarrow (V', E', R')$
else
  $G' \leftarrow \tilde{G}$ ◁ For valid $\tilde{G}$, set $G = \tilde{G}$
end if ◁ 2nd Stage (PC-Aware Generation): $G' \sim \mathbb{P}_{G|\text{2D}}(G' \mid \tilde{G}, \mathcal{F})$
return G'

F. Details of LinkerDPS

F.1. Overview on Diffusion Posterior Sampling

Diffusion Posterior Sampling (DPS)(Chung et al., 2024a) is a novel method for addressing noisy inverse problems using diffusion models. Diffusion models, which are typically employed for generative tasks, works in reverse process of diffusion forward pass that progressively add noise to data.

Formulation of Generative SDE. Formally, diffusion process is governed by following SDE:

$$dx_t = f(x_t, t)dt + g(t)dW_t, \quad t \in [0, T] \tag{19}$$

where $f(x_t, t) = -\frac{1}{2}\beta_t x_t$ is the drift coefficient and $g(t) = \sqrt{\beta_t}$ is the diffusion coefficient for the noise scheduler $\{\beta_t\}$. Further, W_t is the standard Wiener process and dt is an infinitesimal timestep.

To recover the data generating distribution, we can use the reverse SDE defined as

$$dx_t = \left[f(x_t, t) - g^2(t)\nabla_{x_t} \log p(x_t)\right] dt + g(t)d\bar{W}_t, \quad t \in [0, T] \tag{20}$$

where $\bar{W}$ is a reversed Wiener process flows backward from $t = 1$ to $t = 0$. Here, since the score $\nabla_{x_t} \log p(x_t)$ is intractable, the neural score network s_{θ^*} is used for computation where

$$\theta^* = \arg\min_{\theta} \mathbb{E}_{t\sim u(\epsilon,1), x_t\sim p(x_t|x_0), x_0\sim p_{\text{data}}} \left[\|s_\theta(x_t, t) - \nabla_{x_t} \log p(x_t|x_0)\|^2\right] \tag{21}$$

for a very small positive constant $\epsilon \simeq 0$.

Conditional Sampling via DPS. Now, we sample data from the posterior $p(x|y)$ given noisy measurement y, where the likelihood of measurement y given x is formulated as $p(y \mid x) \propto \exp\left(-\frac{\|y-\mathcal{A}(x)\|^2}{\sigma}\right)$ for some $\sigma \in \mathbb{R}$. The authors (Chung et al., 2024a) also show the application in several other likelihoods.

Based on the Bayes' rule, the new score $\nabla_{x_t} \log p(x_t|y)$ is decomposed into

$$\nabla_{x_t} \log p(x_t|y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y|x_t). \tag{22}$$

The proposed DPS approximation (Theorem F.3) makes the intractable term $\nabla_{x_t} \log p(y|x_t)$ computable as

$$\nabla_{x_t} \log p(y|x_t) \approx \nabla_{x_t}\hat{x}(x_t)\nabla_{x_0} \log p(y|x_0 = \hat{x}) \tag{23}$$

where $\hat{x} = \mathbb{E}_{x_0|x_t}[x_0]$. Note that $\mathbb{E}_{x_0|x_t}[x_0]$ can be estimated using the trained score network $s_{\theta^\star}$ as $\frac{1}{\sqrt{1-\bar{\beta}_t}}\left(x_t + \bar{\beta}_t s_{\theta^*}(x_t, t)\right)$ for $\bar{\beta}_t = 1 - \exp\left(-\int_0^t \beta_u du\right)$, since $x_t \sim \mathcal{N}(\sqrt{1-\bar{\beta}_t}x_0, \bar{\beta}_t I)$ and $s_{\theta^*}(x_t, t)$ estimates $(\sqrt{1-\bar{\beta}_t}x_0 - x_t)/\bar{\beta}_t$. Now, (23) is computable using the backward-propagation on tractable computation of $\hat{x}$ and $\log p(y|x_0)$. Consequently, the computable SDE for conditional generation from $p(x \mid y)$ becomes

$$dx_t = \left[f(x_t, t) - g^2(t)(\nabla_{x_t} \log p(x_t) + \nabla_{x_t}\hat{x}_0(x_t)\nabla_{x_0} \log p(y|x_0 = \hat{x}))\right] dt + g(t)d\bar{W}_t, \quad t \in [0, T] \tag{24}$$

F.2. DPS Approximation

Exploiting Definition (F.1), (Chung et al., 2024a) proposed Theorem F.3 about DPS approximation.

Definition F.1 (Jensen Gap (Gao et al., 2017))**.** Let x be a random variable with distribution $p(x)$. For some function f that may or may not be convex, the Jensen gap is defined as

$$\mathcal{J}(f, x \sim p(x)) = \mathbb{E}[f(x)] - f(\mathbb{E}[x]), \tag{25}$$

where the expectation is taken over $p(x)$.

Definition F.2. The general form of the forward model of an inverse problem can be stated as

$$y = \mathcal{A}(x_0) + \epsilon \tag{26}$$

where $y, \epsilon \in \mathbb{R}^n, x_0 \in \mathbb{R}^d$ and $\mathcal{A}(\cdot) : \mathbb{R}^d \mapsto \mathbb{R}^n$ is the forward measurement operator and ϵ is the measurement noise.

Theorem F.3. *For the measurement model defined in Definition (F.2) with $\epsilon \sim \mathcal{N}(0, \sigma^2 I)$, we have*

$$p(y \mid x_t) \approx p(y \mid x_0 = \hat{x}) \tag{27}$$

with $\hat{x} = \mathbb{E}_{x_0|x_t}[x_0]$ where the approximation error can be quantified with the Jensen gap, which is upper bounded by

$$\mathcal{J} \leq \frac{d}{\sqrt{2\pi\sigma^2}} e^{-1/2\sigma^2} \|\nabla_x \mathcal{A}(x)\| m_1, \tag{28}$$

where $\|\nabla_x \mathcal{A}(x)\| := \max_x \|\nabla_x \mathcal{A}(x)\|$ and $m_1 := \int \|x_0 - \hat{x}\| p(x_0|x_t)\, dx_0$.

F.3. Adaptation of DPS Approximation

Inspired by Theorem F.3, we propose Theorem F.4 about LinkerDPS approximation exploiting Lemma (F.5) and Proposition (F.6). Theorem F.4 and the subsequent computation

$$\mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[r_0\right] = \frac{1}{\sqrt{1-\bar{\beta}_t}}\left(r_t + \bar{\beta}_t \cdot s_{\theta^*}(t, r_t \mid V, \mathcal{F})\right), \tag{29}$$

which refers Appendix F.1 and (16), validate (17) and make (b) of (15) computable.

Theorem F.4. *(LinkerDPS Approximation) Suppose we have the cross domain measurement model,*

$$p(\tilde{E}, \tilde{R} \mid r) = \frac{1}{Z} e^{-\left[\phi_1(r|\tilde{R})/2\sigma_1^2 + \phi_2(r|\tilde{E})/\sigma_2\right]} \tag{30}$$

where $\sigma_1, \sigma_2 \in \mathbb{R}$ and ϕ_1, ϕ_2 are defined as

$$\begin{aligned} \phi_1(r \mid \tilde{R}) &:= \|\tilde{R} - r\|^2, \\ \phi_2(r \mid \tilde{E}) &:= \sum_{1 \le i,j \le N} \mathbb{1}_{\tilde{E}_{i,j} \neq \mathbf{0}} \|r_i - r_j\|. \end{aligned} \tag{31}$$

In addition, $\{r_t\}_{t\in[0,1]}$ is the diffusion process of conformation $r_0 = R$ following the forward SDE $dr_t = -\frac{1}{2}\beta_t r_t + \sqrt{\beta_t} dW$. Then, we have

$$\mathbb{E}_{r_0|r_t,V,\mathcal{F}}[p(\tilde{E}, \tilde{R} \mid r_0)] \approx p(\tilde{E}, \tilde{R} \mid \hat{r}), \tag{32}$$

where $\hat{r} = \mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[r_0\right]$ and the approximation error for (32) can be quantified with the Jensen gap $\mathcal{J}$, which is upper bounded by

$$\mathcal{J} \le L m_1, \tag{33}$$

where $L = \frac{1}{Z} \cdot \left(\frac{e^{-1/2}}{\sigma_1} + \frac{N^{3/2}}{\sigma_2}\right)$, and $m_1 := \int \|r_0 - \hat{r}\| p(r_0|r_t, V, \mathcal{F})\, dr_0$.

proof of Theorem F.4.

For $f(r) := p(\tilde{E}, \tilde{R} \mid r)$, the approximation error of (32) becomes the Jensen gap $\mathcal{F}$ as

$$\mathcal{J} = |\mathbb{E}_{r_0|r_t,V,\mathcal{F}}[f(r_0)] - f(\mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[r_0\right])| \tag{34}$$

Let $\hat{r} = \mathbb{E}_{r_0|r_t,V,\mathcal{F}}\left[r_0\right]$ and leveraging Lemma (F.5),

$$|f(r_0) - f(\hat{r})| \le L\|r_0 - \hat{r}\| \tag{35}$$

where $L = \frac{1}{Z} \cdot \left(\frac{e^{-1/2}}{\sigma_1} + \frac{N^{3/2}}{\sigma_2}\right)$. Then, according to Proposition (F.6), $\mathcal{J}$ has upper bound as

$$\mathcal{J} \le L m_1 \tag{36}$$

where $m_1 = \int \|r_0 - \hat{r}\| p(r_0|r_t, V, \mathcal{F})\, dr_0$. □

Lemma F.5. *Let $f(r) := p(\tilde{E}, \tilde{R} \mid r)$ in (30). There exists a constant L such that $\forall r, r^* \in \mathbb{R}^{3N}$,*

$$|f(r) - f(r^*)| \le L\|r - r^*\|, \tag{37}$$

where $L = \frac{1}{Z} \cdot \left(\frac{e^{-1/2}}{\sigma_1} + \frac{N^{3/2}}{\sigma_2}\right)$.

proof of Lemma (F.5).

$$
\begin{aligned}
\max_r \|\nabla_r f(r)\| &= \max_r \|f(r) \cdot \left(-\frac{1}{2\sigma_1^2} \cdot \nabla_r \phi_1(\tilde{R} \mid r) - \frac{1}{\sigma_2} \cdot \nabla_r \phi_2(\tilde{E} \mid r)\right)\| \\
&\leq \max_r \|f(r) \cdot \left(-\frac{1}{2\sigma_1^2} \cdot \nabla_r \phi_1(\tilde{R} \mid r)\right)\| + \max_r \|f(r) \cdot \left(-\frac{1}{\sigma_2} \cdot \nabla_r \phi_2(\tilde{E} \mid r)\right)\| \\
&= \max_r \|f(r) \cdot \left(-\frac{r - \tilde{R}}{\sigma_1^2}\right)\| + \max_r \|f(r) \cdot \left(-\frac{1}{\sigma_2} \cdot \nabla_r \phi_2(\tilde{E} \mid r)\right)\| \\
&\leq \frac{1}{\sigma_1} \cdot \max_r \left(f(r) \cdot \|\frac{r - \tilde{R}}{\sigma_1}\|\right) + \frac{1}{\sigma_2} \cdot \max_r \|f(r) \cdot \nabla_r \phi_2(\tilde{E} \mid r)\| \\
&\leq \frac{1}{Z} \cdot \left(\frac{e^{-1/2}}{\sigma_1} + \frac{N^{3/2}}{\sigma_2}\right)
\end{aligned}
\tag{38}
$$

where the last two inequalities are from

$$
\begin{aligned}
\max_r \left(f(r) \cdot \|\frac{r - \tilde{R}}{\sigma_1}\|\right) &\leq \max_r \left(\frac{e^{-\|r-\tilde{R}\|^2/(2\sigma_1^2)}}{Z} \cdot \|\frac{r - \tilde{R}}{\sigma_1}\|\right) \\
&= \max_{z \in \mathbb{R}^+} \frac{z \cdot e^{-z^2/2}}{Z} \\
&= \frac{e^{-1/2}}{Z}
\end{aligned}
\tag{39}
$$

and

$$
\begin{aligned}
\max_r \|f(r) \cdot \nabla_r \phi_2(\tilde{E} \mid r)\| &\leq \frac{1}{Z} \cdot \max_r \|\nabla_r \phi_2(\tilde{E} \mid r)\| \\
&\leq \frac{1}{Z} \cdot \sqrt{N \cdot (N)^2} \\
&\leq \frac{N^{3/2}}{Z}
\end{aligned}
\tag{40}
$$

Then, we have

$$
\begin{aligned}
|f(r) - f(r^*)| &\leq \max_r \|\nabla_r f(r)\| \cdot \|r - r^*\| \\
&\leq \frac{1}{Z} \cdot \left(\frac{e^{-1/2}}{\sigma_1} + \frac{N^{3/2}}{\sigma_2}\right) \cdot \|r - r^*\|
\end{aligned}
\tag{41}
$$

□

Proposition F.6 (Jensen gap upper bound(Gao et al., 2017)). *Define the absolute cenetered moment as* $m_p := \sqrt[p]{\mathbb{E}[|X - \mu|^p]}$, *and the mean as* $\mu = \mathbb{E}[X]$. *Assume that for* $\alpha > 0$, *there exists a positive number* K *such that for any* $x \in \mathbb{R}^d, |f(x) - f(\mu)| \leq K|x - \mu|^\alpha$. *Then,*

$$
\begin{aligned}
|\mathbb{E}[f(X)] - f(\mathbb{E}[X])| &\leq \int |f(X) - f(\mu)| dp(X) \\
&\leq K \int |x - \mu|^\alpha dp(X) \leq K m_\alpha^\alpha.
\end{aligned}
\tag{42}
$$

F.4. Detailed Computation of LinkerDPS

We show that reverse SDE for LinkerDPS (Equation (15)) is tractable as follows. Substituting Equation (16) and Equation (17) respectively in (a) and (b) of Equation (15) give

$$\begin{aligned} dx_t &= \Big[f(r_t,t) - g^2(t)(s_{\theta^\star}(t, r_t \mid V, \mathcal{F}) + \nabla_{r_t} \log p(\tilde{E}, \tilde{R} \mid \hat{r}))\Big]dt + g(t)d\bar{W}_t \\ &\overset{\text{Chain Rule}}{=} \Big[f(r_t,t) - g^2(t)(s_{\theta^\star}(t, r_t \mid V, \mathcal{F}) + \underbrace{\nabla_{r_t}\hat{r}}_{\texttt{autograd}} \cdot \underbrace{\nabla_{\hat{r}} \log p(\tilde{E}, \tilde{R} \mid \hat{r})}_{(44),(45)})\Big]dt + g(t)d\bar{W}_t \end{aligned} \tag{43}$$

where $\hat{r} = \mathbb{E}_{r_0|r_t,V,\mathcal{F}}[r_0] = \frac{1}{\sqrt{1-\bar{\beta}_t}}\left(r_t + \bar{\beta}_t \cdot s_{\theta^\star}(t, r_t \mid V, \mathcal{F})\right)$ for $\bar{\beta}_t = 1 - \exp\Big(-\int_0^t \beta_u du\Big)$.

We clarify that $\nabla_{r_t}\hat{r} \cdot \nabla_{\hat{r}} \log p(\tilde{E}, \tilde{R} \mid \hat{r})$ is tractable. First, $\nabla_{r_t}\hat{r}$ is computed using `autograd` after the forward calculation. The latter $\nabla_{\hat{r}} \log p(\tilde{E}, \tilde{R} \mid \hat{r})$ is expressed as

$$\nabla_{\hat{r}} \log p(\tilde{E}, \tilde{R}|\hat{r}) = -\nabla_{\hat{r}}\phi_1(\hat{r} \mid \tilde{R}) - \nabla_{\hat{r}}\phi_2(\hat{r} \mid \tilde{E}), \tag{44}$$

where $\phi_1(\hat{r} \mid \tilde{R}) := \|\tilde{R} - \hat{r}\|^2$ and $\phi_2(\hat{r} \mid \tilde{E}) := \sum_{1 \le i,j \le N} \mathbb{1}_{\tilde{E}_{i,j} \neq \mathbf{0}} \|\hat{r}_i - \hat{r}_j\|$.

Indeed, $\nabla_{\hat{r}}\phi_1(\hat{r} \mid \tilde{R})$ and $\nabla_{\hat{r}}\phi_2(\hat{r} \mid \tilde{E})$ are computed in straightforward as

$$\begin{aligned} \nabla_{\hat{r}_i}\phi_1(\hat{r} \mid \tilde{R}) &= 2(\hat{r}_i - \tilde{R}_i)^T \\ \nabla_{\hat{r}_i}\phi_2(\hat{r} \mid \tilde{E}) &= \sum_{1 \le j \le N} \frac{\mathbb{1}_{\tilde{E}_{i,j} \neq \mathbf{0}}(\hat{r}_i - \hat{r}_j)^T}{\|\hat{r}_i - \hat{r}_j\|}, \end{aligned} \tag{45}$$

which makes (44) tractable.

F.5. Conditional Score Estimator

Leveraging pretrained score estimator s_{θ^*} of PC-Aware model (Igashov et al., 2024)

$$s_{\theta^*}(t, v_t, r_t, \mid \mathcal{F}) \approx \nabla_{v_t, r_t} \log p_t(v_t, r_t \mid \mathcal{F}), \tag{46}$$

for joint sampling of atom type $v_0 = V$ and conformation $r_0 = R$, we derive a conditional score estimator for sampling of conformation alone under the fixed atom type. Specifically, the conditional score estimator predicts the score of the conditional distribution $p_t(r_t \mid V, \mathcal{F})$.

To achieve this, we treat the conditional sampling as the inpainting problem and adapt the methods of (Lugmayr et al., 2022) to 3D molecules. To clarify, we formulate the problem as the sampling of molecule $X = (V, R)$, where we know the regions for the atom types V and the remaining regions for conformation R should be sampled via backward diffusion process $\{x_t = (v_t, r_t)\}_{0 \le t \le T}$. Since every reverse step from x_t to x_{t-1} depends solely on x_t, we can alter the known regions for V as long as we keep the correct properties of the corresponding distribution, which is modeled by

$$p(x_t \mid x_0) = \mathcal{N}(\sqrt{1 - \bar{\beta}_t} x_0, \bar{\beta}_t I) \tag{47}$$

for $\bar{\beta}_t = 1 - \exp\Big(-\int_0^t \beta_u du\Big)$.

Thus, using the pretrained reverse SDE $dx_t = (-\frac{1}{2}\beta_t x_t - \beta_t s_{\theta^*}(t, v_t, r_t, \mid \mathcal{F})dt + \sqrt{\beta_t} d\bar{W}_t$ to sample only the unknown region and (47) for the known regions, we get the following reverse SDE for $\{r_t\}_{0 \le t \le T}$.

$$dr_t = \left[-\frac{1}{2}\beta_t r_t - \beta_t (m^R \odot s_{\theta^*}(t, v_t, r_t, \mid \mathcal{F}))\right] dt + \sqrt{\beta_t} d\bar{W}_t, \tag{48}$$

where $v_t \sim \mathcal{N}(\sqrt{1 - \bar{\beta}_t} V, \bar{\beta}_t I)$ and $m^R \odot$ is the masking of output scores to collect only the values in unknown regions for R.

Comparing (48) with the true reverse SDE of r_t for sampling from $p_t(r_t \mid V, \mathcal{F})$: $dr_t = \left[-\frac{1}{2}\beta_t r_t - \beta_t \nabla_{r_t} \log p(r_t \mid V, \mathcal{F}))\right] dt + \sqrt{\beta_t} d\bar{W}_t$, we can define the estimator for conditional score $\nabla_{r_t} \log p(r_t \mid V, \mathcal{F})$ appear in the latter as

$$s_{\theta^*}(t, r_t \mid V, \mathcal{F}) = m^R \odot s_{\theta^*}(t, v_t, r_t \mid \mathcal{F}), \quad \text{where} \quad v_t \sim \mathcal{N}(\sqrt{1 - \bar{\beta}_t} V, \bar{\beta}_t I) \tag{49}$$

where $m^R \odot$ is the masking of output scores to collect only the values in unknown regions for R, and it suffices the following:

$$s_{\theta^*}(t, r_t \mid V, \mathcal{F}) \approx \nabla_{r_t} \log p_t(r_t | V, \mathcal{F}). \tag{50}$$

F.6. Algorithm for LinkerDPS

We describe alorithm of LinkerDPS in Algorithm 4, which is to adopt ancestral sampling (Ho et al., 2020a) with our reverse process discussed in Section 3.3. Similarly to prior work (Chung et al., 2024a), we choose step size $\{\xi_t\}_{t=1}^T$ to be $\xi_t = \frac{\xi'}{\|\nabla_{\hat{r}}\left(\phi_1(\hat{r}|\tilde{R})+\phi_2(\hat{r}|\tilde{E})\right)\|}$ with $\xi' = 0.01$.

Algorithm 4 LinkerDPS

INPUT pretrained score network of PC-Aware model s_{θ^*}, PC-Aware fragment condition $\mathcal{F}$, surrogate molecule $\tilde{G}$, $\phi_1(r \mid \tilde{R}) := \|\tilde{R} - r\|^2$, $\phi_2(r \mid \tilde{E}) := \sum_{1\leq i,j\leq N} \mathbb{1}_{\tilde{E}_{i,j}\neq \mathbf{0}} \|r_i - r_j\|$, step size $\{\xi_t\}_{t=1}^T$, variance level $\{\bar{\beta}_t = 1 - \exp(-\int_0^t \beta_u du)\}_{t=1}^T$
$(\tilde{V}, \tilde{E}, \tilde{R}) \leftarrow \tilde{G}$
$V' \leftarrow \tilde{V}$
$v_0 \xleftarrow[\text{embedding}]{\text{continuous}} V'$
$r_T \sim \mathcal{N}(0, I)$ ◁ run reverse process on r
for $t = T, \ldots, 1$ **do**
  $\epsilon \sim \mathcal{N}(0, I)$
  $v_t \leftarrow \sqrt{1-\bar{\beta}_t} v_0 + \sqrt{\bar{\beta}_t}\epsilon$
  $\hat{s} \leftarrow s_{\theta^*}(t, v_t, r_t \mid \mathcal{F})$ ◁ estimate conditional score $\nabla_{r_t} \log p_t(r_t \mid v_0)$
  $\hat{r} \leftarrow \frac{1}{\sqrt{1-\bar{\beta}_t}}(r_t + \bar{\beta}_t \hat{s})$ ◁ compute $\hat{r} \approx \mathbb{E}_{r_0|r_t,V,\mathcal{F}}[r_0]$
  $z \sim \mathcal{N}(0, I)$
  $r'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} r_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{r} + \sqrt{\bar{\beta}_t} z$
  $r_{t-1} \leftarrow r'_{t-1} - \xi_t \nabla_{r_t}\left(\phi_1(r \mid \tilde{R}) + \phi_2(r \mid \tilde{E})\right)$ ◁ apply guidance from the $\tilde{E}$, $\tilde{R}$
end for
$R' \leftarrow r_0$
return (V', R')

G. Descriptor Optimization

In this experiment, we identify the molecule whose descriptor value is closest to a randomly sampled target value to assess the effectiveness of each generation algorithm. The target value is drawn from an approximated Gaussian distribution of the descriptors of reference molecules in the test dataset. We then select the molecule that minimizes the descriptor distance to the target value. This minimum distance serves as the evaluation metric, where a smaller value indicates that the algorithm better captures a diverse range of molecular properties. We consider five molecular descriptors: Ipc, MolLogP, MolWt, TPSA, and LabuteASA. Table 5 presents the average score of each algorithm across the fragments in the ZINC test dataset. The results highlight the strength of HybridLinkers, as they achieve the best or second-best scores across most descriptors.

Table 5. Comparison of algorithms based on descriptor optimization. The values represent the minimum MAE between the target value and values of generated samples for given fragment.

Method	**Ipc**	**MolLogP**	**MolWt**	**TPSA**	**LabuteASA**
FFLOM	0.216	0.135	0.133	0.145	0.154
DeLinker	0.218	<u>0.085</u>	0.138	0.114	0.141
DiffLinker	0.219	0.114	0.172	0.132	0.177
3DLinker	0.217	0.101	<u>0.124</u>	0.107	<u>0.135</u>
HybridLinker (FFLOM)	**0.209**	0.087	**0.117**	<u>0.099</u>	0.139
HybridLinker (DeLinker)	<u>0.213</u>	**0.064**	0.127	**0.076**	**0.134**

H. Ablation Study

The refinement process using LinkerDPS plays a central role in HybridLinker. To assess the contribution of each component in $p(\tilde{G} \mid V, R)$—namely, the likelihood of atoms, bonds, and conformation, each providing three types of guidance—we conduct an ablation study in Table 6. Variants of LinkerDPS (A-DPS) are tested by selectively removing specific likelihood components. Additionally, we introduce a version where the likelihood of atoms is modeled as a continuous distribution using a Gaussian kernel:

$$p(\tilde{v} \mid v) \propto \kappa(\tilde{v}, v), \tag{51}$$

which allows for a continuous representation of atomic properties. To implement this, we employ the following reverse process:

$$dx_t = \left[f(x_t, t) - g^2(t) \nabla_{x_t} \log p_t(x_t \mid \tilde{V}, \tilde{E}, \tilde{R}) \right] dt + g(t) d\bar{W}_t, \tag{52}$$

where $x_0 = (V, R)$. The LinkerDPS approximation for $p_t(\tilde{V}, \tilde{E}, \tilde{R} \mid x_t)$ is straightforward to adapt, and we apply this approximation to compute the score in (52).

Table 6 reveals that even variants using a single type of guidance outperform baseline methods. Performance improves further when combining two types, with the best results achieved when all three guidance components are incorporated. The marginal gains from additional guidance suggest that each component provides complementary yet overlapping information about molecular topology. Notably, relaxing the condition on atoms reduces molecular diversity, while the strict constraint on atoms ensures the preservation of surrogate atoms, reinforcing its role in maintaining molecular consistency. While validity slightly improves in the absence of inpainting, this is likely due to the repeated generation of certain valid molecules, as reflected in the decline of diversity metrics that count only valid molecules (V+U and V+N).

Table 6. Ablation study on likelihood type in LinkerDPS. The rightmost two columns represent **diversity of valid molecules**. The likelihood components used in each variant are denoted with A (Atom), E (Bond), and R (conformation).

Guidance Type	Unique	Novel	Valid	V+U	V+N
E	57.17	33.55	**76.89**	49.82	29.30
R	57.68	33.47	**76.89**	49.93	28.97
A	65.10	41.79	68.39	53.59	33.71
A + E	<u>67.20</u>	<u>44.10</u>	69.23	<u>54.59</u>	<u>35.08</u>
A (DPS) + E + R	61.61	37.76	75.65	52.12	31.61
A + E + R (Ours)	**68.45**	**44.69**	69.14	**55.10**	**35.17**

I. Additional Discussion on Experimental Results

Impact of surrogate quality on HybridLinker's performance Comparing the two HybridLinker implementations in Table 2, each utilizing FFLOM and DeLinker as surrogate generators, highlights the impact of surrogate quality on HybridLinker's performance. FFLOM exhibits higher Uniqueness and Novelty than DeLinker, and accordingly, HybridLinker(FFLOM) achieves superior scores in these metrics compared to HybridLinker(DeLinker). Conversely, HybridLinker(DeLinker) demonstrates higher validity, reflecting DeLinker's tendency to generate more valid molecules than FFLOM. A similar trend is observed in drug-likeness optimization results presented in Table 3 and Table 5, where DeLinker-produced molecules generally exhibit better drug-likeness, which corresponds to HybridLinker(DeLinker) achieving stronger drug-likeness scores. Likewise, in the molecular descriptor optimization task, implementations using surrogates that perform well for specific descriptors tend to show better optimization results for those properties. Furthermore, the effectiveness of LinkerDPS in molecule refinement is evident in Table 2. While HybridLinker(DeLinker) retains higher validity than HybridLinker(FFLOM), the validity gap between them significantly narrows. This suggests that LinkerDPS successfully refines invalid molecules into valid ones, effectively mitigating the initial disparity between FFLOM and DeLinker.

J. Related Works: Molecule Generation and Guidance on Diffusion Models

Here, we summarize the related works in molecule generation domain.

Table 7. Comparison of molecular generation tasks based on inputs and outputs. It reveals that our setup, linker generation, is a general formulation of fragment-based molecular generation. G represents a complete molecule represented as a 3D graph, while F, another 3D graph, is a partial molecule with $|F| \ll |G|$ in most cases. M denotes the binding pocket of target proteins, providing structural constraints. $\mathcal{T}$ refers to the bonding topology of a molecule, defined by atom types and chemical bonds. $\mathbf{T}$ represents an arbitrary transformation in SE(3) (e.g., rotations and translations).

Task	**Linker Gen. (Our Setup)**	**2D Linker Gen.**	**Scaffold Hop.**	**PROTAC Design**
Input	G_1, G_2	$\mathcal{T}_1, \mathcal{T}_2$	F_1, F_2	G_1, G_2
Output	$G' \mid G_1, G_2 \subset G'$	$\mathcal{T}' \mid \mathcal{T}_1, \mathcal{T}_2 \subset \mathcal{T}'$	$G' \mid F_1, F_2 \subset G'$	$G' \mid \mathbf{T}_1[G_1], \mathbf{T}_2[G_2] \subset G'$
Task	**Target-Aware Drug Design**	**Fragment Growing**	**De Novo Gen.**	**Conformation Gen.**
Input	G_1, G_2, M'	$G^\star$	None	$\mathcal{T}^\star$
Output	$G' \mid G_1, G_2 \subset G, M'$	$G' \mid G^\star \subset G'$	G'	$G' = (\mathcal{T}^\star, R') \mid \mathcal{T}^\star$

Fragment-Based Drug Design. Fragment-Based Drug Design (FBDD)(Jin et al., 2023; Igashov et al., 2024; Torge et al., 2023; Imrie et al., 2020; Huang et al., 2022) is a drug discovery approach that utilizes small molecular fragments and optimizes them into larger, more potent drug candidates. Deep learning-based FBDD encompasses a variety of tasks, each distinguished by its learning objective. **Linker Generation** (Igashov et al., 2024; Huang et al., 2022) is a fundamental task in FBDD, where two molecular fragments are connected to form a complete molecule. Similarly, **Topology Linker Generation** (Jin et al., 2023; Imrie et al., 2020; Zhang et al., 2024) focuses on linking the topological graphs of fragments to generate a complete molecular topology graph. **Scaffold Hopping**(Torge et al., 2023; Zhou et al., 2024) involves replacing the core structure of a given molecule while preserving its biological activity. **PROTAC Design**(Guan et al., 2024; Kao et al., 2023) focuses on generating molecules that incorporate fragment linkers with flexible rotation and translation in 3D space. **Fragment Growing**(Maziarz et al., 2024; Ghorbani et al., 2023) expands single small molecular fragment into larger drug-like structure. Our work specifically addresses Linker Generation, tackling the critical trade-off between diversity and validity observed in existing models. Table 7 Summarizes the tasks in FBDD along with two standard molecular generation tasks—De Novo Generation and Conformation Generation.

Molecule Generation. Molecule generation(Hoogeboom et al., 2022; Vignac et al., 2023; Wang et al., 2024b; Liu, 2021; Madhawa et al., 2019; Schneuing et al., 2024) based on deep learning plays a crucial role in drug discovery and is broadly categorized into **De Novo Molecule Generation**(Peng et al., 2023; Geng et al., 2023; Vignac et al., 2023; Jo et al., 2024), **Fragment-Based Drug Design**(Jin et al., 2023; Igashov et al., 2024; Torge et al., 2023; Imrie et al., 2020; Huang et al., 2022), **Target-Aware Drug Design**(Guan et al., 2023; Corso et al., 2023a; Schneuing et al., 2024; Peng et al., 2022), and **Conformer Generation**(Xu et al., 2022; Kim et al., 2024; Jing et al., 2023; Xu et al., 2021), based on their input and output formulations. Molecule generation can also be classified into three sub-tasks: **Topology Generation**(Shi et al., 2020; Geng et al., 2023; Jin et al., 2023; Jo et al., 2024; Ji et al., 2020; Wang et al., 2024b;a), **Point Cloud Generation**(Hoogeboom et al., 2022; Xu et al., 2023; Igashov et al., 2024; Guan et al., 2023; Schneuing et al., 2024), and **3D Graph Generation**, where deep learning models learn the distributions of molecular bonding topology, spatial coordinates, and 3D molecular graph representations, respectively. Our work falls under Fragment-Based Drug Design, introducing a hybrid approach that leverages pretrained models for topology and point cloud generation in a zero-shot manner.

Guidance on Diffusion Models. Diffusion models(Ho et al., 2020b; Song et al., 2022; Rombach et al., 2022; Ho et al., 2022; Jo et al., 2024; Hoogeboom et al., 2022; Vignac et al., 2023) have demonstrated exceptional performance across various generative tasks, including image, video, graph, and molecular generation. A recent advancement in diffusion models is **conditional generation**(Ho & Salimans, 2022; Chung et al., 2024a;b; Dhariwal & Nichol, 2021), which enables sampling from a conditional distribution based on desired properties. This is achieved by incorporating a guidance term into the backward diffusion process. To compute the guidance term, **Classifier Guidance** (Dhariwal & Nichol, 2021) employs a classifier trained to estimate the likelihood of a given property, while **Classifer-Free Guidance** (CFG) (Dhariwal & Nichol, 2021) replaces the classifier with a conditional diffusion model. **CFG++** (Chung et al., 2024b) is designed to mitigate off-manifold sampling issues, and **Diffusion Posterior Sampling** (DPS) (Chung et al., 2024a) was developed to solve

nonlinear noisy inverse problems. In this paper, we introduce the first DPS-based method for guiding diffusion models in molecular point cloud generation using molecular topology. Our approach introduces a novel energy-based function that effectively bridges topological and spatial molecular representations.