Bridging the Biophysical Gap: Holistic Environmental Awareness for 3D Linker Design

Mengwei Sun[1], **Chengwei Ai**[1], **Xiaoyi Liu**[2], **Shiqiang Ma**[3*], **Qiaozhen Meng**[4*] and **Fei Guo**[1*]
[1]Central South University
[2]Beijing University of Chinese Medicine
[3]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[4]Xiangtan University
{smw010207, ai_chery}@163.com, xiaoyi.liu@bucm.edu.cn, sq.ma@siat.ac.cn, qiaozhenm@xtu.edu.cn, guofei@csu.edu.cn

Abstract

3D molecular linker design is a critical task in structure-based drug discovery, which requires the precise synthesis of chemical bridges to connect fragments within the constrained environment of a protein binding pocket. Existing methods often suffer from environmental blindness, treating the inter-fragment space as a vacuum and yielding candidates with poor binding affinity or severe steric clashes. To address this, we propose LinkerBridge, an equivariant framework that unifies biochemical semantics with physical constraints through two innovations: a Contextual Interaction-Aware Representation module that internalizes pre-existing biochemical semantics, and a Differentiable Physical Guidance mechanism derived from Van der Waals potentials to steer generation away from collision zones. Extensive evaluations on ZINC, GEOM, and BindingMOAD benchmarks, as well as Hsp90 and JNK3 case studies, demonstrate that LinkerBridge significantly outperforms state-of-the-art methods in real-world drug discovery.

1 Introduction

Structure-based drug design (SBDD) is fundamental for modern pharmaceutical research, enabling the precise design of ligands that specifically bind to protein pockets to modulate biological pathways. Within this paradigm, 3D molecular linker design [Bancet *et al.*, 2020] has emerged as a key challenge, particularly for fragment-based drug discovery (FBDD) [Erlanson *et al.*, 2016] and the engineering of proteolysis targeting chimeras (PROTACs) [Békés *et al.*, 2022; Bai *et al.*, 2021; Farnaby *et al.*, 2019]. Unlike de novo generation [Lin *et al.*, 2025; Zhou *et al.*, 2024], it involves synthesizing a chemical bridge to connect pharmacophoric fragments while strictly maintaining their fixed 3D coordinates and bioactive poses. This makes linker design a highly constrained optimization problem within a rigid and sterically crowded microenvironment (Figure 1a). As shown in Figure 1 c, a valid linker must transcend simple geometric connectivity to function as a biochemical conduit, ensuring that its structure and composition satisfy the chemical complementarity and biological activity dictated by the surrounding environment.

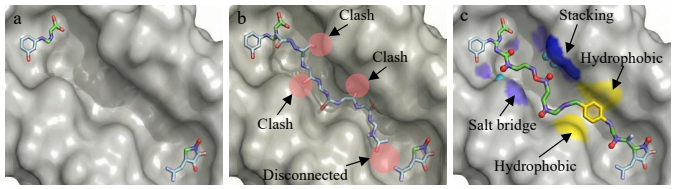

Figure 1: Illustration of difficulties in molecular linker design. (a) Input molecular fragments. (b) Failure Case: Failing to link molecular fragments and exhibiting clashes with pocket. (c) Success Case.

Although with great ability to generate molecules, diffusion models show critical issues when applied to linker design: lacking holistic environmental awareness. Specifically, they regard the space between the segments as a vacuum environment and rely solely on coordinate-based point cloud data, ignoring the complex biochemical and physical interactions between the segments. This kind of environmental blindness is fatal to the design of linkers. First, without perceiving pre-existing environmental semantics (e.g., hydrogen bonding or π-stacking), the model lacks the contextual guidance to rationally orchestrate the identity and spatial arrangement of atoms for effective molecular recognition. Some autoregressive methods [Zhung *et al.*, 2024] have attempted to utilize interaction information during the generation process, however, the interaction labels used in training and inference have a significant gap: the interactions used during training are obtained from real protein-ligand complexes, but during inference, only the protein pocket is used to predict the possible interactions. This disparity often leads to the biochemical information obtained by the model unable to effectively guide the generation process. Second, there are strict spatial constraints in protein-molecular fragments complex structure. However, most diffusion models [Peng *et al.*, 2023; Guan *et al.*, 2023] focus on point-wise denoising, that is, considering each atom separately, concentrating on minimizing the loss based on coordinates, while ignoring the surrounding

*Corresponding author: Shiqiang Ma, Qiaozhen Meng, Fei Guo

neighborhood and boundary information. This loss calculation method leads to the generated linker violating the spatial steric hindrance principle. Coupled with the essence of diffusion random denoising, the generated linker will have severe steric hindrance conflicts with the surrounding environment (Figure 1b), resulting in the generated linker being physically invalid.

To address these issues, we transform the substructure synthesis problem into a context-aware navigation task and propose LinkerBridge (Figure 2), which integrates biochemical semantics and physical constraints into an equivariant diffusion model. First, we identify the non-covalent interactions in protein-fragment or fragment-fragment pairs, and then incorporate them into the diffusion model. By pre-injecting the interaction information, the model is able to perceive the complete biochemical landscape before denoising, thereby being guided to complete and adapt this existing biochemical template during the generation process, ensuring that the linker structure is compatible with the functional requirements of the surrounding context. Furthermore, to prevent space conflicts in highly constrained environment during the linker generation process, we introduce a differentiable physical loss based on van der Waals potential. Specifically, we regard the crowded environment as a differentiable energy surface. At each step of the denoising trajectory, we inject an energy-based gradient to guide the atoms away from the high-energy collision regions and ensure that the synthesized bridges strictly follow the spatial exclusion principle.Through comprehensive experiments in linker design tasks and case studies, LinkerBridge demonstrates superior performance compared to existing approaches.

2 Related Work

Diffusion Models in Drug Discovery. The diffusion model has emerged as a crucial paradigm for 3D molecular tasks in recent years, with its applications spanning various fields such as molecule generation [Hoogeboom *et al.*, 2022; Vignac *et al.*, 2022], conformational generation [Shi *et al.*, 2021; Jing *et al.*, 2022; Xu *et al.*, 2022], and protein-ligand docking [Corso *et al.*, 2022]. Furthermore, the diffusion framework has successfully expanded to areas such as structure-based drug design [Schneuing *et al.*, 2024; Guan *et al.*, 2023; Peng *et al.*, 2023], protein design [Watson *et al.*, 2022; Ingraham *et al.*, 2023; Luo *et al.*, 2022], prediction of protein-ligand complexes [Qiao *et al.*, 2022], and generation of antibody CDRs. These methods typically generate molecular structures under the constraints of complex 3D environments (such as protein pockets or binding sites), and often employ conditional adjustment mechanisms similar to image restoration [Sohl-Dickstein *et al.*, 2015; Lugmayr *et al.*, 2022] to maintain the consistency of the structural context. Although these advancements demonstrate the powerful generative capabilities of diffusion models, there remains a challenge in adapting them to the highly constrained requirements of linker design.

Molecular Linker Design. Molecular linker design has been widely used in the fragment-based drug discovery community. The early learning-based linker design methods [Yang *et al.*, 2020; Imrie *et al.*, 2021; Sheng and Zhang, 2013] mainly employ Transformer-based or autoregressive models on SMILES [Weininger, 1988], such as DeLinker [Imrie *et al.*, 2020] and 3DLinker [Huang *et al.*, 2022]. However, these methods are difficult to achieve E(3) symmetry and are insufficient in their ability to handle large molecular fragments [Elesedy and Zaidi, 2021; Rath and Condurache, 2022] and multiple fragments. To solve this problem, DiffLinker introduces the equivariant diffusion framework, which can connect any number of segments in a single step. Nevertheless, DiffLinker [Igashov *et al.*, 2024] regards the space between fragments as a vacuum environment, lacking interactive perception of the surrounding biochemical landscape, and the generated linker is prone to physical inconsistencies in crowded spaces, resulting in the generated linker being unable to meet the key recognition patterns or causing serious spatial conflicts. The LinkerBridge proposed in this work further integrates context interaction perception and physical guidance mechanisms, thus balancing biological relevance and structural feasibility.

3 Method

3.1 Equivariant 3D Conditional Diffusion Model

We adopt the E(3) equivariant 3D conditional diffusion model [Sohl-Dickstein *et al.*, 2015] as the backbone. Our model takes the structural context u as input and generates M atoms to bridge the given fragment, thereby forming a complete connected molecule. The context u is represented in the form of a 3D atomic point cloud and can include protein-fragment complexes or only the fragment itself.

Joint Diffusion of Continuous and Discrete States

The molecule is represented by the coordinates of its M atoms $r \in \mathbb{R}^{M\times 3}$ and the corresponding one-hot encoded atomic types $h \in \mathbb{R}^{M\times nf}$. To handle the discrete characteristics of atomic types within the Gaussian diffusion framework, the model elevates h to the continuous space by adding Gaussian noise [Hoogeboom *et al.*, 2022]. The joint forward diffusion process $q(z_t|x)$ for $t = 1, \ldots, T$ is defined as follows:

$$q(z_t|x) = \mathcal{N}(z_t; \alpha_t x, \sigma_t^2 I) \tag{1}$$

where $z_t = [z_t^r, z_t^h]$. Due to the properties of Gaussian noise, the intermediate state z_t can be sampled directly from the original data x at any time step t:

$$z_t = \alpha_t x + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \tag{2}$$

The generative reverse process $p(z_{t-1}|z_t, u)$ is modeled as a Markov chain approximating the true denoising transfer:

$$p(z_{t-1}|z_t, u) = \mathcal{N}(z_{t-1}; \mu_\theta(z_t, u, t), \zeta_t^2 I) \tag{3}$$

The neural network φ_θ is trained to predict the added noise $\hat{\epsilon}_t$ [Ho *et al.*, 2020]. The estimated data $\hat{x}$ and the mean of the reverse transition μ_θ are derived as:

$$\hat{x} = \frac{1}{\alpha_t}\left(z_t - \sigma_t \varphi_\theta(z_t, u, t)\right) \tag{4}$$

$$\mu_t(x, z_t) = \frac{\bar{\alpha}_t \sigma_{t-1}^2}{\sigma_t^2} z_t + \frac{\alpha_s \bar{\sigma}_t^2}{\sigma_t^2} \hat{x} \tag{5}$$

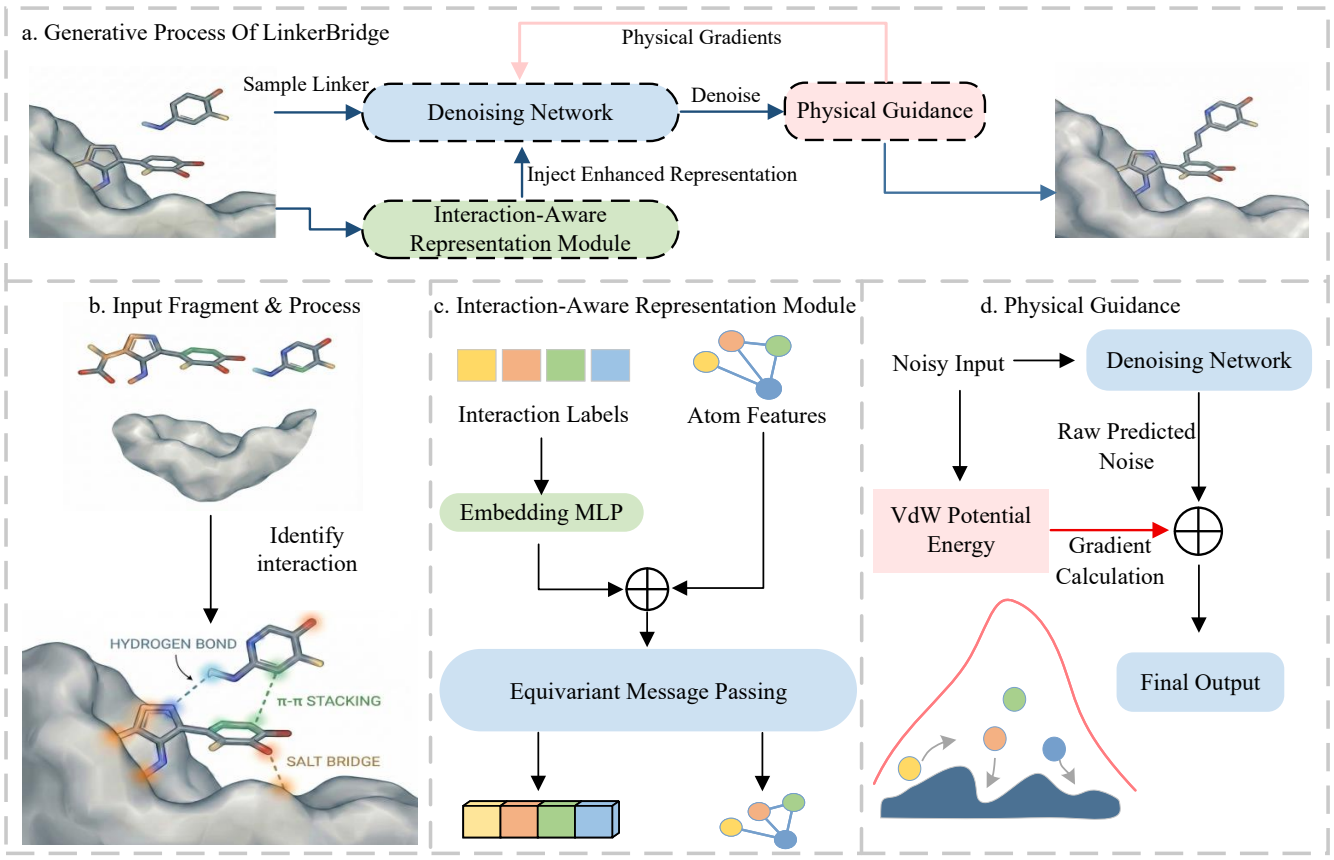

Figure 2: Overview of the LinkerBridge Framework (a) Generative process: Illustrates how the denoising network integrates enhanced representations and physical guidance to sample the connection elements. (b) Input preprocessing step: Identify the interactions (hydrogen bonds, $\pi-\pi$ stacking, salt bridges) between the fixed fragments and the protein pockets. (c) Interaction-Aware Representation Module: Inject interaction information into the message passing process to update the atomic features and coordinates. (d) Physical guidance: Calculate the van der Waals potential to guide the denoising trajectory and avoid conflicts.

E(3)-Equivariant Graph Neural Network

To ensure that the generation process is independent of the global orientation and position of the molecules, the denoising function φ_θ must satisfy $E(3)$-equivariance. We adopt the equivariant graph neural network (EGNN) [Garcia Satorras *et al.*, 2021] to construct the model, where the connected linker atoms z_t and the context u form a fully connected graph. Each layer of the network executes the following message passing update:

$$m_{ij} = \phi_e(h_i^l, h_j^l, \|r_i^l - r_j^l\|^2, t) \tag{6}$$

where the message m_{ij} depends on the scalar features and invariant distances. The coordinate displacement is only applied to the connected linker atoms with noise, while keeping the context coordinates fixed:

$$\Delta r_i = \begin{cases} \sum_{j \neq i} (r_i^l - r_j^l)\phi_{vel}(m_{ij}) & \text{if } i \in \text{linker} \\ 0 & \text{if } i \in \text{context} \end{cases} \tag{7}$$

The feature is updated as follows:

$$h_i^{l+1} = \phi_h(h_i^l, \sum_{j \neq i} m_{ij}) \tag{8}$$

To maintain translational invariance, the model centers the data around an anchor point or the centroid of the entire context $f(u)$ before sampling the initial noise.

In the target-conditioned scenario, protein pockets are represented by atomic point clouds and integrated into a fixed context u. We extend the node features by adding binary indicator tags to distinguish between ligand atoms and protein atoms. Considering that the atomic density in the protein receptor is high, to maintain computational efficiency, we convert the fully connected graph used in the original task to a distance-aware graph construction method. Specifically, connection edges are assigned based on a 4Å distance cutoff value to balance the increase in the number of nodes while retaining the local structural information necessary for spatial compatibility and binding affinity.

3.2 Contextual Interaction-Aware Representation Module

To address the interaction blindness problem in coordinate-based geometric encoding, we introduce a contextual interaction-aware representation integration module. This module is based on a core principle: molecular recognition

is mainly driven by specific physical and chemical patterns rather than merely spatial information. By explicitly characterizing and representing the pre-existing interatomic relationships, we transform the generation process from a point cloud filling task based on vacuum assumptions to a context-guided synthesis process driven by surrounding chemical environment information.

General Interaction Featurization

For each atom i within the molecular complex (including fragment atoms and the available protein pocket atoms), we define an interaction vector $v_i^{int} \in \mathbb{R}^4$. This vector serves as a high-level descriptor for local chemical information, specifically encoding four basic types of non-covalent interactions: hydrogen bonds, salt bridges, hydrophobic contacts, and π-π stacking.

In the target conditioned scenario, these features act as biochemical templates, enabling the model to perceive the specific recognition preferences of the receptor (for example, the donor sites that require hydrogen bond donors). To ensure the generalization ability of the framework in non-target design scenarios, we also apply this featurization as a structural supplement for fragment-fragment relationships. The key point is that for the linker atoms in the denoising process, the interactions v_i^{int} are set to zero vectors. This ensures that the model does not rely on the actual linker interaction during training but instead optimizes the chemical complementarity adaptively based on fixed environmental cues.

Integration into Equivariant Dynamics

To integrate this biochemical context information with the geometric features h_i in the $E(3)$-equivariant framework, we employ a multi-stage integration mechanism. Firstly, the original interaction vectors are mapped to a high-dimensional latent space through a dedicated interaction embedding layer:

$$e_i^{int} = \text{SiLU}(W_{int} v_i^{int} + b_{int}) \tag{9}$$

where the SiLU (Sigmoid Linear Unit) activation function is used to ensure the smoothness of the gradient flow.

To integrate this chemical context information with the geometric features h_i derived from the diffusion process, we adopt a splicing embedding mechanism:

$$\tilde{h}_i = \text{LayerNorm}(W_{comb}[h_i \parallel e_i^{int}] + b_{comb}) \tag{10}$$

where $\parallel$ represents the vector concatenation operation. This enhanced feature $\tilde{h}_i$ will be used as the input for the subsequent equivariant message passing layer (such as EGNN). By directly embedding such biochemical prior information into the latent state, this architecture ensures that each step of message passing and coordinate update can effectively perceive the surrounding biochemical environment. This mechanism transforms the denoised trajectory from a random walk in the coordinate space into a guided navigation process. Compared to the traditional framework that only relies on spatial geometry, LinkerBridge can thus synthesize molecules with better binding affinity and conformational stability.

3.3 Physical Constraint Guidance Based On VdW Force Loss

To ensure the structural integrity and physical validity of the generated molecules, we integrate basic physical chemistry principles into the diffusion process. This method is specifically designed to strictly follow the volume exclusion principle in environments with crowded spaces between fragments.

Formulation of the Van der Waals Potential

An effective 3D linker must navigate through a highly crowded space while avoiding spatial conflicts with predefined segments or rigid protein boundaries. We employ a differentiable potential function based on van der Waals(VdW) interactions to model these spatial constraints. For each predicted linker atom $i \in \mathcal{M}$ and each context atom $j \in \mathcal{U}$ (including fragment atoms and receptor atoms), we define the potential energy according to the repulsion term of the Lennard-Jones model:

$$\mathcal{L}_{VdW}(r) = \sum_{i=1}^{M}\sum_{j=1}^{N} \max\left(0, \epsilon\left[\left(\frac{\sigma_{ij}}{d_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{d_{ij}}\right)^{6}\right]\right) \tag{11}$$

where $d_{ij} = \|\hat{r}_i - r_j\|$ represents the Euclidean distance between the predicted linker coordinates $\hat{r}_i$ and the fixed context atom coordinates r_j Parameter σ_{ij} denotes the equilibrium distance, and its value is the sum of the van der Waals radii of the corresponding atomic types. By applying the $\max(0, \cdot)$ operator, this loss function serves as a differentiable energy barrier: it only penalizes configurations that fall within the repulsion interval ($d_{ij} < \sigma_{ij}$), thereby imposing a high energy penalty for spatial conflicts, while remaining neutral at the equilibrium distance or greater.

End-to-End Joint Optimization

The objective of the standard diffusion model is to minimize the mean square error between the added noise ϵ and the noise predicted by the network φ_θ:

$$\mathcal{L}_{diff} = \mathbb{E}_{x,\epsilon,t}\left[\|\epsilon - \varphi_\theta(z_t, u, t)\|^2\right] \tag{12}$$

Unlike traditional methods that rely on post-processing geometric optimization (often decoupled from the learned generation distribution), we directly integrate the $\mathcal{L}_{VdW}$ into the training objective. Since the predicted coordinates $\hat{r}$ are differentiable with respect to the network parameters θ through the denoising function, we use a joint loss function to optimize the model:

$$\mathcal{L}_{total} = \mathcal{L}_{diff} + \lambda \mathcal{L}_{VdW} \tag{13}$$

Within this framework, $\mathcal{L}_{diff}$ ensures that the model captures the chemical distribution of effective molecules, while λ acts as a regularization weight to control the strength of physical constraints. By backpropagating the gradient of the VdW potential energy $\nabla_{\hat{r}}\mathcal{L}_{VdW}$ to the diffusion network, LinkerBridge can effectively guide its denoising trajectory away from high-energy collision regions. This enables the internalization of physical laws into the model weights, thereby ensuring that the generated linker is not only graphically reasonable but also collision-free and structurally achievable in the crowded microenvironment of the protein pocket.

4 Experiments

4.1 Experimental Settings

Datasets. We conduct a comprehensive evaluation of LinkerBridge on three increasingly complex tasks: (i)Target-free

Dataset	Method	QED ↑	SA ↓	No. of rings ↑	Valid, %	Unique, %	Novel, %	2DFilters, %	Recovery, %	RMSD ↓	SC_RDKit ↑
ZINC	Delinker	0.64	3.11	0.21	98.3	44.2	47.1	84.9	80.2	5.48	0.49
	3DLinker	0.65	3.11	0.23	**99.3**	29.0	41.2	84.2	**94.0**	**0.10**	0.92
	DiffLinker	0.68	3.03	0.26	97.6	22.7	32.4	85.2	83.6	0.32	0.94
	LinkerBridge(ours)	**0.70**	**2.89**	0.25	98.5	24.4	33.1	**86.4**	84.8	0.31	**0.95**
	LinkerBridge(sampled size, ours)	0.68	3.14	**0.37**	93.4	**51.2**	**47.8**	86.1	83.3	0.33	0.92
CASF	Delinker	0.35	4.05	0.35	**95.7**	51.6	55.6	77.3	52.8	11.89	0.39
	DiffLinker	0.40	4.03	**0.38**	90.2	37.3	48.4	83.5	50.2	0.37	0.86
	LinkerBridge(ours)	**0.42**	**3.98**	**0.38**	91.0	39.2	44.4	**86.1**	**52.9**	0.42	**0.87**
	LinkerBridge(sampled size, ours)	0.41	4.10	0.37	72.1	**58.2**	**57.1**	85.2	51.3	**0.35**	0.85

Table 1: Results on dual-fragment design tasks. The best and second best results are marked by bold and underlined.

dual-fragment linking. Using the ZINC [Su *et al.*, 2018] and CASF [Su *et al.*, 2018] datasets, molecular fragment pairs are generated through non-cyclic bond cleavage, resulting in over 438,000 training samples. (ii) Multi-fragment Connecting. For connection tasks involving three or more segments, the GEOM dataset [Axelrod and Gomez-Bombarelli, 2022] is selected. This dataset was generated based on the MMPA/BRICS rules and contains over 285,000 segment sets. (iii) Target-Conditioned Linker Design. To test the protein-conditioned generation ability, a dataset derived from BindingMOAD [Hu *et al.*, 2005] is used. The binding pocket is defined as a $6\mathring{A}$ radius range around the ligand, and the data is classified and numbered according to enzymes. A total of 185,678 training samples are obtained to ensure the evaluation of biological generalization.

Metrics. We conduct a systematic assessment of the generation quality from three core dimensions. First, we evaluate generation quality using standard metrics, including the validity, uniqueness, and novelty of the generated molecules, along with a chemical practicality score: Quantitative Estimate of Drug-likeness (QED) [Bickerton *et al.*, 2012] and the Synthetic Accessibility score (SA) [Ertl and Schuffenhauer, 2009]. The structural fidelity is measured by the following indicators: the recovery rate of the real molecule, the root mean square deviation of atomic coordinates, and the RDKit conformational comparison score SC_{RDKit} [Putta *et al.*, 2005] that integrates geometric and chemical similarity. Second, to verify the validity of the physical constraints in the model, we introduced the van der Waals conflict index. This index is based on the van der Waals radii of atoms and quantifies the degree of spatial conflict between the generated linker and the protein-fragment environment. Third, for the protein-conditioned generation tasks, the predicted binding affinity is further evaluated using the docking score of AutoDock Vina. [Trott and Olson, 2010].

Baselines. In our experiment, we compare LinkerBridge with the sota linker design frameworks, which include the autoregressive model DeLinker [Imrie *et al.*, 2020] and 3DLinker [Huang *et al.*, 2022], as well as DiffLinker based on equivariant diffusion model [Igashov *et al.*, 2024]. For the multi-fragment connection task on the GEOM dataset, we adjust the autoregressive baseline model to enable it to handle multi-segment scenarios through the iterative connection of fragments. According to the standard experimental procedure, the 3D conformation generated by DeLinker is initially produced by the pre-trained ConfVAE model [Xu *et al.*, 2021], and then further optimized by the MMFF force field relaxation [Halgren, 1996] to obtain a more reasonable spatial structure.

4.2 Results on Fragment Pairs

In the fragment-pairs connection task, we evaluate LinkerBridge on the ZINC and CASF benchmarks that focus on the dual-fragment design task. As shown in Table 1, LinkerBridge outperforms the current optimal benchmark models in terms of chemical quality and structural fidelity. On the ZINC dataset, LinkerBridge achieves excellent drug similarity (QED: 0.70) and synthetic accessibility (SA: 2.89) scores, while maintaining 98.5% effectiveness through the built-in chemical constraints. It is worth noting that LinkerBridge demonstrates high robustness across different datasets: it achieves the best RDKit conformational alignment score (ZINC: 0.95; CASF: 0.87), with an atomic coordinate root mean square deviation of only 0.31 Å.

4.3 Results on Multiple Fragments and Pockets

To further evaluate the performance of LinkerBridge in complex scenarios, we conduct a comprehensive test on multi-segment connecting and target conditioning tasks.

Multi-Fragment Linking on GEOM

As shown in Table 2, LinkerBridge demonstrates significant advantages in the one-shot generation task of multi-fragment systems. Under the same complex scenarios, the autoregressive baseline models (such as DeLinker and 3DLinker) fail to generate effective molecules (with effectiveness ¡ 17%) or restore the reference structure, while LinkerBridge maintains high effectiveness (93.7%) and achieves an excellent recovery rate of 87.1%. Compared with DiffLinker, our model also performs better in terms of drug similarity (QED 0.51) and geometric fidelity (SC_RDKit 0.93). These results indicate that the interaction-aware representation method adopted by our model can effectively capture the long-range spatial and electronic correlation information required for simultaneously connecting three or more pharmacophores.

Target-Conditioned Design on BindingMOAD

In the target-conditioned task on the BindingMOAD dataset (Table 2), LinkerBridge outperforms DiffLinker in almost all evaluation metrics by integrating the interaction awareness

Dataset	Method	QED ↑	SA ↓	No. of rings ↑	Valid, %	Unique, %	Novel, %	2DFilters, %	Recovery, %	RMSD ↓	SC_RDKit ↑
GEOM	Delinker	**0.76**	3.59	0.00	0.1	**74.5**	—	**100.00**	0.0	—	0.42
	3DLinker	0.36	3.56	0.00	16.3	73.7	—	94.10	0.0	—	0.72
	DiffLinker	0.49	3.02	0.79	93.4	33.9	68.3	49.70	85.2	0.12	0.92
	LinkerBridge(ours)	0.51	**2.98**	0.81	**93.7**	32.9	68.4	49.50	**87.1**	0.11	**0.93**
	LinkerBridge(sampled size, ours)	0.50	3.23	**0.83**	90.4	65.5	**76.8**	58.1	68.2	**0.09**	0.88
BindingMOAD	DiffLinker	0.45	3.89	1.06	**88.7**	62.5	73.1	60.1	37.3	0.78	0.78
	LinkerBridge(ours)	**0.52**	**3.69**	**1.10**	84.1	**65.3**	**73.5**	**61.2**	**37.5**	**0.72**	**0.80**

Table 2: Results on Multi-Fragment Linking and Pockets.The best and second best results are marked by bold and underlined.

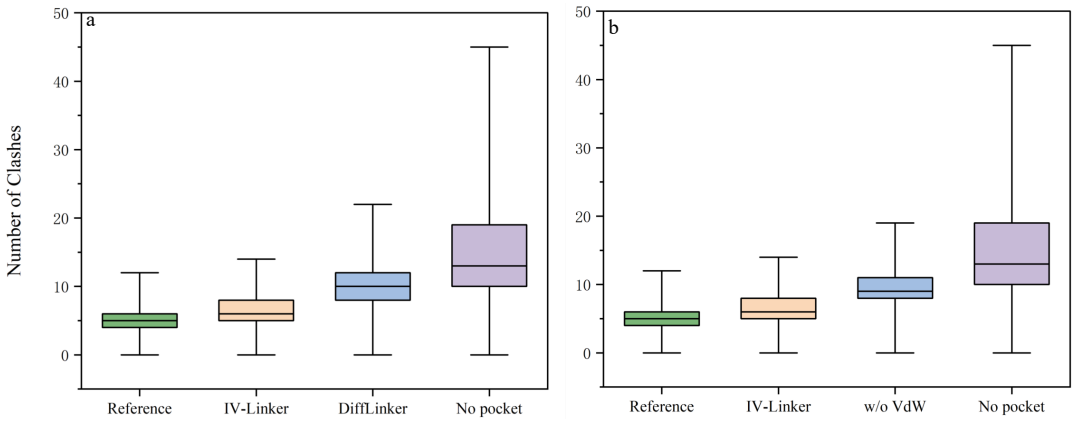

Figure 3: Analysis of steric clash. (a) Comparison between LinkerBridge and other methods. (b) Ablation study on physical constrained loss.

mechanism and the physical guidance strategy. Its drug similarity score (QED: 0.52) and pocket matching performance are excellent. The high uniqueness (65.3%) and structural fidelity (61.2%) further demonstrate that LinkerBridge can effectively screen out candidate molecules that have both physical feasibility and biochemical activity, rather than merely generating graphically reasonable but impractically unfeasible structures.

Steric Clash Analysis

To assess the structural integrity of the generated linker in the biological environment, we quantify the frequency of spatial conflicts between it and the target protein receptor. As illustrated in Figure 3a, the spatial conflict distribution of the molecules generated by LinkerBridge is highly similar to that of the reference complex, and it shows a significant improvement compared to DiffLinker. The significant reduction in spatial conflicts validates the effectiveness of the differentiable physical guidance module. By directly integrating the potential energy function based on van der Waals forces into the denoising process, LinkerBridge successfully implements the volume exclusion principle, ensuring that the synthesized linker not only aligns with the fragment geometry but also achieves the rationality of the physical structure under the rigid constraints of the protein pocket.

4.4 Case Studies

To demonstrate the practical application value of LinkerBridge in diverse drug chemistry workflows, we select two representative tasks for evaluation: fragment-based lead compound optimization for Hsp90, and scaffold hopping design for JNK3 selective inhibitors.

Fragment Linking for Hsp90

For the ATPase binding pocket of Hsp90 [Barker *et al.*, 2010], this study demonstrates the application value of LinkerBridge in fragment-based lead compound optimization (Figure 4). Our model successfully restores the structure of the reference inhibitor (Figure 4d), and generates a series of candidate molecules with Vina scores comparable to or better than the reference compound (Figure 4f). More importantly, by combining interaction-aware representation with physical guidance, LinkerBridge optimizes the protein-ligand interface and successfully restores the key hydrophobic interactions identified by the PLIP tool [Salentin *et al.*, 2015] (Figure 4e). These results confirm that LinkerBridge can transform isolated active fragments into physically feasible high-affinity lead compounds.

Scaffold Hopping for Kinase Selectivity (JNK3)

We evaluate the scaffold hopping ability of LinkerBridge at the JNK3 target site, the core challenge of this task is to achieve selective inhibition for the p38 subtype. Starting from the indazole inhibitor fragment (PDB: 3F13), LinkerBridge not only successfully restores the original indazole backbone, but also generates a highly selective aminopyrazole backbone (Figure 5b, e). The high degree of superposition between the generated structure and the experimental crystal conformation (Figure 5c, f) confirms that our model can effectively explore a diverse range of selective chemical spaces.

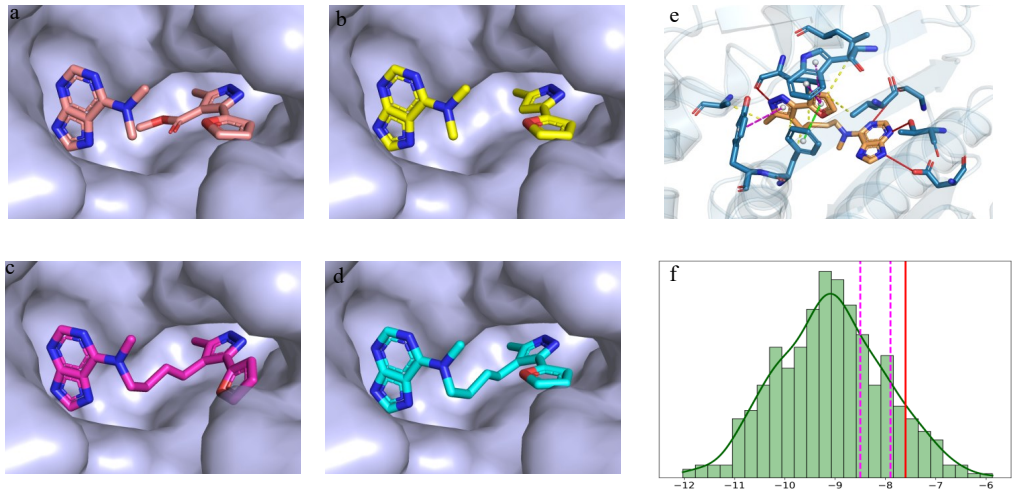

Figure 4: Hsp90 for fragment-based ligand generation. (a) Experimentally identified fragments bound to separate subsites within the ATPase pocket of Hsp90 (PDB code 3HZ1). (b) Starting fragments after removing methyl ester group. (c) Crystal structure of the inhibitor(PDB code 3HZ5). (d) Generated molecule by LinkerBridge. (e) interactions between the generated molecule and the target. (f) Distributions of vina scores for samples generated by LinkerBridge.The red line indicates the reference and the magenta lines indicate the recovered.

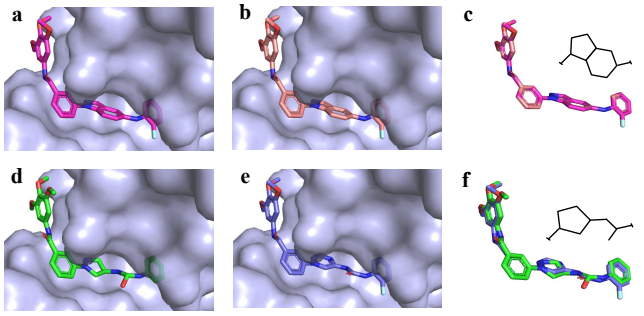

Figure 5: Exploring chemical diversity for improving selectivity of JNK inhibitors. (a) Crystal structures of compounds with indazole. (b) LinkerBridge samples that reproduce indazole. (c) Overlay of real and sampled indazole. (d) Crystal structures of compounds with aminopyrazole. (e) LinkerBridge samples that reproduce aminopyrazole. (f) Overlay of real and sampled aminopyrazole.

4.5 Ablation Study

To assess module necessity, we conduct an ablation study on ZINC (Table 3). The full LinkerBridge outperforms all variants, demonstrating that the combination with biochemical priors and physical constraints is vital for synthesis. Specifically, removing the interaction representation (v^{int}) degrades QED from 0.70 to 0.68, highlighting its role in capturing biochemical context for drug-like generation.

The steric clash analysis (Figure 3b) confirms that removing the VdW loss term would significantly increase the median number of conflicts and produce extreme outliers. When LinkerBridge is able to suppress the degree of atomic overlap to the level of the reference complex, the model without VdW terms exhibits severe physical conflicts within the rigid protein pocket. This validates the necessity of the VdW potential energy for enforcing the principle of volume exclusion, it ensures that the generated linker structure is physically feasible, rather than merely being a geometrically reasonable configuration.

Method	QED ↑	SA ↓	Valid, %	Recovery, %	SC_RDKit↑
LinkerBridge	**0.70**	**2.89**	**98.5**	**84.8**	**0.95**
w/o VdW	0.69	2.94	98.1	84.1	0.94
w/o interaction	0.68	2.98	97.9	83.9	0.94

Table 3: Ablation Study on Key Modules Of LinkerBridge.

5 Conclusion

This study proposes LinkerBridge, a novel equivariant diffusion framework, aiming to overcome the interaction blindness and physical inconsistency problems existing in the design of 3D molecular linker. By introducing a contextual interaction-aware representation module, our model can effectively capture the key biochemical recognition patterns (such as hydrogen bonding and π-stacking interactions) in its surrounding environment. Furthermore, we implement a physical constraint-guided mechanism through a differentiable VdW loss function. This mechanism dynamically adheres to the volume exclusion principle during the denoising process, thereby avoiding spatial conflicts. Experimental results on multiple benchmarks show that LinkerBridge significantly outperforms current methods in generating linkers with high drug similarity, synthetic accessibility, and structural fidelity. Case studies on Hsp90 and JNK3 further validate the practical application value of our model in lead compound optimization and scaffold hopping. By integrating high-fidelity characterization with physical prior knowledge, LinkerBridge bridges the gap between statistical rationality and physical feasibility, providing a powerful tool for structure-based drug discovery research.

Ethical Statement

There are no ethical issues.

Acknowledgments

National Natural Science Foundation of China (Grants No. 62322215, 62532017, 62402488 and 62502050), and Natural Science Foundation of Hunan Province (Grants No. 2026JJ30018 and 2025JJ60387).

References

[Axelrod and Gomez-Bombarelli, 2022] Simon Axelrod and Rafael Gomez-Bombarelli. Geom, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1):185, 2022.

[Bai *et al.*, 2021] Nan Bai, Sven A Miller, Grigorii V Andrianov, Max Yates, Palani Kirubakaran, and John Karanicolas. Rationalizing protac-mediated ternary complex formation using rosetta. *Journal of chemical information and modeling*, 61(3):1368–1382, 2021.

[Bancet *et al.*, 2020] Alexandre Bancet, Claire Raingeval, Thierry Lomberget, Marc Le Borgne, Jean-François Guichou, and Isabelle Krimm. Fragment linking strategies for structure-based drug design. *Journal of medicinal chemistry*, 63(20):11420–11435, 2020.

[Barker *et al.*, 2010] John J Barker, Oliver Barker, Stephen M Courtney, Mihaly Gardiner, Thomas Hesterkamp, Osamu Ichihara, Owen Mather, Christian AGN Montalbetti, Annett Müller, Mario Varasi, et al. Discovery of a novel hsp90 inhibitor by fragment linking. *ChemMedChem*, 5(10):1697–1700, 2010.

[Békés *et al.*, 2022] Miklós Békés, David R Langley, and Craig M Crews. Protac targeted protein degraders: the past is prologue. *Nature reviews Drug discovery*, 21(3):181–200, 2022.

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

[Corso *et al.*, 2022] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv:2210.01776*, 2022.

[Elesedy and Zaidi, 2021] Bryn Elesedy and Sheheryar Zaidi. Provably strict generalisation benefit for equivariant models. In *International conference on machine learning*, pages 2959–2969. PMLR, 2021.

[Erlanson *et al.*, 2016] Daniel A Erlanson, Stephen W Fesik, Roderick E Hubbard, Wolfgang Jahnke, and Harren Jhoti. Twenty years on: the impact of fragments on drug discovery. *Nature reviews Drug discovery*, 15(9):605–619, 2016.

[Ertl and Schuffenhauer, 2009] Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of cheminformatics*, 1(1):8, 2009.

[Farnaby *et al.*, 2019] William Farnaby, Manfred Koegl, Michael J Roy, Claire Whitworth, Emelyne Diers, Nicole Trainor, David Zollman, Steffen Steurer, Jale Karolyi-Oezguer, Carina Riedmueller, et al. Baf complex vulnerabilities in cancer demonstrated via structure-based protac design. *Nature chemical biology*, 15(7):672–680, 2019.

[Garcia Satorras *et al.*, 2021] Victor Garcia Satorras, Emiel Hoogeboom, Fabian Fuchs, Ingmar Posner, and Max Welling. E (n) equivariant normalizing flows. *Advances in Neural Information Processing Systems*, 34:4181–4192, 2021.

[Guan *et al.*, 2023] Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. *arXiv preprint arXiv:2303.03543*, 2023.

[Halgren, 1996] Thomas A Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of computational chemistry*, 17(5-6):490–519, 1996.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[Hoogeboom *et al.*, 2022] Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pages 8867–8887. PMLR, 2022.

[Hu *et al.*, 2005] Liegi Hu, Mark L Benson, Richard D Smith, Michael G Lerner, and Heather A Carlson. Binding moad (mother of all databases). *Proteins: Structure, Function, and Bioinformatics*, 60(3):333–340, 2005.

[Huang *et al.*, 2022] Yinan Huang, Xingang Peng, Jianzhu Ma, and Muhan Zhang. 3dlinker: an e (3) equivariant variational autoencoder for molecular linker design. *arXiv preprint arXiv:2205.07309*, 2022.

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

[Imrie *et al.*, 2020] Fergus Imrie, Anthony R Bradley, Mihaela van der Schaar, and Charlotte M Deane. Deep generative models for 3d linker design. *Journal of chemical information and modeling*, 60(4):1983–1995, 2020.

[Imrie *et al.*, 2021] Fergus Imrie, Thomas E Hadfield, Anthony R Bradley, and Charlotte M Deane. Deep generative design with 3d pharmacophoric constraints. *Chemical science*, 12(43):14577–14589, 2021.

[Ingraham *et al.*, 2023] John B Ingraham, Max Baranov, Zak Costello, Karl W Barber, Wujie Wang, Ahmed Ismail, Vincent Frappier, Dana M Lord, Christopher Ng-Thow-Hing, Erik R Van Vlack, et al. Illuminating protein

space with a programmable generative model. *Nature*, 623(7989):1070–1078, 2023.

[Jing *et al.*, 2022] Bowen Jing, Gabriele Corso, Jeffrey Chang, Regina Barzilay, and Tommi Jaakkola. Torsional diffusion for molecular conformer generation. *Advances in neural information processing systems*, 35:24240–24253, 2022.

[Lin *et al.*, 2025] Haitao Lin, Yufei Huang, Odin Zhang, Siqi Ma, Meng Liu, Xuanjing Li, Lirong Wu, Jishui Wang, Tingjun Hou, and Stan Z Li. Diffbp: Generative diffusion of 3d molecules for target protein binding. *Chemical Science*, 16(3):1417–1431, 2025.

[Lugmayr *et al.*, 2022] Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11461–11471, 2022.

[Luo *et al.*, 2022] Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures. *Advances in Neural Information Processing Systems*, 35:9754–9767, 2022.

[Peng *et al.*, 2023] Xingang Peng, Jiaqi Guan, Qiang Liu, and Jianzhu Ma. Moldiff: Addressing the atom-bond inconsistency problem in 3d molecule diffusion generation. *arXiv preprint arXiv:2305.07508*, 2023.

[Putta *et al.*, 2005] Santosh Putta, Gregory A Landrum, and Julie E Penzotti. Conformation mining: an algorithm for finding biologically relevant conformations. *Journal of medicinal chemistry*, 48(9):3313–3318, 2005.

[Qiao *et al.*, 2022] Zhuoran Qiao, Weili Nie, Arash Vahdat, Thomas F Miller III, and Anima Anandkumar. Dynamic-backbone protein-ligand structure prediction with multiscale generative diffusion models. *arXiv preprint arXiv:2209.15171*, 1, 2022.

[Rath and Condurache, 2022] Matthias Rath and Alexandru Paul Condurache. Improving the sample-complexity of deep classification networks with invariant integration. *arXiv preprint arXiv:2202.03967*, 2022.

[Salentin *et al.*, 2015] Sebastian Salentin, Sven Schreiber, V Joachim Haupt, Melissa F Adasme, and Michael Schroeder. Plip: fully automated protein–ligand interaction profiler. *Nucleic acids research*, 43(W1):W443–W447, 2015.

[Schneuing *et al.*, 2024] Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom L Blundell, Pietro Lio, et al. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4(12):899–909, 2024.

[Sheng and Zhang, 2013] Chunquan Sheng and Wannian Zhang. Fragment informatics and computational fragment-based drug design: an overview and update. *Medicinal Research Reviews*, 33(3):554–598, 2013.

[Shi *et al.*, 2021] Chence Shi, Shitong Luo, Minkai Xu, and Jian Tang. Learning gradient fields for molecular conformation generation. In *International conference on machine learning*, pages 9558–9568. PMLR, 2021.

[Sohl-Dickstein *et al.*, 2015] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. pmlr, 2015.

[Su *et al.*, 2018] Minyi Su, Qifan Yang, Yu Du, Guoqin Feng, Zhihai Liu, Yan Li, and Renxiao Wang. Comparative assessment of scoring functions: the casf-2016 update. *Journal of chemical information and modeling*, 59(2):895–913, 2018.

[Trott and Olson, 2010] Oleg Trott and Arthur J Olson. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

[Vignac *et al.*, 2022] Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. *arXiv preprint arXiv:2209.14734*, 2022.

[Watson *et al.*, 2022] Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models. *BioRxiv*, pages 2022–12, 2022.

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

[Xu *et al.*, 2021] Minkai Xu, Shitong Luo, Yoshua Bengio, Jian Peng, and Jian Tang. Learning neural generative dynamics for molecular conformation generation. *arXiv preprint arXiv:2102.10240*, 2021.

[Xu *et al.*, 2022] Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. *arXiv preprint arXiv:2203.02923*, 2022.

[Yang *et al.*, 2020] Yuyao Yang, Shuangjia Zheng, Shimin Su, Chao Zhao, Jun Xu, and Hongming Chen. Syntalinker: automatic fragment linking with deep conditional transformer neural networks. *Chemical science*, 11(31):8312–8322, 2020.

[Zhou *et al.*, 2024] Xiangxin Zhou, Jiaqi Guan, Yijia Zhang, Xingang Peng, Liang Wang, and Jianzhu Ma. Reprogramming pretrained target-specific diffusion models for dual-target drug design. *Advances in Neural Information Processing Systems*, 37:87255–87281, 2024.

[Zhung *et al.*, 2024] Wonho Zhung, Hyeongwoo Kim, and Woo Youn Kim. 3d molecular generative framework for interaction-guided drug design. *Nature Communications*, 15(1):2688, 2024.