MolFORM: Multi-modal Flow Matching for Structure-Based Drug Design

Jie Huang [*1] **Daiheng Zhang** [*2]

Abstract

Structure-based drug design (SBDD) seeks to generate molecules that bind effectively to protein targets by leveraging their 3D structural information. While diffusion-based generative models have become the predominant approach for SBDD, alternative non-autoregressive frameworks remain relatively underexplored. In this work, we introduce MolFORM, a novel generative framework that jointly models discrete (atom types) and continuous (3D coordinates) molecular modalities using multi-flow matching. To further enhance generation quality, we incorporate a preference-guided fine-tuning stage based on *Direct Preference Optimization* (DPO), using Vina score as a reward signal. We propose a multi-modal flow DPO co-modeling strategy that simultaneously aligns discrete and continuous modalities, leading to consistent improvements across multiple evaluation metrics.

1. Introduction

Structure-based drug design (SBDD) (Anderson, 2003) accelerates drug discovery by utilizing the three-dimensional structures of biological targets, enabling the efficient and rational design of molecules within a defined chemical space. Generative models have recently emerged as a powerful approach for streamlining the SBDD process by directly proposing candidate molecules, thus bypassing the need for exhaustive exploration of large chemical libraries. Advances in this area can be broadly categorized into two directions: autoregressive models (Luo et al., 2021), which formulate molecule generation as a sequential prediction task, and diffusion models (Guan et al., 2023; 2024), which draw inspiration from the iterative refinement process commonly used in image generation.

Despite the variety of non-autoregressive generative models, diffusion-based approaches have become the dominant paradigm. In SBDD, many extensions have been developed on top of diffusion models to better handle protein-ligand interactions, with a particular focus on improving binding affinity through task-specific objectives and interaction-aware designs (Huang et al., 2023; Guan et al., 2024). In parallel, there has been growing interest in exploring alternative non-autoregressive frameworks such as Bayesian Flow Networks (Qu et al., 2024), which have also demonstrated promising results, achieving state-of-the-art performance (Lin et al., 2024) on several benchmark SBDD tasks.

In recent years, flow matching (Liu et al., 2022b; Lipman et al., 2022) has emerged as a widely studied generative modeling framework. Although flow matching is theoretically equivalent to diffusion models under certain conditions (Gao et al., 2024), empirical performance can vary significantly depending on the choice of scheduling strategy. In image generation tasks, flow matching has been successfully scaled to large datasets and has demonstrated strong performance (Esser et al., 2024; Liu et al., 2023). In the domain of AI for science, especially for molecular and protein generation tasks, researchers have also begun to explore the applicability of flow matching (Jing et al., 2024; Geffner et al., 2025; Campbell et al., 2024). Conceptually, flow matching offers a transport mapping interpretation from the perspective of ordinary differential equations (ODEs), providing a flexible modeling framework that enables task-specific adaptations. For structure-based drug design (SBDD), the generative task involves predicting both atom types and their 3D positions, which can be viewed as a combination of discrete and continuous modalities. Motivated by recent advances, we propose **MolFORM**, a novel framework for **Mol**ecular multi-modal **F**low-**O**ptimized **R**epresentation **M**atching. Our results show that carefully designed flow-based models can achieve performance comparable to SOTA diffusion-based approaches.

Furthermore, we incorporate a preference-guided fine-tuning stage using *Direct Preference Optimization* (DPO), leveraging the Vina score as a chemically informed reward signal to refine the base model. We first propose a discrete flow matching variant of DPO under a uniform noising

[*]Equal contribution [1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio, USA [2]Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, New Jersey, USA. Correspondence to: Daiheng Zhang <dz367@rutgers.edu>.

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).

scheme, which enables preference alignment in the atom type space. More importantly, we demonstrate that the full benefit arises from a multi-flow DPO co-modeling strategy that jointly aligns preferences over both discrete atom types and continuous 3D positions. This additional step leads to notable improvements in the quality of generated molecules, with our experiments showing that the gains from DPO fine-tuning surpass those of standard diffusion-based models. To the best of our knowledge, this is the first work to apply preference alignment in a multi-flow setting, highlighting its potential for broader adoption in similar generative tasks.

2. Related work

2.1. Structure-Based Drug Design.

With the increasing availability of structural data, generative models have attracted significant attention for structure-based molecule generation. Early methods (Skalic et al., 2019) utilized sequence generative models to produce SMILES representations from protein contexts. Driven by advancements in 3D geometric modeling, subsequent studies directly generate molecules in 3D space. For instance, Ragoza et al. (2022) employ voxelized atomic density grids within a Variational Autoencoder framework. Other methods use autoregressive models to sequentially place atoms or chemical groups (Luo et al., 2021; Peng et al., 2022), while FLAG (Zhang et al., 2023) and DrugGPS (Zhang & Liu, 2023) leverage chemical priors to construct realistic ligand fragments incrementally. More recently, diffusion models have demonstrated notable success by progressively denoising atom types and coordinates, maintaining SE(3)-equivariant symmetries (Guan et al., 2023; Lin et al., 2022; Schneuing et al., 2023; Huang et al., 2023; Guan et al., 2024; Zhang et al., 2024). Despite these advances, existing models often struggle with generating molecules simultaneously optimized for multiple desirable properties, such as binding affinity, synthesizability, and low toxicity—critical considerations in drug discovery (D Segall, 2012).

2.2. Flow Matching.

Flow matching (Lipman et al., 2022; Liu et al., 2022b) is a continuous-time generative modeling framework that generalizes diffusion models by learning a vector field to transport a simple prior $q(\boldsymbol{x})$ toward the data distribution $p_{\text{data}}(\boldsymbol{x})$. It defines a conditional path $p_t(\boldsymbol{x} \mid \boldsymbol{x}_1)$ that interpolates between $q(\boldsymbol{x})$ and the target $\delta(\boldsymbol{x}-\boldsymbol{x}_1)$, and learns the marginal vector field $v(\boldsymbol{x},t) = \mathbb{E}_{\boldsymbol{x}_1\sim p_t(\boldsymbol{x}_1|\boldsymbol{x})}[u_t(\boldsymbol{x} \mid \boldsymbol{x}_1)]$ using a neural network $v_\theta(\boldsymbol{x},t)$. Diffusion models can be seen as a special case under Gaussian interpolation (Gao et al., 2024). Flow matching has demonstrated strong performance in image and video generation (Liu et al., 2023; Esser et al., 2024), and is gaining traction in scientific domains such as protein generation (Campbell et al., 2024), protein conformation generation (Jing et al., 2024), where it naturally handles both discrete atom types and continuous 3D coordinates.

2.3. Preference alignment of generative models.

While maximizing data likelihood is standard in generative modeling, it often fails to align with downstream user preferences. Reinforcement learning from human feedback (RLHF) (Ziegler et al., 2020; Ouyang et al., 2022) and its variants have been widely adopted to align large language models with human intent. Recent efforts extend these ideas to diffusion models, treating generation as a multi-step decision process (Uehara et al., 2024; Wallace et al., 2023). Direct Preference Optimization (DPO) (Rafailov et al., 2023) offers a simpler alternative by bypassing reinforcement learning and directly optimizing models against pairwise preference data. DPO has shown competitive results in both language and image domains (Wallace et al., 2023; Zhou et al., 2024). Meanwhile, we also observe that some studies (Gu et al., 2024; Cheng et al., 2024) have begun to incorporate preference alignment into structure-based drug design (SBDD), highlighting its potential to improve biological plausibility and design success. However, applying DPO to jointly model discrete and continuous variables remains largely unexplored.

3. Methods

3.1. Problem definitions.

We aim to generate ligand molecules that are capable of binding to specific protein binding sites, by modeling $p(M|P)$. We represent the protein pocket as a collection of N_P atoms, $P = \{(x_P^i, v_P^i)\}_{i=1}^{N_P}$. Similarly, the ligand molecule can be represented as a collection of N_M atoms, $M = \{(x_M^i, v_M^i)\}_{i=1}^{N_M}$, where $x_M^i \in \mathbb{R}^3$ represents atom position and $v_M^i \in [k]$ the k possible atom types. The number of atoms N_M can be sampled from an empirical distribution (Hoogeboom et al., 2022; Guan et al., 2023). For brevity, the ligand molecule is denoted as $M = \{\mathbf{X}, \mathbf{V}\}$ where $\mathbf{X} \in \mathbb{R}^{N_M\times 3}$ and $\mathbf{V} \in [k]^{N_M\times K}$.

3.2. Multi-modal Flow matching

The Conditional Flow Matching (CFM) framework (Lipman et al., 2022; Liu et al., 2022b) learns a time-dependent flow $\psi_t : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ that transports samples from a source distribution p_0 to a target distribution p_1, governed by the ODE $\frac{d}{dt}\mathbf{x}_t = u_t(\mathbf{x}_t)$ with $\mathbf{x}_t = \psi_t(\mathbf{x}_0)$. Since the exact vector field u_t is generally intractable, the conditional formulation enables a tractable objective by conditioning on both source and target samples, where the v_θ is a time-depended, learnable model with parameters θ:

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{t,\mathbf{x}_0,\mathbf{x}_1} \left\| v_\theta(\mathbf{x}_t, t) - u_t(\mathbf{x}_t|\mathbf{x}_0, \mathbf{x}_1) \right\|^2 . \quad (1)$$

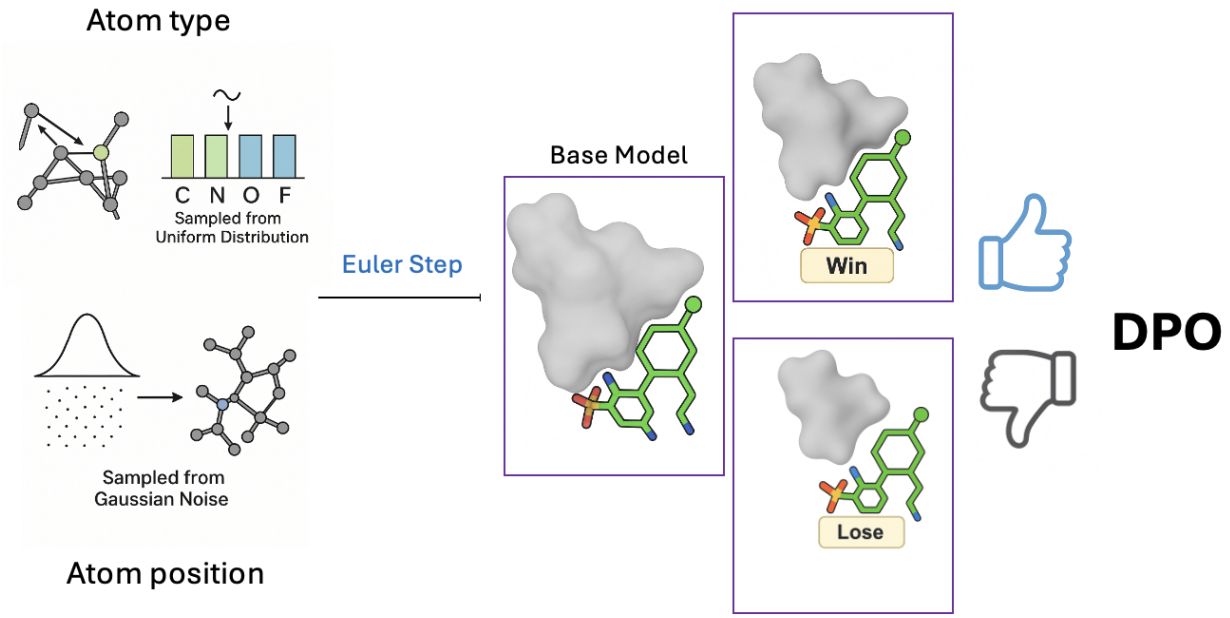

Figure 1. Overview of MolFORM. This workflow can be summarized as two steps: 1) Employs multi-flow generation to construct the base model. 2) Applies DPO to fine-tune the dual modalities, using the Vina score as the reward.

In our structure-based drug design (SBDD) task, we model each atom type as a categorical variable, where $x \in \mathcal{K}$ and $\mathcal{K} = \{1, \ldots, K\}$ is the set of possible atom types with cardinality K. To model the generative process over atom types, we adopt the Discrete Flow Matching (DFM) framework (Campbell et al., 2024), which constructs a time-dependent probability flow $p_t(v)$ that evolves from a known noise distribution $p_0(v)$ to the empirical data distribution $p_1(v) = p_{\text{data}}(v)$ over a time interval $t \in [0, 1]$.

This flow is defined via a family of conditional distributions $\pi_t(v_t \mid v_1)$, referred to as conditional flows, where $v_1 \in [k]$ is the true atom type and v_t is its corrupted version at time t. We adopt the uniform corruption process:

$$\pi_t(v_t \mid v_1) = (1 - t) \cdot \text{Uniform}([k]) + t \cdot \delta_{v_1}(v_t), \quad (2)$$

where $\delta_{v_1}(v_t)$ is the Kronecker delta function (equal to 1 if $v_t = v_1$, and 0 otherwise), and $\text{Uniform}([k]) = 1/k$ for all $v_t \in [k]$. This defines a simple interpolation between a uniform noise distribution $p_0(v) = \text{Uniform}([k])$ at $t = 0$ and the data distribution $p_1(v) = p_{\text{data}}(v)$ at $t = 1$. The marginal distribution at time t is given by averaging over the data distribution:

$$p_t(v_t) = \mathbb{E}_{v_1 \sim p_{\text{data}}} \left[\pi_t(v_t \mid v_1) \right]. \quad (3)$$

To train the model, we learn a denoising distribution $p^\theta(v_1^i \mid \mathbf{v}_t)$ that predicts the original atom type $v_1^i \in [k]$ from a corrupted ligand type sequence $\mathbf{v}_t = \{v_t^i\}_{i=1}^{N_M}$. The model is optimized using a standard cross-entropy loss:

$$\mathcal{L}_{\text{CE}} = \mathbb{E}_{t, \mathbf{v}_1, \mathbf{v}_t \sim \pi_t(\cdot \mid \mathbf{v}_1)} \left[-\sum_{i=1}^{N_M} \log p^\theta(v_1^i \mid \mathbf{v}_t) \right], \quad (4)$$

where $\pi_t(v_t^i \mid v_1^i)$ denotes the uniform corruption process.

To improve training stability, especially when $t \to 1$, we adopt a reparameterized formulation inspired by diffusion-based objectives. Instead of regressing the velocity field directly, we predict the target sample $\hat{\mathbf{x}}_1 = v_\theta(\mathbf{x}_t, t)$ and define the loss in terms of reconstructed flows:

$$\mathcal{L}_{\text{reparam}}(\theta) = \mathbb{E}_{t, \mathbf{x}_0, \mathbf{x}_1} \left\| u_t(\mathbf{x}_t | \hat{\mathbf{x}}_1, \mathbf{x}_0) - u_t(\mathbf{x}_t | \mathbf{x}_1, \mathbf{x}_0) \right\|^2. \quad (5)$$

On Euclidean manifolds, where $u_t(\mathbf{x}_t | \mathbf{x}_1, \mathbf{x}_0) = \mathbf{x}_1 - \mathbf{x}_0$, this reduces to a simple mean squared error:

$$\mathcal{L}_{\text{pos}} = \mathbb{E}_{t, \mathbf{x}_0, \mathbf{x}_1} \left\| \hat{\mathbf{x}}_1 - \mathbf{x}_1 \right\|_2^2. \quad (6)$$

Similarly, we reparameterized the cross-entropy loss to predict the clean sample $\hat{\mathbf{v}}_1$ and match the conditional flows induced by $\hat{\mathbf{v}}_1$ and $\mathbf{v}_1$. Here the loss function is expressed as:

$$\mathcal{L}_{\text{type}} = \mathbb{E}_{t, \mathbf{v}_0, \mathbf{v}_1} \left[\text{CE}\left(\hat{\mathbf{v}}_1, \mathbf{v}_1\right) \right], \quad (7)$$

This reparameterization has been shown to enhance numerical stability, especially when sampling with large t.

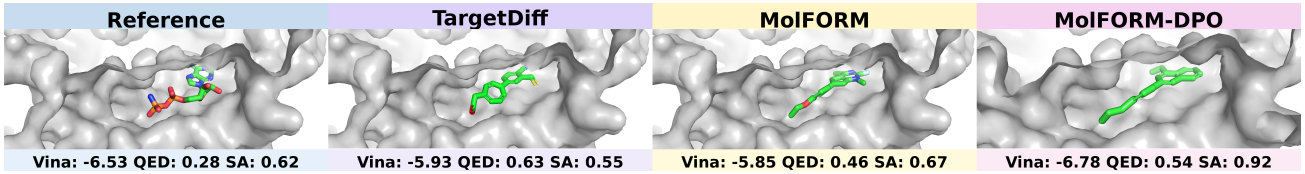

Figure 2. Visualizations of reference molecules and generated ligands for protein pockets (4yhj) generated by Reference, Targetdiff, MolFORM and MolFORM-DPO. Vina score, QED, and SA are reported below.

Chamfer loss To promote accurate geometric alignment between predicted and ground-truth molecular structures, we incorporate a Chamfer loss defined over atomic point clouds. Given two point sets $\hat{\mathbf{x}}_1 = \{\hat{x}_i\}_{i=1}^N$ and $\mathbf{x}_1 = \{x_j\}_{j=1}^M$ representing predicted and reference atomic positions respectively, the Chamfer distance is defined as:

$$\mathcal{L}_{\text{Chamfer}} = \frac{1}{N} \sum_{\hat{x} \in \hat{\mathbf{x}}_1} \min_{x \in \mathbf{x}_1} \|\hat{x} - x\|_2 + \frac{1}{M} \sum_{x \in \mathbf{x}_1} \min_{\hat{x} \in \hat{\mathbf{x}}_1} \|x - \hat{x}\|_2. \tag{8}$$

This bidirectional distance encourages each predicted atom to be close to some reference atom, and vice versa, effectively guiding the flow matching process to learn a distribution that preserves spatial fidelity to the ground-truth.

The final loss is shown below:

$$\mathcal{L} = \mathcal{L}_{\text{pos}} + \mathcal{L}_{\text{type}} + \lambda \cdot \mathcal{L}_{\text{Chamfer}} \tag{9}$$

where λ is a weighting hyperparameter that balances the contribution of the Chamfer loss.

3.3. Sampling

We consider two types of generative sampling procedures, corresponding to the discrete atom types and continuous atomic coordinates.

Continuous sampling (atom positions) For atomic coordinates, we simulate trajectories via the learned velocity field $v_\theta(\mathbf{x}_t, t)$ using Euler integration in N steps:

$$\mathbf{x}_{t+\frac{1}{N}} = \mathbf{x}_t + \frac{1}{N} \cdot v_\theta(\mathbf{x}_t, t), \quad t \in \left\{0, \tfrac{1}{N}, \ldots, \tfrac{N-1}{N}\right\}. \tag{10}$$

Starting from $\mathbf{x}_0 \sim p_0$, where p_0 is a Gaussian distribution, we integrate the dynamics to obtain the final conformation $\mathbf{x}_1$.

Discrete sampling (atom types) Given an initial sample $v_0 \sim p_0$, where p_0 is a discrete uniform distribution over the number of atom types, we simulate the forward trajectory using Euler integration:

$$v_{t+\Delta t} \sim \text{Cat}\left(\delta_{v_t} + R_t(v_t, \cdot)\, \Delta t\right). \tag{11}$$

The rate matrix can be computed as an expectation over conditional flows:

$$R_t(v_t, j) = \mathbb{E}_{v_1 \sim q(v_1 | v_t)} \left[R_t(v_t, j \mid v_1)\right], \tag{12}$$

where the posterior is given by

$$q(v_1 \mid v_t) = \frac{\pi_t(v_t \mid v_1) \cdot p_{\text{data}}(v_1)}{p_t(v_t)}. \tag{13}$$

3.4. Direct Preference Optimization(DPO) on Multi-Flow

Building upon these developments, we introduce a Multi-flow DPO approach that jointly optimizes both continuous and discrete data using Reinforcement Learning from Human Feedback (RLHF).

Given winning samples x_1^w and losing samples x_1^l for a given sample x_1 under a preference $\mathcal{P}$, we construct a dataset of preference pairs $\mathcal{D} = \{(p, x_1^w, x_1^l)\}$, where p denotes the protein condition. Using the loss function in Equation (14), we bypass the need to explicitly estimate a reward function and instead directly optimize the parameterized model p_θ. Here, $\log \sigma$ denotes the log-sigmoid function, β is a scaling (regularization) coefficient, p_θ is the learnable model, and p_{ref} is a fixed reference model.

$$L_{\text{DPO}}(\theta) = -\mathbb{E}_{(x_1^w, x_1^l) \sim \mathcal{D}, p_{1|t}^\theta(x_1|x_t), t \sim \mathcal{U}[0,1]} \Big[\log \sigma \Big(\beta \log \frac{p_\theta(x_1^w \mid p)}{p_{\text{ref}}(x_1^w \mid p)} - \beta \log \frac{p_\theta(x_1^l \mid p)}{p_{\text{ref}}(x_1^l \mid p)}\Big)\Big] \tag{14}$$

The DPO loss can be decomposed into three parts 15, 16 and 17 as below, aligning with Equation (9), to reflect discrete and continuous components:

$$L_{\text{DPO}}^x(\theta) = -\mathbb{E}_{(x_1^w, x_1^l) \sim \mathcal{D},\, t \sim \mathcal{U}[0,1], p_{1|t}^\theta(x_1|x_t)} \Big[\log \sigma \Big(-\beta \big(\|x_1^w - \hat{x}_{1,\theta}^w\|^2 - \|x_1^w - \hat{x}_{1,\text{ref}}^w\|^2 - \|x_1^l - \hat{x}_{1,\theta}^l\|^2 + \|x_1^l - \hat{x}_{1,\text{ref}}^l\|^2\big)\Big)\Big] \tag{15}$$

Model	Vina Score (↓)	Vina Min (↓)	Vina Dock (↓)	Diversity (↑)	QED (↑)		SA (↑)		Static Geometry (↓)		Clash (↓)	
	Avg.	Avg.	Avg.	Avg.	Avg.	Med.	Avg.	Med.	JSD_{BL}	JSD_{BA}	$Ratio_{cca}$	$Ratio_{cm}$
LiGAN	**-6.47**	-7.14	-7.70	0.66	0.46	0.46	**0.66**	**0.66**	0.4645	0.5673	**0.0096**	**0.0718**
3DSBDD	-	-3.75	-6.45	0.70	0.48	0.48	0.63	0.63	0.5024	0.3904	0.2482	0.8683
GraphBP	-	-	-4.57	**0.79**	0.44	0.44	0.64	0.64	0.5182	0.5645	0.8634	0.9974
Pocket2Mol	-5.23	-6.03	-7.05	0.69	0.39	0.39	0.65	0.65	0.5433	0.4922	0.0576	0.4499
TargetDiff	-5.71	-6.43	-7.41	0.72	0.49	0.49	0.60	0.60	0.2659	0.3769	0.0483	0.4920
DiffSBDD	-	-2.15	-5.53	-	0.49	0.49	0.34	0.34	0.3501	0.4588	0.1083	0.6578
DiffBP	-	-	-7.34	-	0.47	0.47	0.59	0.59	0.3453	0.4621	0.0449	0.4077
FLAG	-	-	-3.65	-	0.41	0.41	0.58	0.58	0.4215	0.4304	0.6777	0.9769
D3FG	-	-2.59	-6.78	-	0.49	0.49	**0.66**	**0.66**	0.3727	0.4700	0.2115	0.8571
DecompDiff	-5.18	-6.04	-7.10	0.68	0.49	0.49	**0.66**	**0.66**	0.2576	0.3473	0.0462	0.5248
MolCraft	-6.15	-6.99	-7.79	0.72	0.48	0.48	**0.66**	**0.66**	**0.2250**	**0.2683**	0.0264	0.2691
VoxBind	-6.16	-6.82	-7.68	-	**0.54**	**0.54**	0.65	0.65	0.2701	0.3771	0.0103	0.1890
MolFORM	-5.42	-6.42	-7.50	0.78	0.48	0.49	0.60	0.58	0.3225	0.5535	0.0310	0.4474
MolFORM-DPO	-6.16	**-7.18**	**-8.13**	0.77	0.50	0.51	0.65	0.63	0.3215	0.5584	0.0188	0.2525
Reference	-6.36	-6.71	-7.45	0.48	0.47	0.73	0.74	-	-	-	-	

Table 1. Combined results for binding affinity, chemical properties, and geometry/clash metrics. (↑)/(↓) denote better. Top 2 results are marked in bold and underlined. The result from baseline model is quoted from Lin et al. (2024).

$$L_{\text{DPO}}^{\text{point cloud}}(\theta) = -\mathbb{E}_{(x_1^w, x_1^l)\sim\mathcal{D},\, t\sim\mathcal{U}[0,1], p_{1|t}^{\theta}(x_1|x_t)} \Big[\log\sigma\Big(-\beta\big(\mathcal{L}_{\text{Chamfer}}(x_1^w, \hat{x}_{1,\theta}^w) - \mathcal{L}_{\text{Chamfer}}(x_1^w, \hat{x}_{1,\text{ref}}^w) - \mathcal{L}_{\text{Chamfer}}(x_1^l, \hat{x}_{1,\theta}^l) + \mathcal{L}_{\text{Chamfer}}(x_1^l, \hat{x}_{1,\text{ref}}^l)\big)\Big)\Big] \quad (16)$$

$$L_{\text{DPO}}^{\text{v}}(\theta) = -\mathbb{E}_{(x_1^w, x_1^l)\sim\mathcal{D},\, t\sim\mathcal{U}[0,1]} \Big[\log\sigma\Big(-\beta\big(\mathcal{D}_{\text{ref}}^{\theta}(x_t^w|x_1^w) - \mathcal{D}_{\text{ref}}^{\theta}(x_t^l|x_1^l)\big)\Big)\Big] \quad (17)$$

with

$$\mathcal{D}_{\text{ref}}^{\theta}(x_t|x_1) = \sum_{j\neq x_t} R_t^q(x_t, j|x_1) \log \frac{R_t^{\theta}(x_t, j)}{R_t^{\text{ref}}(x_t, j)} + R_t^{\text{ref}}(x_t, j) - R_t^{\theta}(x_t, j) \quad (18)$$

The values $\hat{x}_{1,\theta}^l$ and $\hat{x}_{1,\theta}^w$ are the predictions from the model p_θ, while $\hat{x}_{1,\text{ref}}^l$ and $\hat{x}_{1,\text{ref}}^w$ are the corresponding predictions from the reference model p_{ref}.

In the discrete space, the rate matrix $R_t(x_t, x_t + \Delta t)$ is defined as the expectation of the conditional rate matrix as shown in the equation (19).

$$R_t(x_t, x_{t+\Delta t}) = \mathbb{E}_{p_{1|t}^{\theta}(x_1|x_t)} \left[R_t^q(x_t, x_{t+\Delta t}|x_1)\right] \quad (19)$$

By choosing the uniform noise initialization, $\text{Uniform}([k])$, the unconditional rate matrix becomes equation (20)(Campbell et al., 2024).

$$R_t(x_t, x_{t+\Delta t}) = \frac{1}{1-t}\, p_{1|t}^{\theta}(x_1 = x_{t+\Delta t} \mid x_t), \quad \text{for } x_{t+\Delta t} \neq x_t \quad (20)$$

By substituting equation (20) into equation (18) and equation (17), we can get the final loss of DPO of the discrete space in equation (21).

$$L_{\text{DPO}}^{v}(\theta) = -\mathbb{E}_{(x_1^w, x_1^l)\sim\mathcal{D},\, t\sim\mathcal{U}[0,1], p_{1|t}^{\theta}(x_1|x_t)} \left[\log\sigma\left(-\frac{\beta}{1-t} \cdot \left(\log\frac{p_{1|t}^{\theta}(x_1^w \mid x_t^w)}{p_{1|t}^{\text{Ref}}(x_1^w \mid x_t^w)} - \log\frac{p_{1|t}^{\theta}(x_1^l \mid x_t^l)}{p_{1|t}^{\text{Ref}}(x_1^l \mid x_t^l)}\right)\right)\right] \quad (21)$$

Equation (21) is the DPO formulation for discrete flow matching under uniform noising.

3.5. Time Scheduling

Training scheduling We adopt a t-sampling strategy that emphasizes large time steps to improve the accuracy of local structural details. Following the scheduling method from Geffner et al. (2025), we sample time as

$$p(t) = 0.02\,\mathcal{U}(0,1) + 0.98\,\mathcal{B}(1.9, 1.0), \quad (22)$$

where $\mathcal{U}$ and $\mathcal{B}$ denote the uniform and Beta distributions, respectively. This schedule reduces weight on early time steps (where global structure is more variable) and concentrates training on later steps to improve geometric fidelity in atom-level predictions.

Sampling scheduling In conventional flow matching, the sampling trajectory typically uses a uniform discretization of time $t \in [0, 1]$ with N equally spaced steps. However, we observe that the model requires finer granularity near $t = 1$, where more detailed structure is synthesized. To better capture this behavior and improve generation quality, we adopt a non-uniform time discretization strategy when integrating the ODE using the Euler method. Specifically,

Method	C-C	C-N	C-O	C=C	C=N	C=O
LiGan	0.4986	0.4146	0.4560	0.4807	0.4776	0.4595
3DSBDD	0.2090	0.4258	0.5478	0.5170	0.6701	0.6448
GraphBP	0.5038	0.4231	0.4973	0.6235	0.4629	0.5986
Pocket2Mol	0.5667	0.5698	0.5433	0.4787	0.5989	0.5025
TargetDiff	0.3101	0.2490	0.3072	0.1715	0.1944	0.3629
DiffSBDD	0.3841	0.3708	0.3291	0.3043	0.3473	0.3647
DiffBP	0.5704	0.5256	0.5090	0.6161	0.6314	0.5296
FLAG	0.3460	0.3770	0.4433	0.4872	0.4464	0.4292
D3GF	0.4244	0.3227	0.3895	0.3860	0.3570	0.3566
DecompDiff	0.2562	0.2007	0.2361	0.2590	0.2844	0.3091
MolCraft	0.2473	0.1732	0.2341	0.3040	0.1459	0.2250
VoxBind	0.3335	0.2577	0.3507	0.1991	0.1459	0.3334
MolFORM	0.4282	0.3419	0.3824	0.2898	0.3239	0.3583

Table 2. JSD Bond Length Comparisons across different methods.

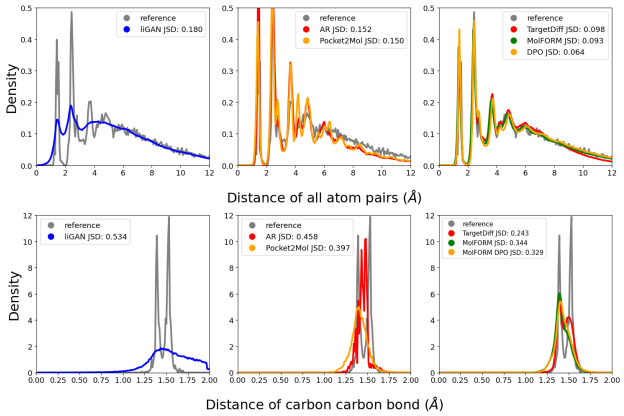

Figure 3. Comparing the distribution for distances of allatom (top row) and carbon-carbon pairs (bottom row) for reference molecules in the test set (gray) and model generated molecules (color).

we allocate 60 steps in the interval $[0, 0.8]$ and 40 steps in $[0.8, 1]$, allowing the model to focus more on the final refinement phase.

4. Experiments

4.1. Experiment Setup

Datasets Our experiments were conducted using the CrossDocked2020 dataset (Francoeur et al., 2020). Consistent with prior studies (e.g., Peng et al., 2022; Guan et al., 2023; 2024), we adhered to the same dataset filtering and partitioning methodologies. We further refined the 22.5 million docked protein binding complexes, characterized by an RMSD $< 1\AA$, and sequence identity less than 30%. This resulted in a dataset comprising 100,000 protein-binding complexes for training, alongside a set of 100 novel complexes designated for testing.

DPO dataset Our data processing strategy for DPO follows the methodology proposed in Gu et al. (2024). We preprocess the dataset into a preference format $\mathcal{D} = \{(\mathbf{p}, \mathbf{m}^w, \mathbf{m}^l)\}$, where $\mathbf{p}$ denotes the protein pocket, $\mathbf{m}^w$ the preferred ligand, and $\mathbf{m}^l$ the less preferred one. For each pocket, we sample two candidate ligands and assign preference based on a user-defined reward, mainly binding energy (e.g., Vina score). Since the affinity labels are continuous, we follow the strategy in Gu et al. (2024) and choose the molecule with the worst score as the dispreferred sample $\mathbf{m}^l$. This encourages a larger reward gap between $\mathbf{m}^w$ and $\mathbf{m}^l$, which has been shown to improve learning.

Model architecture Inspired from recent progress in equivariant neural networks (Satorras et al., 2021), we model the interaction between the ligand molecule atoms and the protein atoms with a SE(3)-Equivariant GNN, the atom hidden embedding and coordinates are updated alternately in each layer, which follows Guan et al. (2023). Our model architecture is plotted in Figure 1.

Baselines. We select all baseline models from CBG-bench (Lin et al., 2024) for comparison with our method. Early structure-based drug design (SBDD) methods are built on voxel grids with deep neural networks, such as LiGAN (Ragoza et al., 2022), which generates atom voxelized density maps using variational autoencoders (VAE) and convolutional neural networks (CNNs), and 3DSBDD (Luo et al., 2022), which predicts atom types on grids with graph neural networks (GNNs) in an auto-regressive manner. The development of equivariant graph neural networks (EGNNs) enables direct generation of 3D atom positions, as seen in Pocket2Mol (Peng et al., 2022) and GraphBP (Liu et al., 2022a), which use auto-regressive strategies with normalizing flows. Diffusion-based methods such as TargetDiff (Guan et al., 2023), DiffBP (Lin et al., 2022), and DiffSBDD (Schneuing et al., 2024) generate atom types and positions using denoising diffusion probabilistic models. Recent methods incorporate domain knowledge to guide generation: FLAG (Zhang et al., 2023) and D3FG (Lin et al., 2023) use fragment motifs for coarse molecular generation; DecompDiff (Guan et al., 2024) uses scaffold and arm clustering with Gaussian process models for atom positions. More recent advances like MolCraft (Qu et al., 2024) and VoxBind (Pinheiro et al., 2024) apply new generative modeling strategies, including Bayesian flow networks and voxel-based diffusion with walk-jump sampling.

4.2. Experiment result

Evaluation We collect all generated molecules across 100 test proteins and evaluate generated ligands from three aspects: **binding affinity**, **molecular properties** and **molecular structures**. For target binding affinity and molecular properties, we present our results under the best setting as **MolFORM**. Following previous work, we utilize AutoDock

Method	3	4	5	6	7	8
Ref.	0.0130	0.0020	0.2855	0.6894	0.0098	0.0003
LiGan	0.2238	0.0698	0.2599	0.4049	0.0171	0.0096
3DSBDD	0.2970	0.0007	0.1538	0.5114	0.0181	0.0116
GraphBP	0.0000	0.2429	0.1922	0.1765	0.1533	0.1113
Pocket2Mol	0.0000	0.1585	0.1822	0.4373	0.1410	0.0478
TargetDiff	0.0000	0.0188	0.2856	0.4918	0.1209	0.0298
DiffSBDD	0.2842	0.0330	0.2818	0.2854	0.0718	0.0193
DiffBP	0.0000	0.2195	0.2371	0.2215	0.1417	0.0707
FLAG	0.0000	0.0682	0.2716	0.5228	0.0996	0.0231
D3FG	0.0000	0.0201	0.2477	0.5966	0.0756	0.0283
DecompDiff	0.0302	0.0378	0.3407	0.4386	0.1137	0.0196
MolCraft	0.0000	0.0022	0.2494	0.6822	0.0489	0.0072
VoxBind	0.0000	0.0062	0.2042	0.7566	0.0232	0.0021
MolFORM	0.0000	0.0291	0.2846	0.4211	0.2033	0.0455

Table 3. Distribution of different ring sizes across various methods.

Vina (Eberhardt et al., 2021) for binding affinity estimation. For **binding affinity**, we report the **Vina Score**, which evaluates the initially generated binding pose; the **Vina Min**, obtained after local energy minimization; and the **Vina Dock**, representing the lowest energy score from a global re-docking procedure using grid-based search. For **molecular properties**, we primarily report **QED** (Quantitative Estimate of Drug-likeness) and **SA** (Synthetic Accessibility). These results are summarized in Table 1. For **molecular structures**, we first evaluate several geometry-related properties following the setup in Lin et al. (2024). These include (i) $\mathbf{JSD}_{\mathrm{BL}}$ and (ii) $\mathbf{JSD}_{\mathrm{BA}}$, which measure the divergence in bond length and bond angle distributions between generated and reference molecules, reflecting structural realism. (iii) $\mathbf{Ratio}_{\mathrm{cca}}$ denotes the proportion of atoms with steric clashes—defined as van der Waals overlaps ≥ 0.4Å—with protein atoms. (iv) $\mathbf{Ratio}_{\mathrm{cm}}$ captures the fraction of generated molecules that contain any such clashes. These results are also included in Table 1. (v) Root Mean Square Deviation (**RMSD**) is a standard metric to evaluate the structural deviation between two sets of atomic coordinates. Lower RMSD indicates higher structural consistency reflects the extent to which rigid fragments preserve their geometric integrity before and after force field optimization, which is shown in Figure 4. Besides, we report the ring portion for chemical assessment in Table 3. For substructure-level evaluation, we analyze the Jensen-Shannon divergence (JSD) between the bond length distributions of reference and generated molecules, as shown in Table 2. Here, we mainly compare with TargetDiff, which is highly similar to our proposed MolFORM.

Result analysis The visualization comparison result is shown in Figure 2. As shown in Table 1, we observed that MolFORM achieved comparable performance to our baseline model TargetDiff in binding affinity and molecular properties. Compared to Targetdiff, the model shows better performance in terms of Clash and Vina Dock. After further DPO fine-tuning, our model achieves state-of-the-art performance across nearly all metrics.

Although we use only binding affinity as the reward, DPO aligns the generation distribution toward higher-quality molecules while preserving the chemical priors from the pretrained model, thus avoiding overfitting to a single objective. As a result, metrics such as QED and SA can also improve. This is further supported by the enhanced quality of generated molecular substructures in Figure 3 and Figure 4. Additionally, we observe that the diversity of generated molecules remains at 0.77 after DPO fine-tuning, indicating that our approach successfully avoids mode collapse while improving generation quality.

We re-trained a version of Targetdiff and applied DPO-based fine-tuning on the model. We observed that applying DPO to Targetdiff led to improvements of 2%, 2%, 2% and 3% in QED, SA, Vina score, and Vina min, while MolFORM achieved improvements of 4%, 8%, 14% and 12% on the same metrics in Table 4. This suggests that our model has greater potential to benefit from DPO fine-tuning. MultiFlow explicitly separates the generation of discrete (e.g., atom types) and continuous (e.g., 3D coordinates) modalities via dedicated flow-based branches. This factorized structure enables DPO to assign fine-grained preference signals to each modality during optimization. As DPO compares generation quality between molecule pairs, the modular design of MultiFlow allows more targeted updates—for example, refining atom types without perturbing geometry, or vice versa. This structural disentanglement enhances the model's responsiveness to reward signals, reduces gradient interference, and ultimately makes preference optimization more effective and stable.

Metric	TargetDiff	TargetDiff-DPO	MolFORM	MolFORM-DPO
Vina Score (↓)	-5.47	-5.58	-5.42	-6.16
Vina Min (↓)	-6.39	-6.59	-6.42	-7.18
QED (↑)	0.46	0.47	0.48	0.50
SA (↑)	0.60	0.61	0.60	0.65

Table 4. Comparison of model performance before and after DPO fine-tuning. The reported numbers are average values here. The results of vanilla TargetDiff are obtained by retraining the model.

Chamfer loss Impact We conducted ablation studies on the Chamfer DPO loss component in Table 5. Our experiments demonstrate that incorporating the Chamfer distance into the DPO objective significantly enhances the model's ability to preserve molecular geometric fidelity.

Sampling efficiency We achieve SOTA sampling performance. While it takes on average 3428s and 6189s for TargetDiff and DecompDiff to generate 100 samples respectively, our model only uses 69s. Thanks to the efficiency of flow matching, our sampling requires only 100 steps, com-

Metric	Chamferw	Chamferwo
QED (↑)	**0.48**	0.41
SA (↑)	**0.60**	0.59
Vina Score (↓)	**-5.42**	-4.67
Vina Min (↓)	**-6.42**	-5.76
JSD_{BL} (↓)	**0.3225**	0.4546
JSD_{BA} (↓)	**0.5535**	0.6173

Table 5. Ablation study: Comparison of model performance with (Chamferw) and without (Chamferwo) Chamfer loss in Vanilla Multi-model flow matching.

pared to the 1000 steps typically used in diffusion models.

Training details The model is trained via the gradient descent method Adam Kingma and Ba (Kingma & Ba, 2014) with init learning rate $= 5 \times 10^{-5}$, betas=(0.95, 0.999), batch size=16, and clip gradient norm=8. During training, we apply data augmentation by adding small Gaussian noise (standard deviation of 0.1) to the protein atom coordinates. Additionally, we use an exponential learning rate decay with a decay factor of 0.6 and a minimum threshold set at 1×10^{-8}. For the DPO model training, a smaller learning rate was applied as rate $= 5 \times 10^{-8}$ with a decay factor of 0.6 and a minimum threshold set at 1×10^{-11} and the beta dpo parameter was set as 5.

We trained our model on two NVIDIA GH200 GPUs. The base model training converged within 12 hours and 100k steps. Moreover, the DPO model training utilized 4 hours and 20k steps.

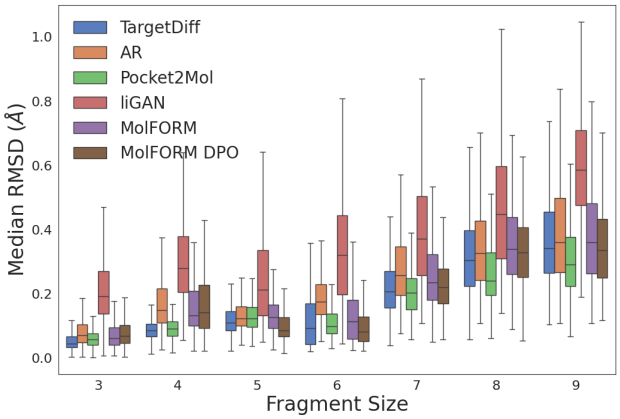

Figure 4. Median RMSD for rigid fragment before and after the force-field optimization.

5. Future Work

Our model builds upon Guan et al. (2023) and achieves comparable or even superior performance across multiple evaluation metrics, indicating that the proposed MultiFlow-based modeling is a promising direction for further exploration. Beyond DPO, other fine-tuning strategies and multi-objective optimization techniques also hold promise for further advancing the performance of our current model. Moreover, we employ uniform corruption for discrete flow matching, as masked variants unexpectedly underperformed in our molecular generation task—contrasting with their success in other scientific domains. Similarly, SDE sampling via Euler-Maruyama yielded inferior results compared to ODE integration, though the specific hyperparameters for stochastic sampling may require further tuning. These findings highlight the importance of tailoring both corruption strategies and sampling algorithms to the specific requirements of SBDD task.

6. Conclusions

In this work, we propose to use a novel Multi flow based generation framework **MolFORM** for the protein-specific molecule generation. Extensive experiments on the CrossDocked2020 benchmark demonstrate the strong performance of MolFORM. Our framework can be further combined with the Direct Preference Optimization (DPO) process, achieving state-of-the-art performance across almost all evaluation metrics. Measurements on molecular substructures indicate that the improvement is not merely due to overfitting on specific metrics, but reflects a holistic enhancement of generation quality. Moreover, our comparison with TargetDiff demonstrates that this DPO-based approach offers greater potential than unimodal fine-tuning.

Acknowledgements

This work is supported by the NSF 2340011. We would like to thank the reviewers for taking the time and effort necessary to review the manuscript. We sincerely appreciate Dr. Chengyue Gong for valuable comments.

Impact Statement

We introduce a multi-flow DPO framework that enables preference-guided generation over both discrete atom types and continuous molecular coordinates. Beyond SBDD, the generality of our framework suggests potential extensions to broader domains of computer-aided design, including small molecule generation, material discovery, and chip layout optimization. We stress the importance of ensuring the responsible deployment of this technology and caution against its misuse in harmful or unethical contexts.

References

Anderson, A. C. The process of structure-based drug design. *Chemistry & biology*, 10(9):787–797, 2003.

Campbell, A., Yim, J., Barzilay, R., Rainforth, T., and Jaakkola, T. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design. *arXiv preprint arXiv:2402.04997*, 2024.

Cheng, X., Zhou, X., Yang, Y., Bao, Y., and Gu, Q. Decomposed direct preference optimization for structure-based drug design. *arXiv preprint arXiv:2407.13981*, 2024.

D Segall, M. Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Current pharmaceutical design*, 18(9):1292–1310, 2012.

Eberhardt, J., Santos-Martins, D., Tillack, A. F., and Forli, S. Autodock vina 1.2. 0: New docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling*, 61(8):3891–3898, 2021.

Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*, 2024.

Francoeur, P. G., Masuda, T., Sunseri, J., Jia, A., Iovanisci, R. B., Snyder, I., and Koes, D. R. Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *Journal of chemical information and modeling*, 60(9):4200–4215, 2020.

Gao, R., Hoogeboom, E., Heek, J., Bortoli, V. D., Murphy, K. P., and Salimans, T. Diffusion meets flow matching: Two sides of the same coin. 2024. URL https://diffusionflow.github.io/.

Geffner, T., Didi, K., Zhang, Z., Reidenbach, D., Cao, Z., Yim, J., Geiger, M., Dallago, C., Kucukbenli, E., Vahdat, A., et al. Proteina: Scaling flow-based protein structure generative models. *arXiv preprint arXiv:2503.00710*, 2025.

Gu, S., Xu, M., Powers, A., Nie, W., Geffner, T., Kreis, K., Leskovec, J., Vahdat, A., and Ermon, S. Aligning target-aware molecule diffusion models with exact energy optimization. *Advances in Neural Information Processing Systems*, 37:44040–44063, 2024.

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. *arXiv preprint arXiv:2303.03543*, 2023.

Guan, J., Zhou, X., Yang, Y., Bao, Y., Peng, J., Ma, J., Liu, Q., Wang, L., and Gu, Q. Decompdiff: diffusion models with decomposed priors for structure-based drug design. *arXiv preprint arXiv:2403.07902*, 2024.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

Huang, Z., Yang, L., Zhou, X., Zhang, Z., Zhang, W., Zheng, X., Chen, J., Wang, Y., Bin, C., and Yang, W. Protein-ligand interaction prior for binding-aware 3d molecule diffusion models. In *The Twelfth International Conference on Learning Representations*, 2023.

Jing, B., Berger, B., and Jaakkola, T. Alphafold meets flow matching for generating protein ensembles. *arXiv preprint arXiv:2402.04845*, 2024.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Lin, H., Huang, Y., Liu, M., Li, X., Ji, S., and Li, S. Z. Diffbp: Generative diffusion of 3d molecules for target protein binding, 2022.

Lin, H., Huang, Y., Zhang, O., Liu, Y., Wu, L., Li, S., Chen, Z., and Li, S. Z. Functional-group-based diffusion for pocket-specific molecule generation and elaboration. *Advances in Neural Information Processing Systems*, 36: 34603–34626, 2023.

Lin, H., Zhao, G., Zhang, O., Huang, Y., Wu, L., Liu, Z., Li, S., Tan, C., Gao, Z., and Li, S. Z. Cbgbench: fill in the blank of protein-molecule complex binding graph. *arXiv preprint arXiv:2406.10840*, 2024.

Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

Liu, M., Luo, Y., Uchino, K., Maruhashi, K., and Ji, S. Generating 3d molecules for target protein binding. *arXiv preprint arXiv:2204.09410*, 2022a.

Liu, X., Gong, C., and Liu, Q. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022b.

Liu, X., Zhang, X., Ma, J., Peng, J., et al. Instaflow: One step is enough for high-quality diffusion-based text-to-image generation. In *The Twelfth International Conference on Learning Representations*, 2023.

Luo, S., Guan, J., Ma, J., and Peng, J. A 3d generative model for structure-based drug design. *Advances in Neural Information Processing Systems*, 34:6229–6239, 2021.

Luo, S., Guan, J., Ma, J., and Peng, J. A 3d generative model for structure-based drug design, 2022.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. Training language models to follow instructions with human feedback, 2022. *URL https://arxiv.org/abs/2203.02155*, 13, 2022.

Peng, X., Luo, S., Guan, J., Xie, Q., Peng, J., and Ma, J. Pocket2mol: Efficient molecular sampling based on 3d protein pockets. In *International Conference on Machine Learning*, pp. 17644–17655. PMLR, 2022.

Pinheiro, P. O., Jamasb, A., Mahmood, O., Sresht, V., and Saremi, S. Structure-based drug design by denoising voxel grids. *arXiv preprint arXiv:2405.03961*, 2024.

Qu, Y., Qiu, K., Song, Y., Gong, J., Han, J., Zheng, M., Zhou, H., and Ma, W.-Y. Molcraft: Structure-based drug design in continuous parameter space. *arXiv preprint arXiv:2404.12141*, 2024.

Rafailov, R., Sharma, A., Mitchell, E., Ermon, S., Manning, C. D., and Finn, C. Direct preference optimization: Your language model is secretly a reward model. *arXiv preprint arXiv:2305.18290*, 2023.

Ragoza, M., Masuda, T., and Koes, D. R. Generating 3d molecules conditional on receptor binding sites with deep generative models. *Chemical science*, 13(9):2701–2713, 2022.

Satorras, V. G., Hoogeboom, E., and Welling, M. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Schneuing, A., Du, Y., Harris, C., Jamasb, A., Igashov, I., Du, W., Blundell, T., Lió, P., Gomes, C., Welling, M., Bronstein, M., and Correia, B. Structure-based drug design with equivariant diffusion models, 2023.

Schneuing, A., Harris, C., Du, Y., Didi, K., Jamasb, A., Igashov, I., Du, W., Gomes, C., Blundell, T. L., Lio, P., et al. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4(12): 899–909, 2024.

Skalic, M., Sabbadin, D., Sattarov, B., Sciabola, S., and De Fabritiis, G. From target to drug: generative modeling for the multimodal structure-based ligand design. *Molecular pharmaceutics*, 16(10):4282–4291, 2019.

Uehara, M., Zhao, Y., Black, K., Hajiramezanali, E., Scalia, G., Diamant, N. L., Tseng, A. M., Levine, S., and Biancalani, T. Feedback efficient online fine-tuning of diffusion models. *arXiv preprint arXiv:2402.16359*, 2024.

Wallace, B., Dang, M., Rafailov, R., Zhou, L., Lou, A., Purushwalkam, S., Ermon, S., Xiong, C., Joty, S., and Naik, N. Diffusion model alignment using direct preference optimization. *arXiv preprint arXiv:2311.12908*, 2023.

Zhang, D., Gong, C., and Liu, Q. Rectified flow for structure based drug design. *arXiv preprint arXiv:2412.01174*, 2024.

Zhang, Z. and Liu, Q. Learning subpocket prototypes for generalizable structure-based drug design. In *International Conference on Machine Learning*, pp. 41382–41398. PMLR, 2023.

Zhang, Z., Min, Y., Zheng, S., and Liu, Q. Molecule generation for target protein binding with structural motifs. In *The Eleventh International Conference on Learning Representations*, 2023.

Zhou, X., Xue, D., Chen, R., Zheng, Z., Wang, L., and Gu, Q. Antigen-specific antibody design via direct energy-based preference optimization. *arXiv preprint arXiv:2403.16576*, 2024.

Ziegler, D. M., Stiennon, N., Wu, J., Brown, T. B., Radford, A., Amodei, D., Christiano, P., and Irving, G. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*, 2020.