MOLWEAVER: SCALABLE FOUNDATION MODEL FOR STRUCTURE-BASED DRUG DISCOVERY

Anonymous authors
Paper under double-blind review

ABSTRACT

Structure-based drug discovery requires designing molecules for target protein pockets and predicting ligand bound poses, both of which require learning molecular connectivity, three-dimensional structure, and ligand–protein interactions. We introduce Molweaver, a family of molecular foundation models ranging from 60M to 500M parameters, pretrained on 100 million SELFIES representations, 1 billion molecular conformers, 3.2 million protein pockets, and 167,000 redocked pocket–ligand complexes with associated properties. Molweaver is a unified transformer framework that learns molecular representations across 2D structure, 3D geometry, and chemical properties. It generates molecular sequences through autoregressive prediction and 3D conformations through flow matching. We adapt Molweaver to downstream structure-based tasks using supervised fine-tuning and reinforcement learning. For pose prediction, Molweaver places 80.2% of ligands within 2 Å RMSD of the reference pose while also passing all applicable PoseBusters checks, comparable to equivariant structure-based models despite using no equivariant inductive biases in the architecture. For molecule generation, docking-based feedback and molecular property optimization steer the model toward compounds with favorable interactions within CrossDock2020 target pockets, yielding substantially improved docking scores, full connectivity, and the lowest clash counts among compared methods. We also show that these metrics improve with model size. These results highlight the potential of large-scale joint pretraining and task-specific adaptation for developing general-purpose models for structure-based drug discovery.

1 INTRODUCTION

Structure-based drug design (SBDD) aims to discover small molecules that can selectively bind to biological targets by exploiting the three-dimensional structure of protein binding pockets. However, designing such a small molecule is a difficult task as the chemical space of possible small molecules is estimated to contain approximately 10^{60} compounds (Kirkpatrick & Ellis, 2004). Efficiently navigating this space requires generative models that can propose molecules with multiple desirable characteristics simultaneously. A successful therapeutic candidate must not only form favorable interactions with a target pocket but also satisfy practical constraints such as synthesizability, solubility, and low toxicity. In addition to molecule generation, SBDD also requires accurate prediction of the bound conformation of a ligand within a pocket, where small errors in geometry can significantly affect the predicted interactions.

Solving these problems requires models capable of capturing the complex physical relationships between molecular connectivity, three-dimensional geometry, and protein-ligand interactions. Recently, generative models have become increasingly effective for SBDD tasks (Luo et al., 2021). Existing approaches explore several modeling paradigms, including autoregressive models (Peng et al., 2022), flow matching methods (Zhou et al., 2025), diffusion models (Guan et al., 2023b), fragment-based methods (Zhang et al., 2023), and Bayesian Flow Network approaches (Qu et al., 2024). Autoregressive approaches naturally support variable-sized molecule generation, while diffusion and flow-based approaches have demonstrated strong capabilities in generating better ligands with good docking scores, but they generate molecule graphs at a size fixed before inference. These approaches typically focus on individual downstream tasks and often rely on specialized architectures designed for either molecule generation or pose prediction. Furthermore, many existing methods

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

are trained primarily on relatively small collections of experimentally determined protein-ligand complexes, such as the CrossDock2020 dataset (Francoeur et al., 2020), limiting their ability to learn broad molecular representations beyond the dataset.

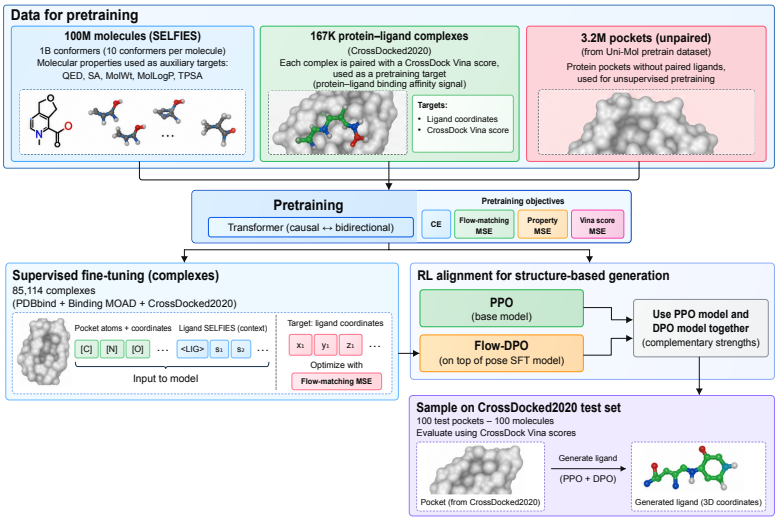


Figure 1: Molweaver training pipeline: joint molecular pretraining, supervised fine-tuning on protein–ligand complexes, and reinforcement-learning alignment for structure-based generation.

Large-scale pretrained transformers (Vaswani et al., 2017) have recently shown promise for molecular modeling. For example, Zholus et al. (2024) and Hu et al. (2025) perform pretraining on about 10 million molecules represented as SMILES strings (Weininger, 1988). However, these models generate ligand coordinates autoregressively and are not specifically designed to produce accurate docked poses, or require additional pose information during inference. This motivates the development of a unified model that can jointly learn molecular representations, generate chemically valid structures, and predict physically plausible ligand conformations.

We introduce Molweaver, a family of structure-based drug discovery foundation models that jointly model discrete molecular representations and continuous three-dimensional structures, trained on 100 million SELFIES with 10 conformers each together with CrossDock2020 docked examples. Connectivity is discrete and fits sequence prediction, whereas coordinates are continuous, so Molweaver generates molecules as SELFIES (Krenn et al., 2020) by autoregressive next-token prediction and ligand coordinates by flow matching (Lipman et al., 2022). Following Transfusion (Zhou et al., 2024), which combines causal language modeling with diffusion-based image generation in one transformer, a single backbone switches between causal and bidirectional attention depending on the objective. We also pretrain on molecular properties computed with RDKit(Landrum & others), namely Quantitative Estimate of Drug-likeness (QED), Synthetic Accessibility (SA), Molecular Weight (MolWt), Molecular LogP (MolLogP), Total Polar Surface Area (TPSA), and AutoDock Vina scores (Trott & Olson, 2010), pushing the model toward molecules that balance drug-likeness, synthetic feasibility, and favorable pocket interactions. We then adapt it to the downstream SBDD tasks by supervised fine-tuning for docked pose prediction and by Proximal Policy Optimization (PPO) (Schulman et al., 2017) and Flow-DPO (Jiao et al., 2024), which feeds AutoDock Vina scores back during molecule generation.

Molweaver uses no explicit SE(3)-equivariant components; instead, random global rotations and translations during preprocessing encourage the model to learn rotational invariance from data, as in AlphaFold3 (Abramson et al., 2024), which likewise relies on augmentation rather than architectural symmetry constraints. We evaluate the generated structures with physical plausibility metrics, including PoseCheck strain analysis (Harris et al., 2023) and the PoseBusters benchmark and its validity criteria (Buttenschoen et al., 2024), showing that Molweaver satisfies important chemical and structural constraints. The main contributions of this work are summarized as follows:

- **Molweaver foundation models for SBDD.** We introduce scalable foundation models pretrained on large-scale molecular, conformational, and pocket datasets to jointly learn molecular representations, 3D structures, and properties.
- **Unified molecule and pose generation.** We combine autoregressive SELFIES generation with flow matching for 3D coordinate generation in a single transformer, enabling molecule design and docked pose prediction. This can be later finetuned for docked pose prediction and achieves comparable performance with other methods on this task.
- **Property-guided optimization.** We incorporate molecular properties and AutoDock Vina feedback during training and use reinforcement learning to optimize generated molecules, which achieve state-of-the-art performance on SBDD metrics.

2 BACKGROUND

2.1 PROBLEM DEFINITION

Let $C_p = (A_p, X_p)$ denote a 10 angstrom protein pocket, where A_p contains atomic features and $X_p \in \mathbb{R}^{N_p \times 3}$ contains atomic coordinates. A ligand is represented by its molecular graph G_l and coordinates $X_l \in \mathbb{R}^{N_l \times 3}$.

Docked-pose prediction. Given a pocket and ligand graph, the task is to model $p_\theta(X_l \mid C_p, G_l)$. Predicted poses should minimize RMSD (Root Mean Square Deviation. See Appendix A.1) to the experimentally determined coordinates, measured in the pocket coordinate frame, while satisfying chemical and physical validity checks.

Pocket-constrained molecule generation. Given a pocket, the task is to model $p_\theta(X_l, G_l \mid C_p)$, generating both the molecular graph and its coordinates. Desired molecules should have favorable docking scores, drug-likeness, and synthetic accessibility, while forming physically plausible poses. For both tasks, validity is assessed using the applicable PoseBusters checks (Buttenschoen et al., 2024). Metric definitions are provided in Appendix A.1.

2.2 FLOW MATCHING

Flow matching learns a time-dependent vector field that transports a noise distribution p_{noise} at $t = 1$ to a data distribution $p_{\text{data}}(\cdot \mid c)$ at $t = 0$ on $\mathbb{R}^d$, where c denotes the conditioning information (Lipman et al., 2022). For ligand coordinates, $d = 3N_l$ once the ligand graph is fixed. Throughout the paper, the subscript 0 denotes clean data and $t = 1$ pure noise, matching X_0 and $U_t = \varepsilon - X_0$ of Section 4.2. Given independent samples $\epsilon \sim p_{\text{noise}} = \mathcal{N}(0, I_d)$ and $x_0 \sim p_{\text{data}}(\cdot \mid c)$, the linear path $x_t = t\epsilon + (1-t)x_0$ has target velocity $\partial x_t / \partial t = \epsilon - x_0$. With $t \sim \mathcal{U}(0,1)$, the conditional flow-matching loss is $\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{c,x_0,\epsilon,t}\left[\|v_\theta(x_t, t, c) - (\epsilon - x_0)\|_2^2\right]$. At inference, ligand coordinates are generated from Gaussian noise using DDIM sampling (Song et al., 2021).

2.2.1 FLOW-DPO

Flow-DPO optimizes preferences between poses of the same ligand under $c = (C_p, G_l)$ (Rafailov et al., 2023; Jiao et al., 2024). Preferred and rejected poses share the same timestep $t \sim \mathcal{U}(0,1)$ and noise $\varepsilon \sim \mathcal{N}(0, I_{3N_l})$. For $b \in \{+,-\}$, we compute $x_t^b = t\varepsilon + (1-t)x_0^b$, $\ell_\theta^b = \mathbb{E}_{c,x_0^b,\epsilon,t}\left\|v_\theta(x_t^b, t, c) - (\varepsilon - x_0^b)\right\|_2^2$, where "+" denotes the preferred data point and "−" denotes the rejected data point. The frozen reference error ℓ_{ref}^b is computed identically. We minimize $z = \beta\left[(\ell_\theta^- - \ell_\theta^+) - (\ell_{\text{ref}}^- - \ell_{\text{ref}}^+)\right]$, $\mathcal{L}_{\text{Flow-DPO}} = -\mathbb{E}\left[\log \text{sigmoid}(z)\right]$, with $\beta = 0.1$, and expectation over preference pairs, timesteps, and noise. This encourages lower relative error on preferred poses compared with the reference, using a velocity-error surrogate rather than an exact likelihood ratio.

3 Related Works

Pocket-aware molecule generation. Earlier models for this task were autoregressive (Peng et al., 2022; Luo et al., 2021; Liu et al., 2022), predicting atoms and their coordinates one at a time conditioned on the pocket. Diffusion (Schneuing et al., 2024; Guan et al., 2023b;a) and flow-based (Zhou et al., 2025) generation generally improve on them on the docking metrics of CrossDock2020, and Bayesian Flow Networks (Qu et al., 2024; Qiu et al., 2025) improve further on both docking and ligand realism, measured by PoseCheck (Harris et al., 2023) and PoseBusters (Buttenschoen et al., 2024). All of these build equivariance into the architecture. Language models have also been tried for SBDD: BindGPT (Zholus et al., 2024) treats molecules, pockets, and coordinates as sequences, and 3DMolFormer (Hu et al., 2025) represents SMILES as discrete tokens and coordinates as numerical ones. 3DMolFormer, however, works in a ligand-centered frame and needs initial atom coordinates to generate a pose, which hands it information about the target pose.

Ligand-pocket docked-pose generation. This task is addressed by both conventional physics-based docking methods (Trott & Olson, 2010; Koes et al., 2013) and learned models (Peng et al., 2026; Alcaide et al., 2024; Corso et al., 2023; Méndez-Lucio et al., 2021), where the learned models make use of equivariant architectures. For evaluating on this task, these works have used the benchmark PoseBusters (Buttenschoen et al., 2024) primarily to evaluate physical validity alongside pose accuracy of an experimentally derived docked pocket and ligand positions, providing an important check beyond geometric agreement alone. We use this benchmark to evaluate Molweaver's docked-pose generation.

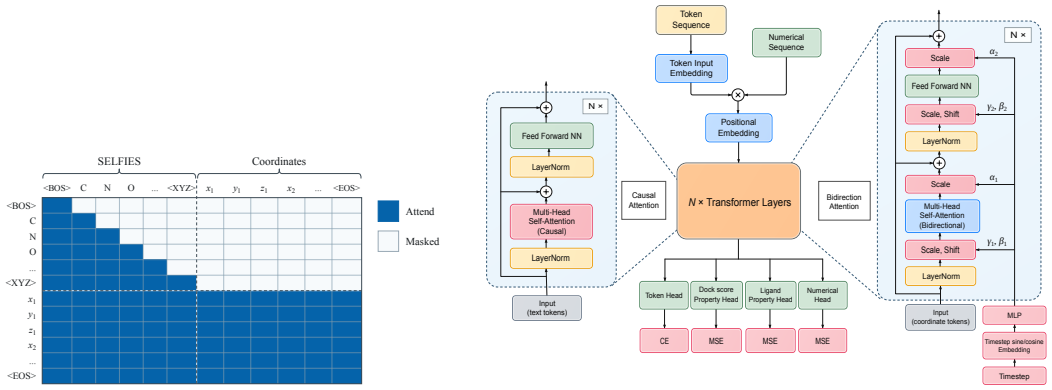
Large-scale molecular pretraining. Molecular language models have explored the concept of large-scale pretraining on molecular data before. Models have been made to learn from large collections of molecular strings such as SMILES and SELFIES (Irwin et al., 2022; Ross et al., 2025; Chitsaz et al., 2025) and have been used mainly in 2D molecular graph generation benchmarks. Furthermore, there exist works (Zhou et al., 2023; Ji et al., 2024) that were jointly pretrained on different modalities, namely 10 million molecular graphs, 1 billion 3D structures generated with the RDKit library, and 3.2 million pockets obtained from PDB (wwPDB consortium, 2019), but these were encoder-based models that were not used for generative tasks, and they even explored scaling laws on the predictive benchmarks. Later works (Alcaide et al., 2024; Zholus et al., 2024; Hu et al., 2025) have adapted the same dataset for SBDD tasks. Our focus is pretraining on a larger dataset and observing whether scaling model sizes trained on this dataset will help us surpass the results of specialist architectures trained for the SBDD tasks.

4 Methodology

4.1 Model Architecture

Molweaver uses a shared GPT-NeoX backbone (Black et al., 2022) for both discrete sequence prediction and continuous coordinate generation, as shown in Figure 2(b). Following Hu et al. (2025) and Golkar et al. (2024), the model reads a token sequence and a numerical sequence in parallel, and each token s_i is embedded and scaled by its value $v_i \in \mathbb{R}$: $\mathbf{h}_i = v_i E(s_i)$. We write ligands as SELFIES, pockets as atom tokens, and atomic coordinates as `[x]`, `[y]`, and `[z]` tokens. SELFIES and pocket tokens have numerical value $v_i = 1$, while coordinate tokens carry the noised coordinate values. Molecular properties and docking scores are handled the same way after standardizing them with the mean and variance of each quantity.

Attention is causal for SELFIES and pocket atom prediction, and bidirectional inside coordinate blocks. We condition coordinate generation on the noise timestep using adaptive layer normalization (AdaLN), following DiT (Peebles & Xie, 2023), which modulates the scale and shift of the normalized activations and the residual contributions of the attention and feed-forward branches. The backbone feeds four heads: a token head for SELFIES and pocket atoms, a numerical head for coordinate velocity, a property head that predicts QED, TPSA, MolWt, MolLogP, and SA together, and a docking-score head (Figure 2b).



(a) Attention mask (b) Molweaver architecture

Figure 2: Molweaver attention mask (causal during SELFIES generation and training, but bidirectional during coordinate pose flow-matching denoising) and transformer architecture.

4.2 PRETRAINING

Sequence construction and attention. We organize each training example into discrete sequence blocks and coordinate blocks, using next-token prediction for the former and flow matching for the latter. Attention is causal between blocks, so each block sees the preceding context but not the later ones. Within a discrete block it is causal, and the token head models the SELFIES as $p_\theta(s \mid c) = \prod_{i=1}^{L} p_\theta(s_i \mid s_{<i}, c)$, where c is whatever was placed before the sequence. Within a coordinate block attention is bidirectional, so all velocities are predicted jointly. Standalone ligands and pockets use $[S_l; X_l]$ and $[A_p; X_p]$, while pocket–ligand complexes use $[A_p; X_p; S_l; X_l]$. Ligand SELFIES in a complex therefore attend to the full pocket, and ligand coordinates additionally attend to the full SELFIES sequence and to their own block. Figure 2(a) shows the transition between the two attention modes. Full sequence formats and masks are in Appendix B.3, and the preprocessing and augmentation we apply to the coordinates are in Appendix B.1.

Coordinate noising and flow matching. Let $X_0 \in \mathbb{R}^{N\times 3}$ be a clean coordinate block after augmentation and scaling. We sample one timestep $t \sim \mathcal{U}(0, 1)$ for the whole block and a standard Gaussian noise matrix ε, and noise the coordinates as described in Section 2.2. The entries of X_t become the numerical values of the `[x]`, `[y]`, and `[z]` tokens. The numerical head predicts the velocity of this interpolation from the noisy coordinates, the timestep, and the preceding context c. The target is $U_t = \partial X_t/\partial t = \varepsilon - X_0$ and the prediction is $\widehat{U}_t = v_\theta(X_t, t, c)$, matched with a squared-error flow-matching objective (Lipman et al., 2022). So the model predicts a velocity, not the clean coordinates or the noise. Each coordinate block starts with an `<XYZ>` token whose numerical value is t, giving the embedding $tE(\texttt{<XYZ>})$, and the same t is passed as a sine/cosine embedding to the AdaLN layers (Appendix A.2).

Property and docking-score supervision. Molecular properties and docking scores appear in two formats. In the conditioning format, the standardized QED, MolWt, MolLogP, SA, TPSA, and Vina values precede the ligand SELFIES and coordinates. They are inputs only and give no regression targets. In the prediction format, the property and docking-score heads predict them after reading the ligand SELFIES, without seeing the values as inputs. Conditioning on the Vina score also lets training account for how much the docking scores differ across the paired examples. Both formats are shown in Appendix B.3.

Pretraining objective. We combine token prediction, coordinate flow matching, and property and docking-score regression. The token loss $\mathcal{L}_{\text{CE}}$ is the usual cross-entropy over the supervised token positions, and $\mathcal{L}_{\text{vel}}$ is the MSE over the velocity, to which both pocket and ligand coordinate blocks contribute when they are training targets. The property and Vina losses, $\mathcal{L}_{\text{prop}}$ and $\mathcal{L}_{\text{Vina}}$, are MSE between the predicted and target values in standardized units. Only prediction targets enter these sums, so values supplied as conditioning inputs are excluded, and a loss is zero when its target set is

empty. The overall loss is

$$\mathcal{L}_{\text{pretrain}} = \mathcal{L}_{\text{CE}} + 10\,\mathcal{L}_{\text{vel}} + 5\,\mathcal{L}_{\text{prop}} + 5\,\mathcal{L}_{\text{Vina}}. \tag{1}$$

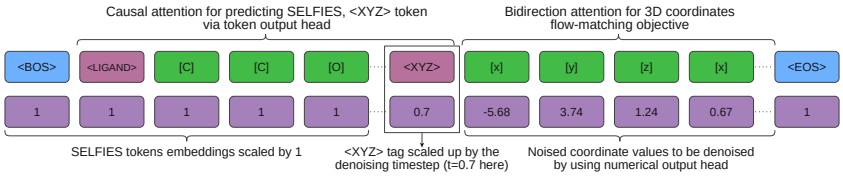


Figure 3: Ligand sequence for pretraining.

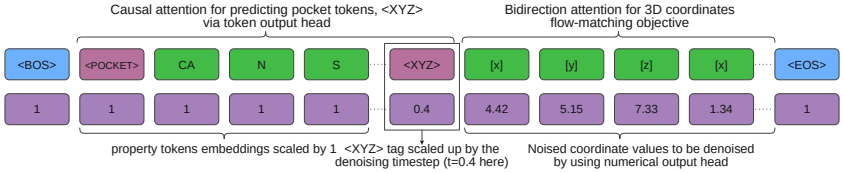


Figure 4: Pocket sequence for pretraining.

4.3 Docked-Pose Fine-Tuning

We fine-tune the pretrained models to predict ligand coordinates given the pocket and the ligand graph. Each example provides the pocket $C_p = (A_p, X_p)$ and the ligand graph G_l as SELFIES, which together form the conditioning context $c = (C_p, G_l)$. Only the ligand coordinates are noised; the pocket and the ligand graph stay fixed. As in pretraining, we sample $t \sim \mathcal{U}(0, 1)$ and Gaussian noise ε, build $X_{l,t} = (1 - t)X_{l,0} + t\varepsilon$ from the clean, augmented and rescaled coordinates $X_{l,0}$, and train with the velocity MSE alone.

Confidence-based pose ranking. After pose fine-tuning we train two confidence models on candidate poses sampled from a small subset of the fine-tuning data. One ranks poses by their RMSD to the reference structure, favoring lower RMSD. The other ranks them by PoseBusters validity and PoseCheck strain, favoring PoseBusters-valid (PB-valid) poses with low strain. Keeping them separate lets us judge pose accuracy apart from physical plausibility. Both are trained independently on pairwise comparisons between poses of the same pocket–ligand complex. For model $k \in \{\text{RMSD}, \text{phys}\}$, with f_k^+ and f_k^- the scores of a preferred and a rejected pose under its own criterion, we use the logistic ranking loss

$$\mathcal{L}_{\text{rank}}^{(k)} = \mathbb{E}_{(+,-)\sim\mathcal{D}_k}\left[\log\left(1 + \exp(f_k^- - f_k^+)\right)\right], \tag{2}$$

where $\mathcal{D}_k$ holds the preference pairs for model k, so a higher score on the rejected pose is penalized.

Guided sampling and pose selection. We use classifier-free guidance (CFG) (Ho & Salimans, 2021) when generating poses, combining the two predictions as in Appendix A.3. At each sampling step we run the model twice on the same noisy coordinates and timestep, once with the pocket and the ligand SELFIES, giving the conditional v_{pocket}, and once with the ligand SELFIES alone, giving the condition-dropped v_{ligand}, so $w = 1$ recovers the pocket-conditioned velocity. Coordinates are then generated from Gaussian noise by DDIM sampling (Song et al., 2021) with the guided velocities. We use $w = 3$ and 20 DDIM steps, chosen from an ablation on the fine-tuned Molweaver-500M checkpoint (Figure 9) that scored one sampled pose per complex on RMSD below 2 Å alone, with no best-of-N sampling and no confidence-based selection. Accuracy rises from roughly 60% at $w = 1$ to about 65% at $w = 3$, then flattens. Steps matter much less, with 20 to 60 nearly indistinguishable and only the 10-step setting degrading sharply at large guidance scales, where the coarse discretization cannot track the amplified velocity field. At test time we draw 250 poses per pocket–ligand pair and keep the one the two confidence models rank best, that is, the lowest predicted RMSD along with PB validity and low strain, which needs no reference ligand coordinates. Drawing many candidates and keeping the best-ranked one is a form of inference-time scaling. The pose is improved by spending more compute at sampling rather than by further training, and the sample budget can be raised or lowered without retraining the model or the confidence models.

4.4 POCKET-CONSTRAINED MOLECULE GENERATION

We optimize molecular graph generation and pose generation in two stages. First, proximal policy optimization (PPO) steers SELFIES generation toward favorable Vina Dock scores, QED, and synthetic accessibility. We then fix the PPO-trained SELFIES generator and optimize the coordinate model using Flow-DPO, with preferences based on direct Vina scores, PoseBusters validity, and PoseCheck strain.

4.4.1 STEERING SELFIES GENERATION

We use PPO (Schulman et al., 2017) to update the pocket-conditioned SELFIES generator using rewards assigned to completed molecules. For each generated ligand G_l, the docked-pose fine-tuned model produces a single pose X_l, which is used for docking evaluation. This is a best-of-1 prediction. The 250-pose confidence-ranking procedure is not used during PPO.

The reward combines the Vina Dock score with QED and SA:

$$R = \frac{1}{3}\left[-V_{\mathrm{dock}}(G_l, X_l; C_p) + \frac{1}{2}\,\mathrm{QED}(G_l) + \frac{1}{2}\,\mathrm{SA}(G_l)\right], \tag{3}$$

where V_{dock} is obtained through an AutoDock Vina docking search (Trott & Olson, 2010), and SA uses the higher-is-better convention reported in Table 1. PPO alternates between sampling SELFIES sequences, evaluating their rewards, and updating the generator using the clipped policy objective described in the Appendix.

4.4.2 STEERING POSE GENERATION

After PPO, we freeze the SELFIES generator and initialize the coordinate model from the docked-pose fine-tuned checkpoint. For each pocket–ligand example, we generate eight candidate poses and evaluate their direct Vina Score, PoseBusters validity, and PoseCheck strain (Harris et al., 2023). We do this mainly to align with Vina scores, since the previously fine-tuned model may not generalize to unseen pockets and to new molecules outside the fine-tuning dataset. The mean Vina score on the PoseBusters benchmark was around -4.2, so that benchmark may not test the maximization of Vina scores adequately. We favor poses with lower direct Vina scores while requiring PB validity and low strain. The resulting pose preferences are used to optimize the coordinate model with Flow-DPO (Jiao et al., 2024), using the docked-pose fine-tuned checkpoint as the fixed reference model. Preference pairs share the same pocket and ligand graph, so this stage optimizes ligand placement and conformation without changing molecular connectivity. The Flow-DPO objective is given in section 2.2.1.

5 EXPERIMENTS AND RESULTS

5.1 DOCKED-POSE PREDICTION

Data and fine-tuning setup. We fine-tune Molweaver on a combination of CrossDock2020, PDBbind (Wang et al., 2004), and Binding MOAD (Hu et al., 2005), using the same data split as Peng et al. (2026). The fine-tuning data is already de-duplicated against the PoseBusters evaluation set so that no evaluation complex is seen at any stage of training. This leaves 85,114 complexes, which we fine-tune on with the flow-matching velocity loss alone, as described in Section 4.3. Preprocessing follows the same method used during pretraining. The pocket and ligand are centered on the pocket centroid, the same random rotation is applied to both, and all coordinates are divided by 5, so the model continues to see the ligand in the pocket frame rather than in a ligand-centered frame. We train for 50K steps with a maximum learning rate of 7.5×10^{-5} decayed on a cosine schedule to 2×10^{-5}, weight decay 0.01, gradient clipping 1.0, and a global batch size of 256.

The two confidence models are trained afterwards on a held-out subset of 1,000 complexes from the same fine-tuning data. For each complex we sample 250 poses from the fine-tuned model, which gives 250,000 candidate poses in total, and form preference pairs within each complex using RMSD to the reference for the first model, and PoseBusters validity together with PoseCheck strain for the second. Both are trained with the logistic ranking loss of Equation 2.

Evaluation. We evaluate on the 428 pocket–ligand complexes of the PoseBusters benchmark (Buttenschoen et al., 2024) and report the percentage of complexes whose predicted pose has a symmetry-corrected heavy-atom RMSD below 2 Å *and* passes every applicable PoseBusters check. This is stricter than RMSD alone, since a pose can land close to the reference while still having distorted bond geometry or clashing with the pocket. Poses are generated as described in Section 4.3: 250 samples per complex with CFG at $w = 3$, followed by confidence-based selection, which uses no reference coordinates and so is available at inference time. With 20 DDIM steps and two forward passes per step for guidance, this budget amounts to $250 \times 20 \times 2 = 10^4$ network evaluations per complex, which dominates the inference cost. Figure 5 reports the results. The results given for Molweaver are the mean of five runs, each with a different seed.

Results. Molweaver-500M reaches 80.2%, ahead of the classical docking programs Vina and Gold(Verdonk et al., 2003) and of the earlier learned docking models EquiBind (Stärk et al., 2022), TankBind (Lu et al., 2022), DeepDock (Méndez-Lucio et al., 2021), and DiffDock (Corso et al., 2023). It is also slightly ahead of the current specialist models, Uni-Mol Docking V2 (Alcaide et al., 2024), SurfDock (Cao et al., 2025) with force-field minimization, and PocketXMol (Peng et al., 2026) with tuned ranking and best-of-100 sampling. Accuracy increases monotonically with model size, from 74.5% at 60M to 76.8% at 100M, 78.3% at 300M, and 80.2% at 500M, with every model trained on the same data for the same number of steps, and even the 60M model is already competitive with the specialist baselines.

5.2 Pocket-Constrained Molecule Generation

Setup. We align the pretrained model for pocket-constrained generation in the two stages of Section 4.4 using PPO on the base pose model to steer SELFIES generation, then Flow-DPO initialized from the docked-pose fine-tuned checkpoint, using the PPO model as the frozen SELFIES prior. Running them in this order matters, since PPO changes which molecules are proposed while Flow-DPO only changes where they are placed, so the pose stage is optimized against the molecule distribution it will actually see at test time. PPO runs for 200 steps with a batch size of 8, and Flow-DPO uses two preference pairs per pocket–ligand example at the same batch size. Both stages use a maximum learning rate of 7.5×10^{-5} decayed on a cosine schedule to 2×10^{-5}, with weight decay 0.01 and gradient clipping 1.0.

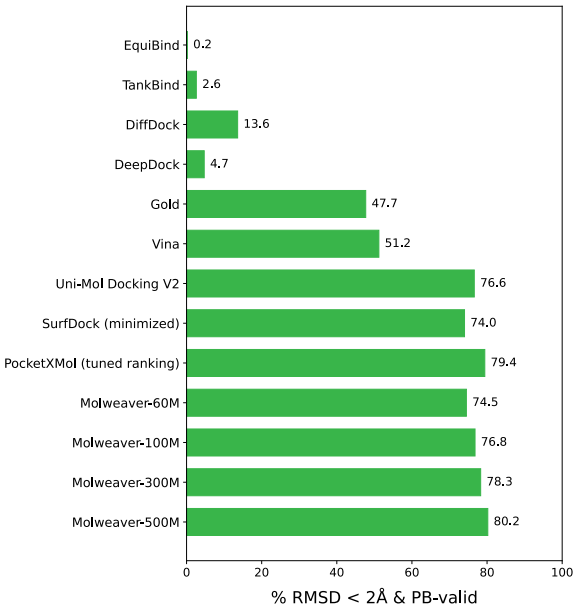


Figure 5: PoseBusters results on 428 complexes: percentage of poses with RMSD below 2 Å that are also PB-valid. Baseline values are taken from Peng et al. (2026).

Evaluation. We evaluate on the CrossDock2020 test set, generating 100 molecules for each of 100 pockets with the 60M, 100M, 300M, and 500M models. Both alignment stages run online against each test pocket. The model proposes molecules and poses, Vina and the PoseBusters and PoseCheck routines score them, and the model is updated on that feedback for the pocket it is being evaluated on. This is per-target optimization rather than zero-shot generation, unlike the baselines in Table 1, and the comparison should be read with that in mind. For every molecule we sample 250 poses and pick one with the PB-valid and low-strain confidence model; the RMSD model is not used here, since a generated molecule has no reference pose to rank against. Table 1 reports docking scores, PoseBusters validity, connectivity, molecular properties, and diversity, with baselines

reproduced from Qiu et al. (2025), and Table 2 reports PoseCheck strain and clashes, with baselines from the results reported by Zhou et al. (2025) and Qiu et al. (2025).

Table 1: Molecular generation results. In the ranked columns, **bold**, underlined, and dotted-underlined values indicate first, second, and third place, respectively. Ties receive the same formatting; CrossDock is not ranked.

Method	PB-valid Avg. (↑)	Vina Score (↓)		Vina Dock (↓)		Connected Avg. (↑)	QED Avg.	SA Avg.	Div. Avg.
		Avg.	Med.	Avg.	Med.				
CrossDock	95.0%	−6.36	−6.46	−7.45	−7.26	–	0.48	0.73	–
Pocket2Mol (Peng et al., 2022)	72.3%	−5.14	−4.70	−7.15	−6.79	96.3%	0.57	0.76	0.69
FLAG (Zhang et al., 2023)	16.0%	45.85	36.52	−4.84	−5.56	97.1%	**0.61**	0.63	0.70
TargetDiff (Guan et al., 2023a)	50.5%	−5.47	−6.30	−7.80	−7.91	90.4%	0.48	0.58	0.72
DiffSBDD (Schneuing et al., 2024)	37.6%	−1.44	−4.91	−7.14	−7.3	93.2%	0.47	0.58	**0.73**
DecompDiff (Guan et al., 2023b)	71.7%	−5.19	−5.27	−7.03	−7.16	82.9%	0.51	0.66	**0.73**
MolCRAFT (Qu et al., 2024)	84.6%	−6.55	−6.95	−7.67	−7.82	96.7%	0.50	0.67	0.72
MolPilot (Qiu et al., 2025)	**95.9%**	−6.88	−7.03	−7.92	−7.92	97.9%	0.56	0.74	0.69
Molweaver-60M	75.7%	−6.43	−6.73	−8.88	−9.12	**100.0%**	0.52	0.73	**0.73**
Molweaver-100M	77.8%	−6.71	−7.03	−9.16	−9.36	**100.0%**	0.55	0.78	0.71
Molweaver-300M	77.6%	−6.89	−7.11	−9.24	−9.47	**100.0%**	0.57	0.76	0.72
Molweaver-500M	79.6%	**−7.22**	**−7.48**	**−9.59**	**−9.78**	**100.0%**	0.57	**0.84**	0.71

Table 2: PoseCheck strain energy and clashes. SE reports the median strain energy, and Clash reports the mean clash count.

Method	SE Median (↓)	Clash Avg. (↓)
TargetDiff (Guan et al., 2023a)	1233	10.77
ALiDiff (Gu et al., 2024)	56055	8.68
KGDiff (Qian et al., 2023)	84186	7.82
PAFlow (Zhou et al., 2025)	90239	7.74
LiGAN (Ragoza et al., 2022)	18693.8	3.40
Pocket2Mol (Peng et al., 2022)	**189**	5.62
MolCRAFT (Qu et al., 2024)	195	7.09
Molweaver-500M	205.8	**1.22**

Results. Molweaver-500M has the lowest mean and median Vina Score and Vina Dock values in Table 1 and places third among PB-validity among other baselines. Connectivity is complete at every model size, a consequence of generating SELFIES rather than atom sets. Mean SA is the highest in the table showcasing the model ability to design better synthesizable compounds. The diversity matches the baselines, so the docking gains do not come from collapsing onto one scaffold, and the docking columns improve monotonically with model size. In Table 2 the same model has the lowest mean clash count of any listed method, below LiGAN and well below the diffusion and flow baselines, while its median strain energy sits close to Pocket2Mol and MolCRAFT and two to three orders of magnitude below every previous diffusion and flow-matching method, namely ALiDiff, KGDiff, and PAFlow. The model therefore places ligands that neither push into the pocket nor strain internally, which is what the low-strain confidence model was trained to select for. Though PB validity is lower than MolCRAFT and MolPilot but still better than other baselines. Further scaling of model size and inference can actually help this metric.

6 Conclusion and future work

We introduced Molweaver, a family of foundation models that learn molecular graphs, 3D geometry, and pocket–ligand structure in one transformer, by pairing autoregressive SELFIES prediction with flow matching over coordinates. On docked-pose prediction the fine-tuned models recover the bound pose well enough to say they have picked up the interactions between a ligand and its pocket, and they do so with no equivariant components anywhere in the architecture, only rotation augmentation during preprocessing. On pocket-constrained generation, feedback from Vina, PoseCheck, and PoseBusters steers the model through PPO and Flow-DPO toward molecules with better docking scores, full connectivity, and fewer clashes than the methods we compare against, though at a PB-valid rate below the equivariant baselines. Accuracy improves with model size on both tasks we trained for. Our model is capable of producing realistic poses with favourable docking scores.

Several things we did not do are worth naming. Our conformers come from ETKDGv3 with MMFF relaxation, which is cheap but well short of experimentally determined or DFT-relaxed structures. Better conformers, larger models, and a pretraining run long enough to find where the scaling flattens are the directions we would pursue next. Optimizing for other properties during pocket-constrained generation is also left for future work.

AI Use Statement

Generative AI tools assisted with language editing, LaTeX formatting, and in helping with development of the codebase.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, Alexander I. Cowen-Rivers, Andrew Cowie, Michael Figurnov, Fabian B. Fuchs, Hannah Gladman, Rishub Jain, Yousuf A. Khan, Caroline M. R. Low, Kuba Perlin, Anna Potapenko, Pascal Savy, Sukhdeep Singh, Adrian Stecula, Ashok Thillaisundaram, Catherine Tong, Sergei Yakneen, Ellen D. Zhong, Michal Zielinski, Augustin Žídek, Victor Bapst, Pushmeet Kohli, Max Jaderberg, Demis Hassabis, and John M. Jumper. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024. doi: 10.1038/s41586-024-07487-w.

Eric Alcaide, Zhifeng Gao, Guolin Ke, Yaqi Li, Linfeng Zhang, Hang Zheng, and Gengmo Zhou. Uni-mol docking v2: Towards realistic and accurate binding pose prediction, 2024. URL https://arxiv.org/abs/2405.11769.

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. doi: 10.1038/nchem.1243. URL https://doi.org/10.1038/nchem.1243.

Sidney Black, Stella Biderman, Eric Hallahan, Quentin Anthony, Leo Gao, Laurence Golding, Horace He, Connor Leahy, Kyle McDonell, Jason Phang, Michael Pieler, Usvsn Sai Prashanth, Shivanshu Purohit, Laria Reynolds, Jonathan Tow, Ben Wang, and Samuel Weinbach. GPT-NeoX-20B: An open-source autoregressive language model. In *Proceedings of BigScience Episode #5—Workshop on Challenges & Perspectives in Creating Large Language Models*, pp. 95–136. Association for Computational Linguistics, 2022. doi: 10.18653/v1/2022.bigscience-1.9. URL https://aclanthology.org/2022.bigscience-1.9/.

Martin Buttenschoen, Garrett M. Morris, and Charlotte M. Deane. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science*, 15(9):3130–3139, 2024. doi: 10.1039/D3SC04185A. URL http://dx.doi.org/10.1039/D3SC04185A.

Duanhua Cao, Mingan Chen, Runze Zhang, Zhaokun Wang, Manlin Huang, Jie Yu, Xinyu Jiang, Zhehuan Fan, Wei Zhang, Hao Zhou, Xutong Li, Zunyun Fu, Sulin Zhang, and Mingyue Zheng. SurfDock is a surface-informed diffusion generative model for reliable and accurate protein–ligand complex prediction. *Nature Methods*, 22(2):310–322, 2025. doi: 10.1038/s41592-024-02516-y. URL https://doi.org/10.1038/s41592-024-02516-y.

Kamran Chitsaz, Roshan Balaji, Quentin Fournier, Nirav Pravinbhai Bhatt, and Sarath Chandar. NovoMolGen: Rethinking molecular language model pretraining. arXiv preprint arXiv:2508.13408, 2025. URL https://arxiv.org/abs/2508.13408v2.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. DiffDock: Diffusion steps, twists, and turns for molecular docking. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2210.01776.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics*, 1:8, 2009. doi: 10.1186/1758-2946-1-8. URL https://doi.org/10.1186/1758-2946-1-8.

Paul G. Francoeur, Tomohide Masuda, Jocelyn Sunseri, Andrew Jia, Richard B. Iovanisci, Ian Snyder, and David R. Koes. 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. doi: 10.1021/acs.jcim.0c00411. URL https://doi.org/10.1021/acs.jcim.0c00411.

Siavash Golkar, Mariel Pettee, Michael Eickenberg, Alberto Bietti, Miles Cranmer, Geraud Krawezik, Francois Lanusse, Michael McCabe, Ruben Ohana, Liam Parker, Bruno Régaldo-Saint Blancard,

Tiberiu Tesileanu, Kyunghyun Cho, and Shirley Ho. xVal: A continuous numerical tokenization for scientific language models. arXiv preprint arXiv:2310.02989, 2024. URL https://arxiv.org/abs/2310.02989.

Siyi Gu, Minkai Xu, Alexander Powers, Weili Nie, Tomas Geffner, Karsten Kreis, Jure Leskovec, Arash Vahdat, and Stefano Ermon. Aligning target-aware molecule diffusion models with exact energy optimization. *Advances in Neural Information Processing Systems*, 37:44040–44063, 2024. URL https://arxiv.org/abs/2407.01648.

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. In *International Conference on Learning Representations*, 2023a. URL https://arxiv.org/abs/2303.03543.

Jiaqi Guan, Xiangxin Zhou, Yuwei Yang, Yu Bao, Jian Peng, Jianzhu Ma, Qiang Liu, Liang Wang, and Quanquan Gu. DecompDiff: Diffusion models with decomposed priors for structure-based drug design. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 11827–11846. PMLR, 2023b. URL https://proceedings.mlr.press/v202/guan23a.html.

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. doi: 10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

Charles Harris, Kieran Didi, Arian R. Jamasb, Chaitanya K. Joshi, Simon V. Mathis, Pietro Liò, and Tom Blundell. Benchmarking generated poses: How rational is structure-based drug design with generative models? arXiv preprint arXiv:2308.07413, 2023. URL https://arxiv.org/abs/2308.07413.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. In *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*, 2021. URL https://arxiv.org/abs/2207.12598. arXiv:2207.12598.

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. doi: 10.1002/prot.20512.

Xiuyuan Hu, Guoqing Liu, Can Chen, Yang Zhao, Hao Zhang, and Xue Liu. 3DMolFormer: A dual-channel framework for structure-based drug discovery. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2502.05107.

Ross Irwin, Spyridon Dimitriadis, Jiazhen He, and Esben Jannik Bjerrum. Chemformer: A pre-trained transformer for computational chemistry. *Machine Learning: Science and Technology*, 3 (1):015022, 2022.

Xiaohong Ji, Zhen Wang, Zhifeng Gao, Hang Zheng, Linfeng Zhang, Guolin Ke, and Weinan E. Uni-Mol2: Exploring molecular pretraining model at scale. arXiv preprint arXiv:2406.14969, 2024. URL https://arxiv.org/abs/2406.14969.

Rui Jiao, Xiangzhe Kong, Wenbing Huang, and Yang Liu. 3d structure prediction of atomic systems with flow-based direct preference optimization. *Advances in Neural Information Processing Systems*, 37:110197–110217, 2024.

Peter Kirkpatrick and Clare Ellis. Chemical space. *Nature*, 432:823–823, 2004. doi: 10.1038/432823a. URL https://doi.org/10.1038/432823a.

David Ryan Koes, Matthew P. Baumgartner, and Carlos J. Camacho. Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. *Journal of Chemical Information and Modeling*, 53(8):1893–1904, 08 2013. ISSN 1549-9596. doi: 10.1021/ci300604z. URL https://doi.org/10.1021/ci300604z.

Mario Krenn, Florian Häse, AkshatKumar Nigam, Pascal Friederich, and Alán Aspuru-Guzik. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. *Machine Learning: Science and Technology*, 1(4):045024, 2020. doi: 10.1088/2632-2153/aba947. URL https://arxiv.org/abs/1905.13741.

Greg Landrum and others. RDKit: Open-source cheminformatics. https://www.rdkit.org. Accessed: 2026-09-26.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. arXiv preprint arXiv:2210.02747, 2022. URL https://arxiv.org/abs/2210.02747.

Meng Liu, Youzhi Luo, Kanji Uchino, Koji Maruhashi, and Shuiwang Ji. Generating 3D molecules for target protein binding. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 13912–13924. PMLR, 2022. URL https://proceedings.mlr.press/v162/liu22m.html.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Bkg6RiCqY7.

Wei Lu, Qifeng Wu, Jixian Zhang, Jiahua Rao, Chengtao Li, and Shuangjia Zheng. TankBind: Trigonometry-aware neural networks for drug-protein binding structure prediction. *Advances in Neural Information Processing Systems*, 35, 2022.

Shitong Luo, Jiaqi Guan, Jianzhu Ma, and Jian Peng. A 3D generative model for structure-based drug design. In *Advances in Neural Information Processing Systems*, volume 34, pp. 6229–6239. Curran Associates, Inc., 2021. URL https://arxiv.org/abs/2203.10446.

Oscar Méndez-Lucio, Mazen Ahmad, Ehecatl Antonio del Rio-Chanona, and Jörg Kurt Wegner. A geometric deep learning approach to predict binding conformations of bioactive molecules. *Nature Machine Intelligence*, 3(12):1033–1039, 2021. doi: 10.1038/s42256-021-00409-9. URL https://doi.org/10.1038/s42256-021-00409-9.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4195–4205, 2023. URL https://arxiv.org/abs/2212.09748.

Xingang Peng, Shitong Luo, Jiaqi Guan, Qi Xie, Jian Peng, and Jianzhu Ma. Pocket2Mol: Efficient molecular sampling based on 3D protein pockets. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 17644–17655. PMLR, 2022. URL https://proceedings.mlr.press/v162/peng22b.html.

Xingang Peng, Ruihan Guo, Fenglin Guo, Ziyi Wang, Jiayu Sun, Jiaqi Guan, Yinjun Jia, Yan Xu, Yanwen Huang, Muhan Zhang, Jian Peng, Xinquan Wang, Chuanhui Han, Zihua Wang, and Jianzhu Ma. Unified modeling of 3d molecular generation via atomic interactions with pocketxmol. *Cell*, 189(7):1904–1922.e28, February 2026. ISSN 0092-8674. doi: 10.1016/j.cell.2026.01.003. URL http://dx.doi.org/10.1016/j.cell.2026.01.003.

Hao Qian, Wenjing Huang, Shikui Tu, and Lei Xu. KGDiff: towards explainable target-aware molecule generation with knowledge guidance. *Briefings in Bioinformatics*, 25(1):bbad435, 12 2023. ISSN 1477-4054. doi: 10.1093/bib/bbad435. URL https://doi.org/10.1093/bib/bbad435.

Keyue Qiu, Yuxuan Song, Zhehuan Fan, Peidong Liu, Zhe Zhang, Mingyue Zheng, Hao Zhou, and Wei-Ying Ma. Piloting structure-based drug design via modality-specific optimal schedule. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 50619–50644. PMLR, 2025. URL https://arxiv.org/abs/2505.07286.

Yanru Qu, Keyue Qiu, Yuxuan Song, Jingjing Gong, Jiawei Han, Mingyue Zheng, Hao Zhou, and Wei-Ying Ma. MolCRAFT: Structure-based drug design in continuous parameter space. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 41749–41768. PMLR, 2024. URL https://arxiv.org/abs/2404.12141.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://arxiv.org/abs/2305.18290.

Matthew Ragoza, Tomohide Masuda, and David Ryan Koes. Generating 3d molecules conditional on receptor binding sites with deep generative models. *Chemical Science*, 13(9):2701–2713, 2022. doi: 10.1039/D1SC05976A. URL https://doi.org/10.1039/D1SC05976A.

Sereina Riniker and Gregory A. Landrum. Better informed distance geometry: Using what we know to improve conformation generation. *Journal of Chemical Information and Modeling*, 55(12): 2562–2574, December 2015. doi: 10.1021/acs.jcim.5b00654. URL https://doi.org/10.1021/acs.jcim.5b00654.

Jerret Ross, Brian Belgodere, Samuel C. Hoffman, Vijil Chenthamarakshan, Jiri Navratil, Youssef Mroueh, and Payel Das. GP-MoLFormer: A foundation model for molecular generation. *Digital Discovery*, 4(10):2684–2696, 2025. doi: 10.1039/D5DD00122F. URL https://doi.org/10.1039/D5DD00122F.

Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom L Blundell, Pietro Lio, Max Welling, Michael Bronstein, and Bruno Correia. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4(12):899–909, Dec 2024. ISSN 2662-8457. doi: 10.1038/s43588-024-00737-x. URL https://doi.org/10.1038/s43588-024-00737-x.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017. URL https://arxiv.org/abs/1707.06347.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL https://arxiv.org/abs/2010.02502.

Hannes Stärk, Octavian Ganea, Lagnajit Pattanaik, Regina Barzilay, and Tommi Jaakkola. EquiBind: Geometric deep learning for drug binding structure prediction. In *International Conference on Machine Learning*, pp. 20503–20521. PMLR, 2022.

Benjamin I. Tingle, Khanh G. Tang, Mar Castanon, John J. Gutierrez, Munkhzul Khurelbaatar, Chinzorig Dandarchuluun, Yurii S. Moroz, and John J. Irwin. ZINC-22—a free multi-billion-scale database of tangible compounds for ligand discovery. *Journal of Chemical Information and Modeling*, 63(4):1166–1176, 2023. doi: 10.1021/acs.jcim.2c01253. URL https://doi.org/10.1021/acs.jcim.2c01253.

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. doi: 10.1002/jcc.21334. URL https://doi.org/10.1002/jcc.21334.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pp. 5998–6008, 2017. URL https://arxiv.org/abs/1706.03762.

Marcel L. Verdonk, Jason C. Cole, Michael J. Hartshorn, Christopher W. Murray, and Richard D. Taylor. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, 52(4):609–623, 2003. doi: 10.1002/prot.10465.

Renxiao Wang, Xueliang Fang, Yipin Lu, and Shaomeng Wang. The PDBbind database: Collection of binding affinities for protein–ligand complexes with known three-dimensional structures. *Journal of Medicinal Chemistry*, 47(12):2977–2980, 2004. doi: 10.1021/jm030580l.

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, 02 1988. ISSN 0095-2338. doi: 10.1021/ci00057a005. URL https://doi.org/10.1021/ci00057a005.

wwPDB consortium. Protein data bank: the single global archive for 3d macromolecular structure data. *Nucleic Acids Research*, 47(D1):D520–D528, 01 2019. ISSN 0305-1048. doi: 10.1093/nar/gky949. URL https://doi.org/10.1093/nar/gky949.

Zaixi Zhang, Yaosen Min, Shuxin Zheng, and Qi Liu. Molecule generation for target protein binding with structural motifs. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=Rq13idF0F73.

Artem Zholus, Maksim Kuznetsov, Roman Schutski, Rim Shayakhmetov, Daniil Polykovskiy, Sarath Chandar, and Alex Zhavoronkov. BindGPT: A scalable framework for 3D molecular design via language modeling and reinforcement learning. arXiv preprint arXiv:2406.03686, 2024. URL https://arxiv.org/abs/2406.03686.

Chunting Zhou, Lili Yu, Arun Babu, Kushal Tirumala, Michihiro Yasunaga, Leonid Shamis, Jacob Kahn, Xuezhe Ma, Luke Zettlemoyer, and Omer Levy. Transfusion: Predict the next token and diffuse images with one multi-modal model. arXiv preprint arXiv:2408.11039, 2024. URL https://arxiv.org/abs/2408.11039.

Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-Mol: A universal 3D molecular representation learning framework. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=6K2RM6wVqKu.

Jingyuan Zhou, Hao Qian, Shikui Tu, and Lei Xu. Prior-guided flow matching for target-aware molecule design with learnable atom number. In *Advances in Neural Information Processing Systems*, 2025. URL https://arxiv.org/abs/2509.01486.

A Additional Background

A.1 Evaluation criteria

Pose accuracy and validity. For predicted coordinates $\widehat{X}_l$ and reference coordinates $X_l^\star$, the symmetry-corrected heavy-atom RMSD (Root Mean Square Deviation) is

$$\mathrm{RMSD}_{\mathrm{sym}}(\widehat{X}_l, X_l^\star) = \min_{\pi \in \mathcal{S}(G_l)} \left(\frac{1}{|\mathcal{H}_l|} \sum_{i \in \mathcal{H}_l} \left\| \widehat{x}_{l,i} - x^\star_{l,\pi(i)} \right\|_2^2 \right)^{1/2}, \tag{4}$$

where $\mathcal{H}_l$ denotes ligand heavy atoms and $\mathcal{S}(G_l)$ denotes symmetry-equivalent atom correspondences. Both poses are expressed in the same pocket frame, without independently aligning the ligand. PoseBusters validity is assessed separately from RMSD through the applicable molecular consistency, intramolecular geometry, and intermolecular clash checks (Buttenschoen et al., 2024).

Docking and molecular properties. Vina Score evaluates the generated pose, whereas Vina Dock reports the best score found through a specified docking search (Trott & Olson, 2010). Lower values are preferred for both; a favorable Vina Dock score does not establish the quality of the original generated pose. Higher QED indicates greater drug-likeness (Bickerton et al., 2012). The original synthetic accessibility score, $\mathrm{SA}_{\mathrm{raw}} \in [1, 10]$, is lower for molecules estimated to be easier to synthesize (Ertl & Schuffenhauer, 2009). A higher-is-better rescaling is $\mathrm{SA}_{\mathrm{norm}} = (10 - \mathrm{SA}_{\mathrm{raw}})/9$.

A.2 Adaptive LayerNorm (AdaLN)

Adaptive LayerNorm conditions a transformer block on the flow timestep through learned scale and shift parameters (Peebles & Xie, 2023). For a hidden state $h \in \mathbb{R}^{d_h}$ at one token position, with $d_h = d_{\mathrm{model}}$ the model width and $\mathbf{1} \in \mathbb{R}^{d_h}$ the all-ones vector, write the normalization as

$$\mathrm{LN}(h) = \frac{h - \mu(h)\mathbf{1}}{\sqrt{\sigma^2(h) + \epsilon_{\mathrm{LN}}}}, \qquad \mu(h) = \frac{1}{d_h} \sum_{j=1}^{d_h} h_j, \qquad \sigma^2(h) = \frac{1}{d_h} \sum_{j=1}^{d_h} (h_j - \mu(h))^2, \tag{5}$$

with $\epsilon_{\mathrm{LN}} > 0$ for numerical stability, where $\mu(h)$ and $\sigma^2(h)$ are the mean and variance over the d_h components of that token's hidden state rather than over the batch or the sequence. The timestep t is embedded with sine and cosine features and passed through an MLP that produces six vectors in $\mathbb{R}^{d_h}$, namely $\gamma_1, \beta_1, \alpha_1$ for the attention branch and $\gamma_2, \beta_2, \alpha_2$ for the feed-forward branch:

$$(\gamma_1, \beta_1, \alpha_1, \gamma_2, \beta_2, \alpha_2) = \mathrm{MLP}_\theta(e(t)) \,. \tag{6}$$

Each pair (γ_i, β_i) rescales and shifts the normalized activations, and each α_i gates the residual contribution of its branch. With $\odot$ denoting element-wise multiplication, a block updates h as

$$\begin{aligned} h' &= h + \alpha_1 \odot \mathrm{Attn}((\mathbf{1} + \gamma_1) \odot \mathrm{LN}(h) + \beta_1) \,, \\ h'' &= h' + \alpha_2 \odot \mathrm{FFN}((\mathbf{1} + \gamma_2) \odot \mathrm{LN}(h') + \beta_2) \,. \end{aligned} \tag{7}$$

Writing the scale as $\mathbf{1} + \gamma_i$ means the block reduces to a standard pre-norm transformer layer when the modulation outputs are zero, which is how we initialize them, so the timestep conditioning starts as the identity and is learned from there. Only coordinate blocks are modulated this way. Discrete blocks use ordinary LayerNorm.

A.3 Classifier-Free Guidance (CFG)

Classifier-free guidance combines a conditional and a condition-dropped prediction without a separate classifier. We use no condition dropout schedule during training, because the pretraining mixture already supplies both predictions: standalone ligand examples $[S_l; X_l]$ teach a velocity conditioned on the ligand alone, and pocket–ligand examples $[A_p; X_p; S_l; X_l]$ teach one conditioned on the pocket as well (Section 4.2). Since attention is causal between blocks, dropping the leading pocket block leaves the ligand blocks unchanged, so a single network supplies both branches at sampling time and the ligand-only behavior learned in pretraining is retained through docked-pose fine-tuning. Write v_θ

for the predicted velocity, x_t for the noised ligand coordinates at timestep t, $c = (C_p, G_l)$ for the full conditioning context, and $w \geq 0$ for the guidance scale. The guided velocity is

$$v_\theta^{\mathrm{CFG}}(x_t, t, c) = v_\theta(x_t, t, \varnothing) + w\left[v_\theta(x_t, t, c) - v_\theta(x_t, t, \varnothing)\right], \tag{8}$$

where $\varnothing$ denotes the same input with the pocket block $[A_p; X_p]$ removed and the ligand SELFIES retained, so $v_\theta(x_t, t, c) = v_{\mathrm{pocket}}$ and $v_\theta(x_t, t, \varnothing) = v_{\mathrm{ligand}}$ of Section 4.3. The scale $w = 1$ recovers the conditional field, while $w > 1$ amplifies the conditional difference.

A.4 Reinforcement Learning

Reinforcement learning optimizes a conditional policy $\pi_\theta(y \mid c)$ using a reward $R(y, c)$ on completed outputs, where θ are the model parameters, c is the conditioning context and y the completed output. In our setting c is the pocket C_p, y is the generated ligand SELFIES S_l, R is the reward of Equation 3.

A.4.1 Proximal Policy Optimization (PPO)

PPO updates a policy using rollouts collected from a previous policy π_{old} (Schulman et al., 2017). For autoregressive generation, the state s_k contains the context and generated prefix, and the action a_k is the next token. Defining $\rho_k(\theta) = \pi_\theta(a_k \mid s_k)/\pi_{\mathrm{old}}(a_k \mid s_k)$, the clipped policy objective is

$$\mathcal{J}_{\mathrm{PPO}}(\theta) = \widehat{\mathbb{E}}_k\left[\min\left(\rho_k(\theta)\widehat{A}_k,\ \mathrm{clip}\left(\rho_k(\theta), 1-\epsilon, 1+\epsilon\right)\widehat{A}_k\right)\right], \tag{9}$$

where $\widehat{A}_k$ estimates the advantage under π_{old}, $\widehat{\mathbb{E}}_k$ averages over collected rollouts, and $0 < \epsilon < 1$ controls clipping. Maximizing this objective limits the incentive for large policy changes, rather than imposing a hard KL constraint.

B Training information

B.1 Dataset preparation and augmentation

We used the first 100 million SMILES from the NovoMol-Gen dataset (Chitsaz et al., 2025), which was derived from ZINC22 (Tingle et al., 2023) and had already been filtered for diversity and validity. We converted each SMILES to SELFIES and generated 10 conformers per molecule using ETKDGv3 (Riniker & Landrum, 2015) in RDKit, followed by MMFF (Halgren, 1996) relaxation. This gave us 1 billion ligand conformers for pretraining.

Table 3: Repetition counts in the pretraining mixture.

Dataset	Repeats per example
SELFIES	10
Ligand conformers	1
Unpaired pockets	10
CrossDock2020	20

We also used the 3.2 million preprocessed unpaired protein pockets from Uni-Mol (Zhou et al., 2023). For paired examples, we used CrossDock2020 (Francoeur et al., 2020), a synthetic dataset of docked poses generated using Vina. We kept only examples below the 1 Å filtering threshold and examples that were converted to SELFIES and our pocket format without error, leaving approximately 167,000 pocket–ligand pairs. We repeated examples in the pretraining mixture according to Table 3.

Every example is centered, rotated, and rescaled, in that order. Standalone ligand conformers and unpaired pockets are centered at their own geometric centroids. The centroid is computed as the mean of the atomic coordinates along each axis:

$$\mathbf{c}_s = \begin{pmatrix}\bar{x}_s\\ \bar{y}_s\\ \bar{z}_s\end{pmatrix} = \frac{1}{N_s}\sum_{i=1}^{N_s}\begin{pmatrix}x_{s,i}\\ y_{s,i}\\ z_{s,i}\end{pmatrix}, \qquad s \in \{l, p\}, \tag{10}$$

where N_l is the number of atoms in the ligand graph G_l, and N_p is the number of pocket atoms represented by A_p. For standalone examples, the clean training coordinates are

$$X_{s,0} = \frac{1}{5}\left(X_s - \mathbf{1}_{N_s}\mathbf{c}_s^\top\right)R^\top, \qquad s \in \{l, p\}, \tag{11}$$

where $\mathbf{1}_{N_s}$ is a column vector of ones and $R \in \mathrm{SO}(3)$ is a randomly sampled rotation matrix. Thus, we subtract the centroid from every atom, apply a global rotation, and divide the coordinates by 5.

For CrossDock2020 pocket–ligand pairs, we use the pocket centroid to center both structures, rather than centering the ligand independently. We also apply the same rotation to the pocket and ligand:

$$X_{p,0} = \frac{1}{5}\left(X_p - \mathbf{1}_{N_p}\mathbf{c}_p^\top\right) R^\top,$$
$$X_{l,0} = \frac{1}{5}\left(X_l - \mathbf{1}_{N_l}\mathbf{c}_p^\top\right) R^\top. \tag{12}$$

Using a shared center and rotation keeps the ligand in the pocket coordinate frame and preserves their relative placement and orientation. A fresh R is drawn uniformly from $\mathrm{SO}(3)$ for every example and applied to the whole coordinate set, so internal geometry is unchanged, and the division by 5 is applied after centering and rotation in all cases.

B.2 Pretraining configuration

All Molweaver models share the same GPT-NeoX backbone (Black et al., 2022), a 352-token vocabulary, an attention head dimension of 64, and identical data and optimization settings; they differ only in depth and width (Table 4). Parameter counts include embeddings, AdaLN modulation, and the four output heads.

We pretrain each model for one epoch ($\sim$290K steps) on eight MI300X GPUs using fused AdamW (Loshchilov & Hutter, 2019), with a learning rate of 10^{-4} on a constant schedule with warmup ratio 0.01, weight decay 0.1, and gradient clipping at 1.0. Ligand-only examples are centered at the ligand centroid and paired examples at the pocket centroid; we apply a fresh random rotation in $\mathrm{SO}(3)$ and divide coordinates by 5.0.

Sequence lengths span more than an order of magnitude, from a standalone ligand of a few tens of tokens to a pocket–ligand complex of $4{,}356$ tokens, so a single batch size would either waste most of the compute on padding or be limited by the longest examples. We therefore group examples into length buckets, draw each update from one bucket, and halve the per-GPU batch size as the bucket length doubles, which keeps the padded token count per GPU roughly constant (Table 5).

Table 4: Molweaver model configurations. Head features is the shared bottleneck width feeding the output heads.

Model	Params	Layers	d_{model}	Heads	FFN width	Head features
Molweaver-60M	58.79M	7	768	12	1,536	3,072
Molweaver-100M	101.32M	13	768	12	1,536	3,072
Molweaver-300M	301.12M	17	1,024	16	4,096	4,096
Molweaver-500M	508.30M	23	1,152	18	4,608	4,608

Table 5: Pretraining hyperparameters and length-bucket batching. Effective batch sizes count examples per optimizer update across eight GPUs.

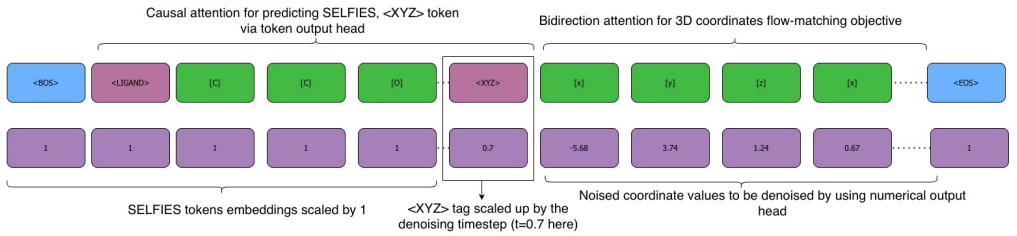
Optimization and execution

Parameter	Value
GPUs	8× MI300X
Epochs	1 ($\sim$290K steps)
Optimizer	`adamW`
Learning rate	10^{-4}
Weight decay	0.1
LR schedule	Constant with warmup
Warmup ratio	0.01
Gradient accumulation	1
Gradient clipping	1.0

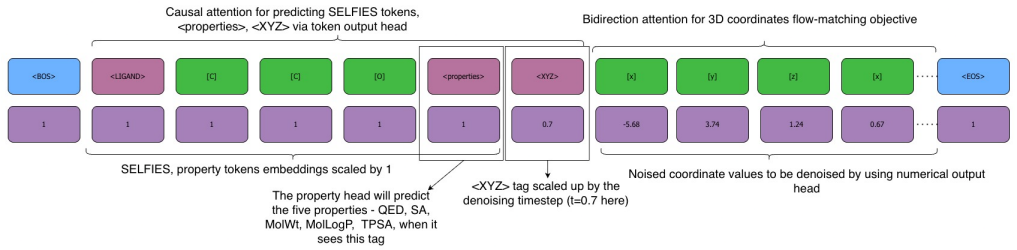
Length-bucket batching

Maximum sequence length	Per-GPU batch size	Effective batch size
162	256	4,096
256	256	4,096
512	128	2,048
1,024	64	1,024
2,048	32	512
4,356	32	512

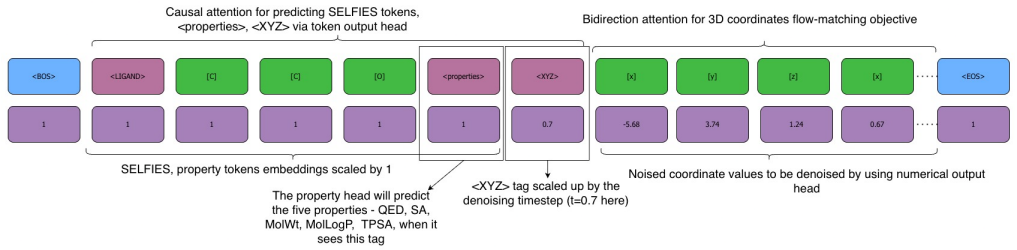
B.3 INPUT SEQUENCES AND ATTENTION MASK



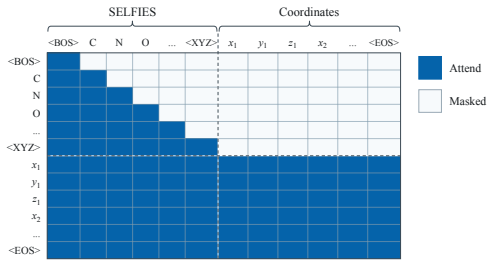
(a) Ligand-only sequence



(b) Ligand sequence with properties to be predicted

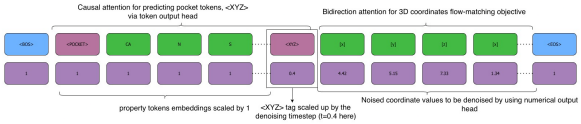


(c) Ligand sequence with properties in context

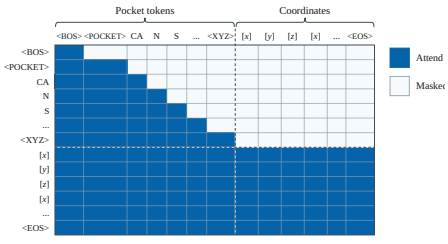


(d) Attention mask of the ligand sequence

Figure 6: Different types of ligand sequence and the attention mask.

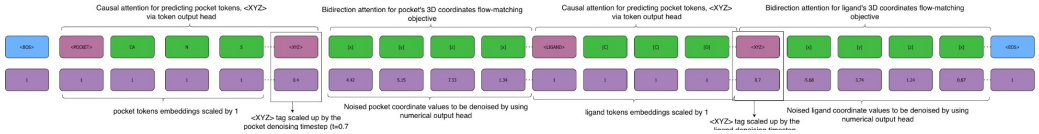


(a) Pocket sequence

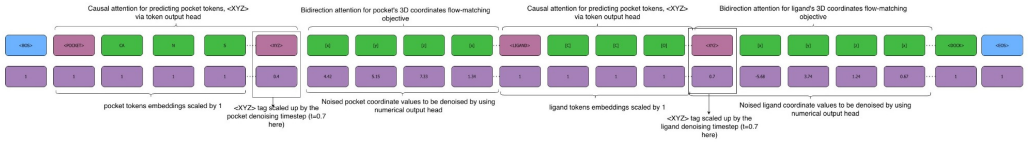


(b) Pocket attention mask

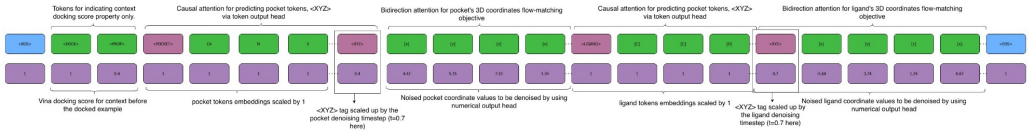
Figure 7: Pocket input sequence and attention mask.



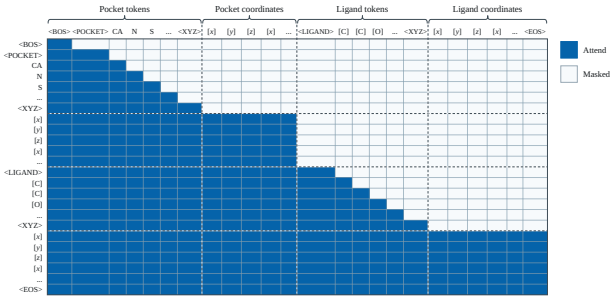
(a) Docked sequence



(b) Docked sequence with the docking score to be predicted



(c) Docked sequence with the docking score in context



(d) Attention mask of the docked sequence

Figure 8: Docked input sequence and attention mask.

B.4 CFG SCALE ABLATION

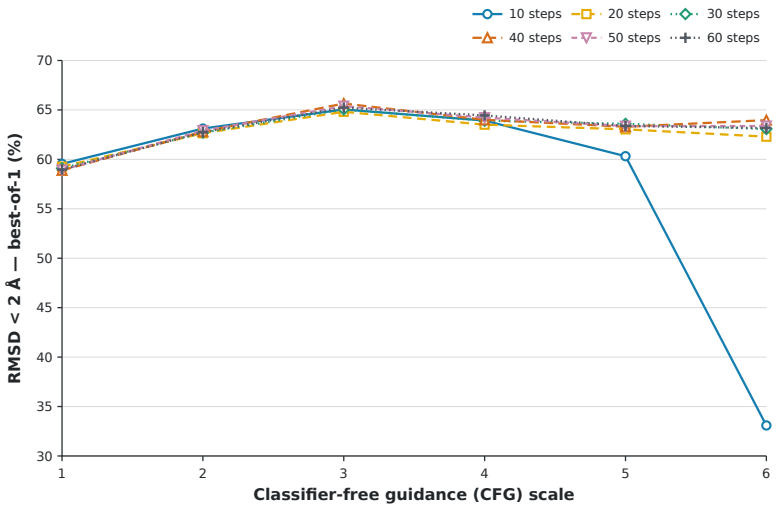


Figure 9: Sampling ablation for the fine-tuned Molweaver-500M checkpoint. The percentage of best-of-1 poses with RMSD below 2 Å is plotted against the classifier-free guidance scale for different numbers of sampling steps.

C DOCKED EXAMPLES

This section contains the docked positions of generated ligands sampled from CrossDock2020 test pockets. The poses shown were generated directly by Molweaver-500M.

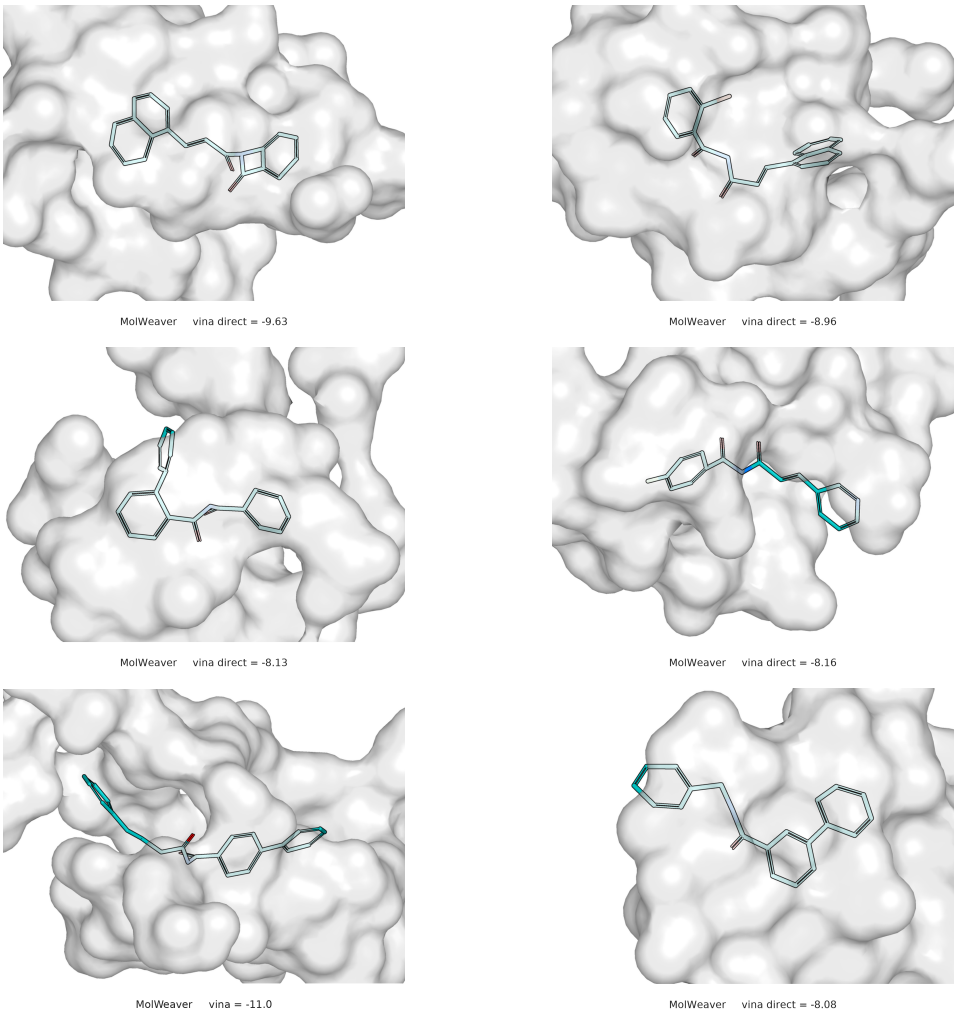


Figure 10: Close-up views of molecules generated by Molweaver-500M in six CrossDock2020 test pockets, each shown in the pose the model produced directly, with no post-hoc redocking or minimization.