PIPCDiff: Pocket-Informed Pharmacophore Conditioning for Diffusion-Based Molecular Generation

Qi Zhang[1], Ningxin Qin[1], Chenyang Liu[3], Chenran Jiang[3], jiean Chen[3], Bin Zeng[2(✉)], and Bingding Huang[1(✉)]

[1] School of Artificial Intelligence, Shenzhen Technology University, Shenzhen, China
huangbingding@sztu.edu.cn
[2] College of Pharmacy, Shenzhen Technology University, Shenzhen, China
zengbin@sztu.edu.cn
[3] High-throughput Screening Center, Shenzhen Bay Laboratory, Shenzhen 518132, China

Abstract. De novo molecular generation for structure-based drug design (SBDD) aims to generate ligand structures that are not only geometrically compatible with a protein pocket but also functionally aligned with its interaction requirements. Recent diffusion-based 3D generative models have achieved substantial progress in modeling pocket–ligand spatial relationships. However, they typically emphasize geometric fitting and do not explicitly encode target-specific pharmacophore preferences derived from the pocket itself. As results, the generated molecules may satisfy spatial constraints while remaining suboptimal in functional feature arrangement. Here we propose PIPCDiff, a Pocket-Informed Pharmacophore Conditioning Diffusion framework built upon DiffGui. PIPCDiff introduces a pocket-adaptive functional prior into diffusion-based molecular generation by learning target-specific pharmacophore distributions directly from protein pocket representations, without relying on reference ligands or handcrafted pharmacophore templates. To preserve this functional guidance during generation, the inferred prior is injected into the denoising network through an implicit timestep-wise conditioning mechanism. In addition, an auxiliary atom-level pharmacophore supervision head is incorporated during training to strengthen local functional representation learning. Experiments on the PDBbind dataset demonstrate that PIPCDiff consistently improves pharmacophore match rate (PMR) and 3D structural feasibility relative to representative baselines, while achieving improved Vina scores and competitive molecular property metrics as well. These results indicate that pocket-derived pharmacophore priors can serve as an effective functional constraint for diffusion-based ligand generation and provide a practical strategy for improving structure-function alignment in SBDD.

Keywords: Structure-Based Drug Design · Diffusion Models · Pharmacophore Modeling · 3D Molecular Generation

D.-S. Huang et al. (Eds.): ICIC 2026, LNBI 16669, pp. 1–13, 2027.
https://doi.org/10.1007/978-981-92-3492-9_38

1 Introduction

Drug discovery remains a long-cycle, high-cost, and high-risk process [1, 2], in which the accurate modeling of protein–ligand interactions is one of the central bottlenecks. Within this context, structure-based drug design (SBDD) seeks to exploit the three-dimensional architecture of target binding sites to guide the identification and optimization of small molecules [3, 4]. As a key task in SBDD, de novo molecular design aims to generate novel ligands for a given protein pocket that simultaneously satisfy binding-site compatibility, chemical validity, and favorable drug-like properties. Recent advances in 3D deep generative modeling have significantly improved this process by enabling the joint modeling of atomic identities, molecular topology, and spatial coordinates.

Among current generative paradigms, denoising diffusion probabilistic models (DDPM) have emerged as one of the most effective frameworks for 3D molecular generation [5]. When combined with E(3)-equivariant neural networks [6, 7], diffusion models can iteratively generate ligand structures under pocket constraints while preserving geometric consistency. Pocket-conditioned models such as DiffGui [8] further improve this paradigm by jointly modeling atom and bond generation and incorporating additional conditional signals related to molecular properties. Despite these advances, most existing approaches remain primarily oriented toward geometric compatibility and global property guidance, whereas the explicit modeling of functional constraints within the pocket is still limited.

In practical medicinal chemistry, geometric fit alone is rarely sufficient. Productive binding depends not only on whether a molecule can occupy the pocket, but also on whether its functional groups are arranged in a way that matches the interaction preferences of the target environment, such as hydrogen-bond donor/acceptor patterns, aromatic features, hydrophobic regions, and ionizable sites. These interaction-relevant chemical features are naturally summarized by the concept of a pharmacophore [9, 10]. Therefore, incorporating pharmacophore-level information into molecular generation offers a promising route toward better structure–function consistency.

However, existing pharmacophore-aware generative strategies still suffer from important limitations (Fig. 1). First, the extraction of pharmacophore priors often relies on heuristic rules, predefined templates, or known reference ligands [11]. While such strategies can be useful in well-characterized systems, they become less suitable when prior ligand knowledge is limited, especially for apo or weakly annotated pockets. Second, even when pharmacophore information is used as a condition in generative models, it is often introduced in a relatively static manner, for example as an initial prompt or a one-off condition. In diffusion models, such a strategy may lead to progressive weakening of functional guidance during long reverse denoising trajectory, particularly when the denoising objective is dominated by geometric reconstruction. Consequently, the final generated molecules may only partially preserve the intended structure–function relationship.

To address these issues, we propose PIPCDiff (Pocket-Informed Pharmacophore Conditioning Diffusion), a pharmacophore-aware diffusion framework built upon DiffGui. The central idea of PIPCDiff is to infer a target-specific pharmacophore prior

distribution directly from the protein pocket, and then inject this prior into the denoising network throughout the reverse diffusion process. In this way, functional information is not treated as an external, predefined prompt, but as a pocket-derived latent prior that continuously guides molecular generation. In addition, PIPCDiff introduces an atom-level auxiliary supervision branch that aligns denoising representations with pharmacophore annotations, thereby strengthening local functional semantics during training. More specifically, PIPCDiff contains two collaborative components. First, the Pocket Pharmacophore Prior Network (P3N) encodes pocket atomic features and predicts a graph-level pharmacophore distribution without requiring reference ligands, which serves as the continuous conditioning signal for the denoising network. Second, the Atom Pharmacophore Supervision Head (APSH) provides atom-level pharmacophore supervision during training, encouraging the model to learn representations that are not only geometrically meaningful but also functionally interpretable. Together, these components enable PIPCDiff as a diffusion-based SBDD framework that strengthens ligand generation by bridging pocket geometry, target-adaptive pharmacophore priors, and persistent functional conditioning within a unified equivariant generative model.

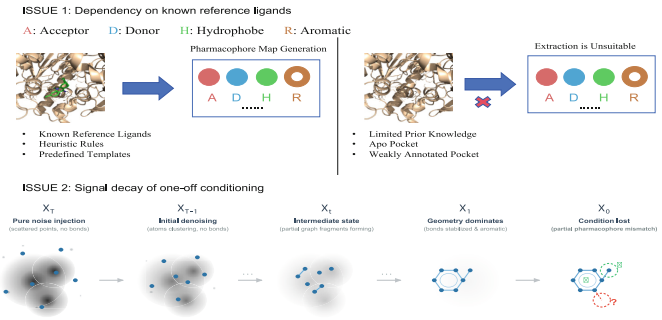


Fig. 1. Limitations of existing pharmacophore-aware generative strategies.

2 Methodology

2.1 Standard Diffusion Formulation and DiffGui Backbone

In the context of structure-based drug design, given a protein pocket P, our objective is to model the conditional distribution $p(M| P)$ for generating a ligand molecule $M = (x_M, v_M)$, where $x \in \mathbb{R}^3$ denotes atomic coordinates and $v \in \mathbb{R}^d$ denotes atom-type features. We adopt the Denoising Diffusion Probabilistic Model (DDPM) architecture, which consists of two processes:

Forward diffusion process: a Markov chain progressively perturbs clean data M_0 with Gaussian noise until it approaches a simple prior $M_T \sim \mathcal{N}(0, I)$:

$$q(M_{1:T} \mid M_0) = \prod_{t=1}^{T} q(M_t \mid M_{t-1}) \tag{1}$$

Reverse generation process: a parameterized model p_θ iteratively denoises from the noisy states under the pocket condition P:

$$p_\theta(M_{0:T} \mid P) = p(M_T)\prod_{t=1}^{T} p_\theta(M_{t-1} \mid M_t, P) \tag{2}$$

Our method is built upon DiffGui, which unifies atom diffusion and bond diffusion within a single denoising framework. The backbone consists of three major stages: (1) Conditional encoding: an E(3)-equivariant graph neural network is used to encode the pocket geometric and interaction context, while conditional embeddings are fused into the latent representation to provide target-aware guidance [12, 13]; (2) Property guidance: molecular property information (e.g., binding affinity and drug-likeness) is incorporated as conditional variables $\boldsymbol{c}$ during both training and sampling, enabling the diffusion process to favor molecules with improved pharmacological properties (Eq. 3); (3) Diffusion generation: ligand atomic coordinates and bond types are jointly predicted during the reverse denoising process, allowing the model to generate structurally coherent and chemically plausible 3D molecules under pocket constraints (Eq. 4).

$$p_\theta(M_{0:T} \mid P, c) = p(M_T)\prod_{t=1}^{T} p_\theta(M_{t-1} \mid M_t, P, c) \tag{3}$$

$$L = L_{\text{pos}} + L_{\text{node}} + L_{\text{edge}} + \lambda_{\text{len}} L_{\text{bond-len}} \tag{4}$$

2.2 Overall Framework of PIPCDiff

As illustrated in Fig. 2, PIPCDiff extends the DiffGui backbone by introducing pocket-derived pharmacophore priors and integrating them into the denoising process through an implicit conditional modulation mechanism. This design enables the model to capture functional-level constraints while preserving the original E(3)-equivariant generative architecture.

The overall framework consists of three tightly coupled components:

(1) **Pocket Pharmacophore Prior Network (P3N)** predicts a graph-level pharmacophore distribution directly from pocket representations, without requiring reference ligands. Specifically, the atomic features of the input pocket are first embedded and then aggregated by mean pooling to obtain a global pocket representation. This global representation is subsequently passed through a multi-layer perceptron (MLP) to predict the pharmacophore prior probability distribution (Eq. 5). During training, the target prior is defined as the mean of atom-level pharmacophore multi-hot labels over the ligand graph, and the module is optimized with a binary cross-entropy loss (Eq. 7).

The predicted graph-level prior is mapped into the denoiser condition space using a MLP projector and fused with the original conditional embedding by element-wise addition. At each reverse diffusion step t, the fused condition embedding is injected into both node-category and edge-category conditional embedding branches (Eq. 6). This MLP-based additive fusion operates entirely in the native conditioning space of the denoiser, preserving architectural consistency while providing persistent pharmacophore guidance throughout the full denoising trajectory.

(2) **Atom Pharmacophore Supervision Head (APSH)** provides an auxiliary atom-level supervision branch during training to enhance the learning of local functional representation. It takes the ligand node representations from the denoiser as input and predicts per-atom pharmacophore probabilities. Supervision is provided by RDKit-derived atom-level pharmacophore multi-hot labels, and optimization is performed using binary cross-entropy loss (Eq. 8). This auxiliary objective encourages the denoiser to learn node representations that are functionally discriminative at the local level.

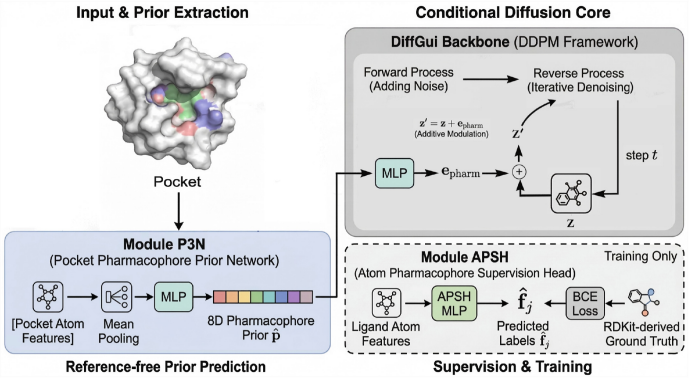


Fig. 2. Overall framework of PIPCDiff

Together, these components establish a dual-granularity learning framework that jointly models global pharmacophore priors and local atom-level functional features. All modules are jointly optimized during training. During inference, the model relies only on the predicted pocket-derived priors and the implicit conditioning mechanism to guide the diffusion process, without introducing additional supervisory branches or substantial computational overhead.

The core operations for prior prediction and conditional injection are defined as:

$$Z_{\text{pharm}} = \text{MLP}_{\text{target}}\big(\text{MeanPool}\big(W_p f_{\text{pocket}}\big)\big), \;\; \hat{p} = \sigma\big(Z_{\text{pharm}}\big) \tag{5}$$

$$z'_{\text{node}} = z_{\text{node}} + \text{MLP}_{\text{cond}}(\hat{p}) \tag{6}$$

where z_{node} represents the node embeddings in the EGNN layer, and the same symmetric additive formulation is applied to edge embeddings.

The dual-granularity supervision losses are defined as:

$$L_{\text{P3N}} = \text{BCEWithLogits}\big(Z_{\text{pharm}}, p\big) \tag{7}$$

$$L_{APSH} = \frac{1}{N}\sum_{j=1}^{N} \text{BCEWithLogits}\big(\text{MLP}_{ligand}\big(h_j\big), f_j\big) \tag{8}$$

3 Experiments

3.1 Dataset and Implementation Details

We use the PDBbind dataset for training, validation, and evaluation [14], following the same filtering and data-splitting protocol adopted by DiffGui to ensure fair comparison. PDBbind is a widely used benchmark for structure-based drug design and contains experimentally resolved 3D structures of protein-ligand complexes together with binding affinity annotations [15]. Starting from the PDBbind 2020 release and applying the above processing protocol, we obtain a high-quality subset comprising 17,327 complexes for training, 1,825 for validation, and 100 for testing.

To ensure both fairness and reproducibility, all experiments are conducted on top of the DiffGui backbone, and we strictly follow its original training protocol and hyperparameter configurations. The overall training objective is optimized using the same strategies as in the DiffGui baseline, with the optimizer, learning rate, batch size, and diffusion process settings kept identical unless otherwise specified. During the inference phase, we sample 100 ligand candidates for each testing protein pocket to comprehensively evaluate the generative performance. All training and sampling procedures are conducted on NVIDIA GPUs with CUDA support, and mixed-precision training is enabled to improve computational efficiency and reduce memory footprint.

3.2 Baselines

We compare PIPCDiff against three representative pocket-conditioned 3D molecular generation baselines. (1) TargetDiff is an early diffusion-based framework for target-aware 3D molecular generation [16]. It demonstrates the feasibility of generating ligand structures directly under protein pocket constraints. (2) DecompDiff is an end-to-end diffusion framework that incorporates decomposition priors and emphasizes molecular validity during generation. It supports multiple prior settings including reference prior, pocket prior, and optimal prior [17]. (3) DiffGui is a target-conditioned E(3)-equivariant diffusion model that jointly models atom and bond diffusion to capture their interdependence, while incorporating molecular property conditioning as well. Since our method is built upon DiffGui, this baseline is particularly important for isolating the effect of the proposed pharmacophore-aware implicit conditioning on pocket-guided molecular generation quality.

3.3 Evaluation Metrics

We adopt a diverse set of evaluation metrics to comprehensively assess the quality of generated molecules from the perspectives of chemical validity, 3D structural quality, functional matching, binding potential, and drug-likeness. (1) Atom Stability refers to the proportion of generated atoms with chemically valid valence states.(2) Molecular Stability refers to the proportion of generated molecules in which all constituent atoms are chemically stable. (3) 3D Structural Feasibility is evaluated using the PoseBusters suite to check geometric quality [18]. (4) Novelty is the ratio of generated molecules that do not appear in the training dataset. (5) Reference Similarity is computed by comparing the

Morgan fingerprints of the generated molecules with the corresponding native reference ligands. (6) Interaction Similarity represents the similarity of protein-ligand interaction fingerprints between the generated poses and the native complexes. (7) Pharmacophore Match Rate (PMR) measures the congruence between the generated 3D ligands and essential pharmacophoric features within the binding site. This metric is particularly relevant for assessing the effectiveness of the proposed pharmacophore-aware design. (8) Docking Score is estimated by AutoDock Vina [19], especially reporting three specific measurements. The Vina Score evaluates the native pose of the generated ligand, Vina Min reflects the binding affinity after local energy minimization, and Vina Dock denotes the best possible score derived from redocking. (9) QED stands for quantitative estimation of drug-likeness, which integrates multiple molecular property descriptors into a single score [20]. (10) SA denotes the synthetic accessibility that measures the relative difficulty of chemical synthesis for a given molecules [21].

4 Results

4.1 Qualitative Results

We conducted a series of experiments on the PDBbind dataset to evaluate the performance of PIPCDiff in structure-based molecular generation. Using multiple evaluation metrics, we compared PIPCDiff with representative baseline models in terms of chemical validity, binding quality, 3D geometric feasibility, and pharmacophore alignment.

As shown in Table 1, PIPCDiff achieves overall performance comparable to DiffGui. Although a slight decrease in molecular stability is observed, it is accompanied by improved pharmacophore precision. In particular, the improvement in atom-level stability suggests enhanced sensitivity to local chemical features. Moreover, the higher reference similarity and competitive interaction similarity demonstrate the advantage of PIPCDiff in preserving essential protein–ligand recognition patterns.

Table 1. Basic molecular metrics of generated molecules for the PDBbind dataset.

Model	Atom. Stab.(↑)	Mol. Stab.(↑)	Novel.(↑)	Sim. Ref.(↑)	Inter. Sim.(↑)
TargetDiff	**0.9443**	0.3751	**0.9932**	0.5060	0.7541
DecompDiff	0.8927	**0.4277**	0.9859	0.4511	**0.8844**
DiffGui	0.8815	0.2718	0.9888	0.5457	0.7755
PIPCDiff	0.9001	0.2686	0.9855	**0.5485**	0.7656

As shown in Table 2, PIPCDiff also demonstrates favorable performance in terms of binding affinity proxies and physicochemical drug-likeness. The generated molecules exhibit a better average Vina Score (-6.789) than those produced by DiffGui (−6.232), indicating that the incorporation of pharmacophore priors not only improves feature alignment but also translates into more favorable binding energetics in the generated native poses. In addition, PIPCDiff attains the highest QED score (0.604) and improved

synthetic accessibility (0.66), surpassing DiffGui. These results suggest that the proposed method is able to maintain chemical feasibility while generating molecules with richer functional constraints.

One limitation is observed in the Vina Dock evaluation after conformer resampling, where PIPCDiff (-8.780) does not outperform DiffGui (−9.015). A possible explanation is that the strong prior guidance in PIPCDiff favors higher-quality initial 3D poses, but may also restrict scaffold diversity or conformational flexibility. Consequently, the generated molecules may explore torsional space less broadly during subsequent redocking and therefore be less likely to reach deeper docking minima. Additionally, DecompDiff slightly surpass PIPCDiff in synthetic accessibility (0.67), which is consistent with its fragment-based decomposition strategy that naturally favors synthetically tractable molecular frameworks. Overall, PIPCDiff achieves a strong balance among structural quality, target-binding potential, and developability.

Table 2. Average molecular properties of reference ligands and generated molecules

Model	Vina Score(↓)	Vina Mini(↓)	Vina Dock(↓)	QED(↑)	SA(↑)
	Avg	Avg	Avg	Avg	Avg
Ref	-6.894	-7.679	-8.378	0.451	0.70
TargetDiff	-5.858	-6.814	-8.032	0.482	0.59
DecompDiff	-3.591	-4.714	-5.601	0.525	**0.67**
DiffGui	-6.232	-8.025	**-9.015**	0.572	0.59
PIPCDiff	**-6.789**	**-8.076**	-8.780	**0.604**	0.66

We further evaluated the 3D structural feasibility and geometric quality of the generated molecules using the PoseBusters suite. As shown in Table 3, PIPCDiff achieves consistent improvements over DiffGui across all evaluated geometric metrics, with the Overall score increasing from 48.6 to 61.9. In particular, PIPCDiff achieves the best performance in resolving internal steric clashes, reaching an Internal Clash score of 97.3 compared to 91.7 for DiffGui. This reduction suggests that the pharmacophore-aware conditioning mechanism effectively guides atoms toward more compatible spatial regions within the binding environment. In addition, PIPCDiff consistently outperforms DiffGui in Bond Length, Bond Angle, and Internal Energy, indicating improved geometric rationality and structural refinement. Among these metrics, the most pronounced improvement is observed in Internal Clash, suggesting that persistent pharmacophore-aware conditioning is particularly beneficial for reducing steric conflicts during molecular generation.

Despite these improvements, DecompDiff remains superior in the Overall score, Bond Angle, and Internal Energy metrics, likely due to its explicit fragment-based decomposition prior, which promotes low-energy local motifs and favorable geometric patterns. Nevertheless, PIPCDiff substantially improves the structural realism of DiffGui and shows particular strength in clash-free atom placement, highlighting the benefit of persistent pharmacophore-guided conditioning.

Table 3. Evaluation of 3D structural feasibility and geometric quality.

Model	Overall(↑)	Bond Length(↑)	Bond Angle(↑)	Int. Clash(↑)	Int. Energy(↑)
TargetDiff	58.5	**98.3**	75.9	92.1	70.2
DecompDiff	**68.0**	89.9	**86.9**	94.5	**80.8**
DiffGui	48.6	92.2	73.4	91.7	68.5
PIPCDiff	61.9	95.7	78.0	**97.3**	74.4

For qualitative visualization, we selected 6g1v (PDBid) from the PDBbind test set as a representative protein target. Each model generated 100 molecules, and for each method we visualized the ligands with the best docking scores among those passing PoseBusters filters (Fig. 3). Compared with the baseline models, PIPCDiff produces molecules with clearly defined chemical structures and more favorable pocket complementarity. This qualitative result further supports the idea that pharmacophore priors enable the model to better capture the spatial and functional correspondence between the pockets and ligands.

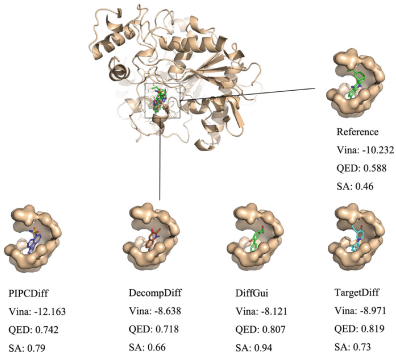


Fig. 3. Molecules generated for target from the PDBbind test set

In terms of pharmacophore alignment, PIPCDiff demonstrates improved global awareness of binding-site functional requirements. As shown in Table 4, PIPCDiff achieves an overall pharmacophore match rate (PMR) of 0.5019 under a 3 Å tolerance threshold, outperforming DiffGui (0.4745) and other representative baselines. Notably, PIPCDiff remains competitive even under the stricter 2 Å threshold, indicating that the improvement is not limited to coarse spatial matching. These results suggest that the pocket-adaptive priors predicted by the P3N module and injected through conditioning injection mechanism can provide effective functional guidance throughout the diffusion trajectory. In particular, the model shows reliable matching for major pharmacophoric categories such as hydrophobic centers and aromatic features. This improved alignment implies not only better structural consistency with the target pocket, but also stronger potential to form physicochemical interactions during binding.

Table 4. Pharmacophore matching performance (2 Å and 3 Å).

Metrics	TargetDiff		DecompDiff		PIPCDiff		DiffGui	
	2 Å	3 Å	2 Å	3 Å	2 Å	3 Å	2 Å	3 Å
Overall	0.3265	0.4743	0.2035	0.3320	**0.3504**	**0.5019**	0.3314	0.4745
Acceptor	**0.3162**	**0.4919**	0.2309	0.3865	0.2948	0.4488	0.2570	0.4024
Donor	**0.3431**	**0.5417**	0.2461	0.4110	0.2583	0.4341	0.2871	0.4571
Aromatic	0.1546	0.2171	0.0731	0.1160	**0.2624**	**0.3666**	0.1838	0.2616
Hydrophobe	0.5207	0.6897	0.2712	0.4148	**0.6184**	**0.7888**	0.6011	0.7652
Lumped Hydrophobe	0.1533	0.2104	0.0663	0.1028	**0.3730**	**0.5271**	0.3003	0.4428
NegIonizable	0.0828	0.1040	0.0483	0.0622	**0.1095**	**0.1375**	0.0640	0.0780
PosIonizable	0.1172	0.1784	0.0860	0.1383	0.0984	0.1567	**0.1362**	**0.1993**
ZnBinder	0.0869	0.1102	0.0550	0.0704	**0.1172**	**0.1425**	0.0628	0.0740

4.2 Ablation Study

To explore the contribution of each module, we perform three ablation experiments based on the baseline DiffGui. We first add only the APSH module for atom-level supervision, then adopt only the P3N module for pharmacophore distribution prediction and graph-level supervision, and finally compare single-step and multi-step injection strategies. The single-step strategy injects conditional signals only at the initial diffusion stage, while the multi-step strategy continuously introduces pharmacophore prior throughout the reverse diffusion process.

Table 5 shows that adding individual modules effectively improves pharmacophore matching accuracy, and integrating all modules achieves the best performance. The APSH module improves overall accuracy by 3.9%, mainly enhancing the recognition of Acceptor and Aromatic pharmacophores. The P3N module yields a 4.8% overall gain with similar performance improvements. Compared with the multi-step injection of PIPCDiff, the single-step strategy obtains slight gains in Acceptor and Donor matching but underperforms greatly on Aromatic and Hydrophobe pharmacophores. Combining all modules increases the overall accuracy by 5.7%. Meanwhile, the baseline model still maintains advantages in Donor pharmacophore recognition.

Table 5. Ablation study of PIPCDiff on modules and injection strategies.

Model	PMR(↑)	Acceptor(↑)	Donor(↑)	Aromatic(↑)	Hydrophobe(↑)
DiffGui	0.3314	0.2570	**0.2871**	0.1838	0.6011
DiffGui + APSH	0.3445	0.2908	0.2825	0.2028	0.5843

(*continued*)

Table 5. (*continued*)

Model	PMR(↑)	Acceptor(↑)	Donor(↑)	Aromatic(↑)	Hydrophobe(↑)
DiffGui + P3N	0.3476	0.2795	0.2441	0.2414	0.5719
PIPCDiff (single-step)	0.3478	**0.3041**	0.2612	0.2092	0.5562
PIPCDiff	**0.3504**	0.2948	0.2583	**0.2624**	**0.6184**

To assess the predictive performance of our P3N module, we measured the prediction quality using Mean Absolute Error (MAE) and compared the P3N module's predictions against a global mean baseline. Additionally, we conducted a progressive pocket-masking experiment. As shown in Table 6, the results indicate that as the masking ratio of the input pocket features decreases from 70% to 30%, the MAE gradually improves. When no masking is applied, the P3N module reduces the MAE from 0.105 to 0.088, resulting in a significant 16.2% improvement. These differences confirm the effectiveness of explicitly predicting priors from pocket features.

Table 6. MAE improvement of P3N module with pocket-masking experiment.

Condition	MAE(↓)	Relative MAE_global(↑)
Global Mean Baseline	0.105	0.0%
Full Pocket	**0.088**	**+16.2%**
Mask 30%	0.089	+15.2%
Mask 50%	0.091	+13.3%
Mask 70%	0.098	+6.7%

5 Conclusion

In this work, we propose PIPCDiff, a pharmacophore-aware diffusion framework for structure-based molecular generation. Built upon the DiffGui backbone, PIPCDiff introduces pocket-derived pharmacophore priors and integrates them into the denoising process through an implicit multi-step conditioning mechanism, while further strengthening local functional learning with atom-level pharmacophore supervision.

Experimental results on the PDBbind dataset demonstrate that PIPCDiff improves pharmacophore alignment, 3D structural feasibility, and binding-affinity-related proxy metrics, particularly Vina Score, compared with representative baseline methods. These findings indicate that pocket-derived pharmacophore priors can serve as an effective functional constraint for diffusion-based ligand generation and can enhance the structure–function consistency of generated molecules in structure-based drug design.

In future work, we plan to explore the integration of persistent pharmacophore-aware conditioning with fragment-based molecular assembly strategies. Combining

fragment-level geometric priors with pocket-adaptive pharmacophore guidance may further improve docking geometry, structural validity, and functional alignment during molecular generation. In addition, future studies may investigate diversity-aware regularization or adaptive conditioning strategies to better balance pharmacophore prior preservation and conformational flexibility. We will also evaluate the framework on broader cross-target and generalization-oriented benchmarks to further assess its applicability in practical drug discovery scenarios.

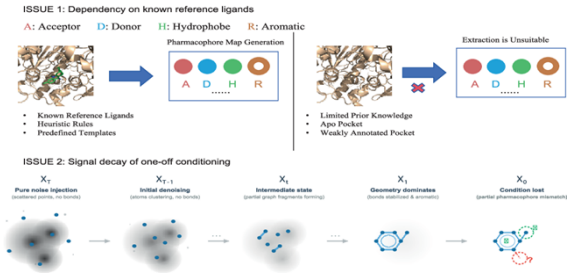
Funding. This study was supported by the Shenzhen Science and Technology Program (KJZD20240903095605007, JCYJ20250604145033042) and the Shenzhen Medical Research Fund (D250402003).

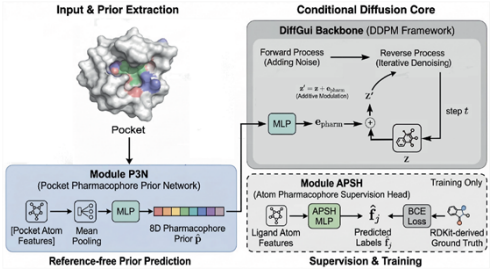
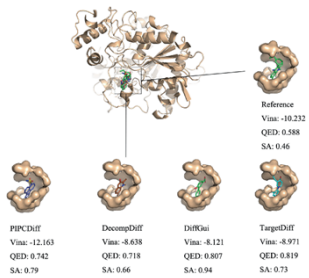
References

1. Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., Greyson, D.: The cost of drug development: a systematic review. Health Policy. **100**(1), 4–17 (2011)
2. Maxmen, A.: Busting the billion-dollar myth: how to slash the cost of drug development. Nature. **536**, 388–390 (2016)
3. Anderson, A.C.: The process of structure-based drug design. Chem. Biol. **10**(9), 787–797 (2003)
4. Ferreira, L.G., Dos Santos, R.N., Oliva, G., Andricopulo, A.D.: Molecular docking and structure-based drug design strategies. Molecules. **20**(7), 13384–13421 (2015)
5. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Adv. Neural Inf. Proces. Syst. **33**, 6840–6851 (2020)
6. Satorras, V.G., Hoogeboom, E., Welling, M.: E (n) equivariant graph neural networks. In: International Conference on Machine Learning, pp. 9323–9332. PMLR (2021)
7. Hoogeboom, E., Satorras, V.G., Vignac, C., Welling, M.: Equivariant diffusion for molecule generation in 3d. In: International Conference on Machine Learning, pp. 8867–8887. PMLR (2022)
8. Hu, Q., et al.: Target-aware 3d molecular generation based on guided equivariant diffusion. Nat. Commun. **16**(1), 7928 (2025)
9. Wolber, G., Langer, T.: Ligandscout: 3-d pharmacophores derived from proteinbound ligands and their use as virtual screening filters. J. Chem. Inf. Model. **45**(1), 160–169 (2005)
10. Yang, S.Y.: Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discov. Today. **15**(11–12), 444–450 (2010)
11. Li, B., Ran, T., Yang, Y., Xu, M., Chen, H.: Pharmacophore Model Guided 3d Molecular Generation through Diffusion Model (2025)
12. Alakhdar, A., Poczos, B., Washburn, N.: Diffusion models in de novo drug design. J. Chem. Inf. Model. **64**(19), 7238–7256 (2024)
13. Peng, X., Luo, S., Guan, J., Xie, Q., Peng, J., Ma, J.: Pocket2mol: efficient molecular sampling based on 3d protein pockets. In: International Conference on Machine Learning, pp. 17644–17655. PMLR (2022)
14. Wang, R., Fang, X., Lu, Y., Wang, S.: The pdbbind database: collection of binding affinities for protein- ligand complexes with known three-dimensional structures. J. Med. Chem. **47**(12), 2977–2980 (2004)
15. Wang, R., Fang, X., Lu, Y., Yang, C.Y., Wang, S.: The pdbbind database: methodologies and updates. J. Med. Chem. **48**(12), 4111–4119 (2005)

16. Guan, J., Qian, W.W., Peng, X., Su, Y., Peng, J., Ma, J.: 3d equivariant diffusion for target-aware molecule generation and affinity prediction. arXiv preprint (2023). https://arxiv.org/abs/2303.03543
17. Guan, J., et al.: Decompdiff: diffusion models with decomposed priors for structure-based drug design. arXiv preprint (2024). arXiv:2403.07902
18. Buttenschoen, M., Morris, G.M., Deane, C.M.: Posebusters: Ai-based docking methods fail to generate physically valid poses or generalise to novel sequences. Chem. Sci. **15**(9), 3130–3139 (2024)
19. Trott, O., Olson, A.J.: Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. **31**(2), 455–461 (2010)
20. Bickerton, G.R., Paolini, G.V., Besnard, J., Muresan, S., Hopkins, A.L.: Quantifying the chemical beauty of drugs. Nat. Chem. **4**(2), 90–98 (2012)
21. Ertl, P., Schuffenhauer, A.: Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J. Chem. **1**(1), 8 (2009)

Alternative Texts for Your Images, Please Check and Correct them if Required

Page no	Fig/Photo	Thumbnail	Alt-text Description
3	Fig1		The diagram explains two main issues in generating pharmacophore maps for drug design. The first issue shows that using known reference ligands, heuristic rules, and predefined templates allows creating a pharmacophore map with features labeled as acceptor (red), donor (blue), hydrophobe (green), and aromatic (brown). However, when prior knowledge is limited, such as with an empty binding pocket or weakly annotated pocket, this extraction method fails, indicated by a red cross. The second issue illustrates signal decay in one-off conditioning through five stages: starting with pure noise injection where points are scattered without bonds; initial denoising where atoms cluster but no bonds form; an intermediate state with partial graph fragments forming; a stage where geometry dominates with stabilized bonds and aromatic rings; and finally, condition loss where partial pharmacophore mismatches occur. This sequence highlights how the signal weakens over time, affecting the accuracy of pharmacophore generation. The image uses arrows to show progression and color-coded labels to identify chemical features, emphasizing challenges in drug design modeling when relying on limited data or single conditioning events.

Page no	Fig/Photo	Thumbnail	Alt-text Description
5	Fig2		A diagram explaining a process for predicting pharmacophore features in molecular pockets using a conditional diffusion model. On the left, a 3D molecular pocket is analyzed by Module P3N, which extracts pocket atom features, applies mean pooling, and uses a multilayer perceptron (MLP) to produce an eight-dimensional pharmacophore prior. This prior feeds into the DiffGui backbone, representing the core diffusion model with a forward process that adds noise and a reverse process that iteratively denoises, modulated by the pharmacophore prior. On the right, during training only, Module APSH uses ligand atom features to predict labels through an APSH MLP, compares predictions to ground truth from RDKit using binary cross-entropy loss, and refines the model. The flow shows how input features are transformed and combined to guide the diffusion process for accurate pharmacophore prediction, highlighting the integration of prior extraction, diffusion modeling, and supervised training.
9	Fig3		A protein structure is shown with a highlighted binding site containing a small molecule in green. Four smaller close-up views display different molecules docked in the same binding site, each with a unique color: blue, brown, green, and light blue. Each smaller view includes three scores labeled Vina, QED, and SA, which measure binding strength, drug-likeness, and surface area, respectively. The reference

Page no	Fig/Photo	Thumbnail	Alt-text Description
			molecule has a Vina score of minus 10.232, QED of 0.588, and SA of 0.46. The other molecules show varied scores, indicating differences in how well they fit and their properties. This comparison helps identify molecules with better binding and drug-like qualities for potential drug design.