

---

# Latent-Guided Equivariant Diffusion for Controlled Structure-Based De Novo Ligand Generation

---

Tuan Le<sup>\*1,2</sup> Julian Cremer<sup>\*1,3</sup> Djork-Arné Clevert<sup>1</sup> Kristof Schütt<sup>1</sup>

## Abstract

We propose PoLiGenX for *de novo* ligand design using latent-conditioned, target-aware equivariant diffusion. Our model leverages the conditioning of the generation process on reference molecules within a protein pocket to produce shape-similar *de novo* ligands that can be used for target-aware hit expansion and hit optimization. The results of our study showcase the efficacy of PoLiGenX in ligand design. Docking scores indicate that the generated ligands exhibit superior binding affinity compared to the reference molecule while preserving the shape. At the same time, our model maintains chemical diversity, ensuring the exploration of diverse chemical space. The evaluation of Lipinski’s rule of five suggests that the sampled molecules possess a higher drug-likeness than the reference data. This constitutes an important step towards the controlled generation of therapeutically relevant *de novo* ligands tailored to specific protein targets.

## 1. Introduction

In recent years, the intersection of artificial intelligence (AI) and drug discovery has witnessed remarkable strides, with the potential to revolutionize the traditional approaches to identifying novel therapeutic compounds. Among these innovations, AI-enabled structure-based drug discovery has emerged as a promising research avenue, in particular in form of equivariant target-aware diffusion models. By conditioning the diffusion process on the receptors of proteins, these models exhibit a remarkable capacity to generate *de novo* ligands with enhanced affinity (Peng et al., 2022; Guan et al., 2023; Schneuing et al., 2023; Le et al., 2024). Failing to consider the essential chemical properties for target bind-

ing can lead to a significant lack of specificity and result in ineffective drug candidates. Moreover, these candidates must exhibit favorable absorption, distribution, metabolism, excretion (ADME), and toxicity profiles. Designing ligands from scratch without addressing these critical properties may produce molecules with poor bioavailability or potential toxicity, thereby limiting their therapeutic potential. This challenge is further exacerbated by the often sparse and noisy data available for developing effective machine learning models. However, machine learning shows considerable promise during the hit expansion phase of drug discovery. This crucial stage involves enhancing and exploring the chemical space around promising hits already identified through high-throughput screening or other methods. In this study, we introduce PoLiGenX (**P**ocket-based **L**igand **G**enerator for hit **eX**ansion) that generates ligands *de novo* within a protein binding pocket. Unlike previous models, PoLiGenX starts with a seed molecule, such as a hit candidate or an initial scaffold, and iteratively refines and modifies it to improve its efficacy. We enhance the capabilities of the existing equivariant diffusion model, EQGAT-diff (Le et al., 2024), by incorporating a latent encoding as a condition. It is derived from an invariant graph neural network that is jointly trained to process 3D molecular inputs. The setup ensures that the newly generated ligands retain structural characteristics of the seed molecules while undergoing necessary chemical modifications and diversification. Our proposed approach adds a new level of control to the process of generating *de novo* ligands, aligning it more closely with the specific needs of targeted drug design, particularly during the hit expansion phase.

## 2. Related Work

Deep generative modeling in the life sciences has become a promising research area. Recent work by Xu et al. (2022); Jing et al. (2022) uses Denoising Diffusion Probabilistic Models (DDPMs) (Sohl-Dickstein et al., 2015; Ho et al., 2020; Kingma et al., 2021; Song et al., 2021) to predict the 3d coordinates of molecules with the help of 3d equivariant graph neural networks. In the *de novo* setting, another line of research focuses on directly generating the atomic coordinates and elements, using autoregressive models (Gebauer

---

<sup>\*</sup>Equal contribution <sup>1</sup>Pfizer Research & Development <sup>2</sup>Freie Universität Berlin <sup>3</sup>University Pompeu Fabra. Correspondence to: Tuan Le <tuan.le@pfizer.com>, Julian Cremer <julian.cremer@pfizer.com>.

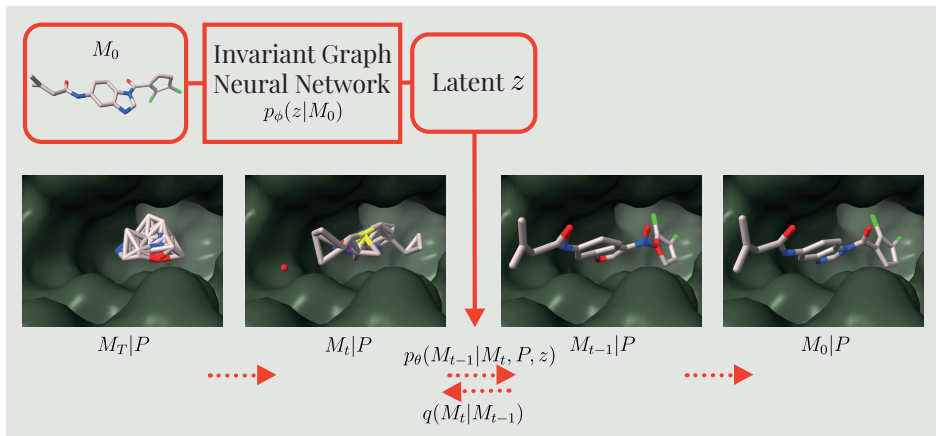


Figure 1. Graphical model of the proposed latent diffusion model. The encoded ligand  $z$  serves as input to the diffusion model  $p_\theta$  to steer the generation process of new ligands  $M_0$ .

et al., 2019; 2022; Luo & Ji, 2022). Hoogeboom et al. (2022) introduced E(3) equivariant diffusion model (EDM) for *de novo* molecule design that simultaneously learns atomic elements next to the coordinates while treating chemical elements as continuous variables to utilize the formalism of DDPM. Follow-up work leverages EDM and develop diffusion models for linker design (Igashov et al., 2022) and structure-based ligand modeling (Schneuing et al., 2023; Guan et al., 2023; Le et al., 2024). In the context of shape-conditioned molecule generation Adams & Coley (2023) (SQUID) and Chen et al. (2023) (ShapeMol) recently proposed to incorporate the shape of a seed molecule into the generation process. Both approaches use an equivariant surface encoding of a seed molecule, whereby SQUID uses variational auto-encoding on graphs and focuses more on fragment-based design. ShapeMol is an adaption of SQUID in 3d space leveraging an equivariant diffusion model. However, both works do not include a protein receptor condition. We propose to use a simple approach employing reference molecules in a latent representation, as outlined in more details below.

### 3. Methods

**Problem Formulation and Notation** We investigate the generation of molecular structures  $M$  in a *de novo* setting conditioned on a protein pocket  $P$ , i.e., building a generative model  $p_\theta(M|P)$ . For this, we use the EQGAT-diff framework proposed by Le et al. (2024). In this setup, a noisy ligand  $M_t = (X_t, H_t, E_t)$  — representing perturbed atomic coordinates, element types, and bond features — is used, and the diffusion model  $p_\theta$  predicts the uncorrupted data modalities  $(\hat{X}_0, \hat{H}_0, \hat{E}_0)$ , because the distribution  $M_{t-1}|M_t$  depends on both  $M_t$  and  $\hat{M}_0$ . Specifically, for continuous coordinates, the reverse distribution adheres to a multivariate Gaussian model, while for discrete-valued modalities,

it follows a categorical distribution. We refer to Le et al. (2024) for further details.

While models like EQGAT-diff, TargetDiff or DiffSBDD generate ligands in context of a protein pocket, they do not constraint the generated ligands to preserve properties like shape or chemical similarity during training. In contrast, we include a latent variable  $z \in \mathbb{R}^K$  that relates to the input molecule  $\hat{M}_0$ . The latent  $z$  may serve as a shape conditioning that also comprises chemical information like the atom composition of the molecule  $\hat{M}_0$ . The graph encoder  $q_\phi: \mathcal{X}^M \rightarrow \mathbb{R}^K$  is invariant to permutation, rotation and translation of atoms (Winter et al., 2022; Le et al., 2022).

Following Adams & Coley (2023), chemical similarity of two molecules is measured as the Tanimoto similarity of ECFP4 fingerprints (2048 bits) computed by RDKit, whereby shape similarity is defined by Gaussian descriptions of molecular shape in form of atom-centered Gaussians and calculated by the volume overlaps between them as in Adams & Coley (2023).

**PoLiGenX** To model the dependence on variable  $z$ , we include a variational distribution  $q_\phi(z|M_0)$  similar to Luo & Hu (2021); Zeng et al. (2022) and obtain the ELBO

$$\begin{aligned}
 p_\theta(M_0|P) &= \mathbb{E}_{q(M_{1:T}|M_0)q_\phi(z|M_0)} \left[ \frac{p_\theta(M_0, M_{1:T}, z|P)}{q(M_{1:T}|M_0)q_\phi(z|M_0)} \right] \\
 &\geq \mathbb{E}_{q(M_{1:T}|M_0)q_\phi(z|M_0)} \left[ \log \frac{p_\theta(M_0, M_{1:T}, z|P)}{q(M_{1:T}|M_0)q_\phi(z|M_0)} \right] \\
 &= \mathbb{E}_{q(M_1|M_0)q_\phi(z|M_0)} [\log p_\theta(M_0|M_1, P, z)] \\
 &+ \mathbb{E}_{q(M_T|M_0)q_\phi(z|M_0)} \left[ \log \frac{p(M_T|z)}{q(M_T|M_0)} \right] \\
 &- D_{KL}(q_\phi(z|M_0)||p(z)) - \sum_{t=2}^T \mathbb{E}_{q(M_t|M_0)q_\phi(z|M_0)} [L_{t-1}],
 \end{aligned} \tag{1}$$

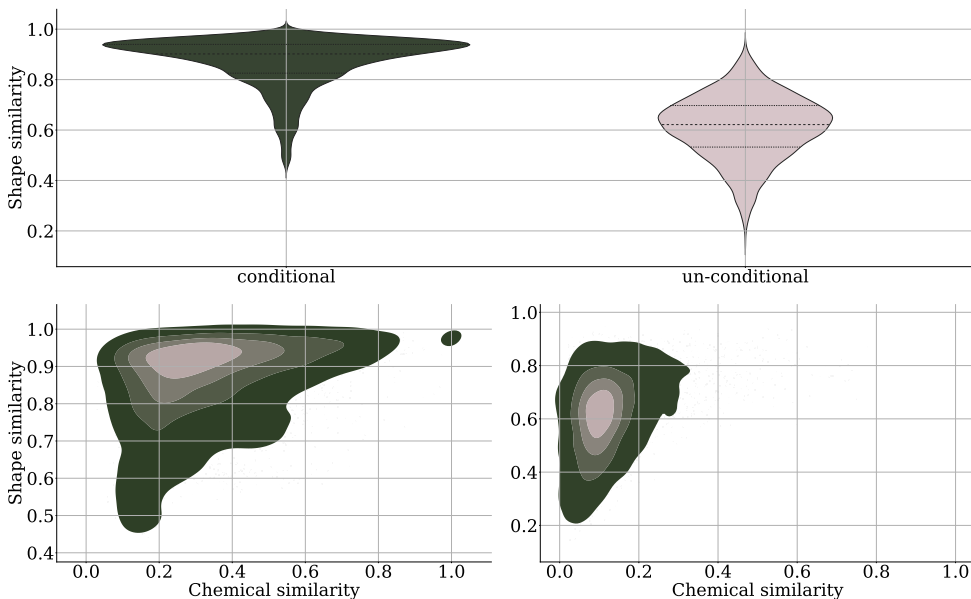


Figure 2. **Top:** Violin plot of the Tanimoto shape similarity evaluated across all test targets of the CrossDocked dataset. PoLiGenX (left) is compared to EQGAT-diff (right). In the conditional setting the model generates significantly more shape-similar molecules. **Bottom:** Heatmap histogram comparing PoLiGenX (left) with EQGAT-diff with respect to Tanimoto shape and chemical similarity on the CrossDocked test set. The brighter the color the higher the molecule count.

where the diffusion loss  $L_{t-1}$  is per timestep and defined as  $L_{t-1} = D_{KL}(q(M_{t-1}|M_t, M_0)||p_\theta(M_{t-1}|M_t, P, z))$ .

We extend the diffusion model by a conditioning on  $z$  and train  $p_\theta(M|P, z)$  to minimize the KL divergence to the tractable reverse distribution, which is achieved when predicting the original data points  $\hat{M}_0$  (Ho et al., 2020; Austin et al., 2021; Le et al., 2024). Similar to prior works, we optimize the diffusion  $L_{t-1}$  by drawing steps per minibatch instead of the entire trajectory. We adopt a Gaussian prior for the latent distribution, i.e.,  $p(z) \sim N(0, I)$  and enforce a smooth latent space by choosing the maximum mean discrepancy (MMD) loss (Tolstikhin et al., 2018) over the KL divergence. The prior distribution for the ambient data space, i.e.,  $M_T$  is a 0-CoM Gaussian for coordinates and empirical categorical distribution for discrete data types from the training set as discussed in Le et al. (2024). During training, we sample a batch of pocket-ligand pairs and a step  $t \in \{1, \dots, 500\}$ . Next, we encode the ligands  $M_0$  into latents  $z$ , apply the noise process to the ligands to obtain  $M_t$  and minimize the diffusion loss while providing  $z$  as an additional input via adaptive layer normalization (Huang & Belongie, 2017) next to the protein pocket  $P$ . We refer to the supplementary material for further details including the derivation of the ELBO.

## 4. Results

We train PoLiGenX using the CrossDocked2020 (Francoeur et al., 2020) dataset, following the same dataset splits as found in previous research (Luo et al., 2021; Peng et al., 2022; Guan et al., 2023; Schneuing et al., 2023; Le et al., 2024). Unlike other models, PoLiGenX incorporates not only the protein pocket as a condition for generating novel ligands but also utilizes a latent embedding of a ligand from the dataset as an initial condition. This distinctive approach positions PoLiGenX differently from the mentioned models — it is specifically designed to perform tasks akin to hit expansion by enhancing specificity, chemical diversity, and binding affinity, rather than operating solely as a target-aware, but unconditional *de novo* model. In the following, we evaluate if PoLiGenX effectively maintains the structural shape of the seed molecule while promoting chemical diversity.

Fig. 2 (top) shows the evaluation of the mean shape similarity on the CrossDocked test set for both PoLiGenX (conditional) and EQGAT-diff (unconditional). The test set comprises 100 ligand-pocket complexes for which 100 ligands each were sampled and the Tanimoto shape similarity measured against the reference ligands. PoLiGenX exhibits significantly higher shape similarities across complexes. However, we aim to preserve the shape between reference and sample *without* sacrificing chemical diversity to ensure an efficient exploration of chemical space. Fig. 2 (bottom) shows the distribution of shape similarity against

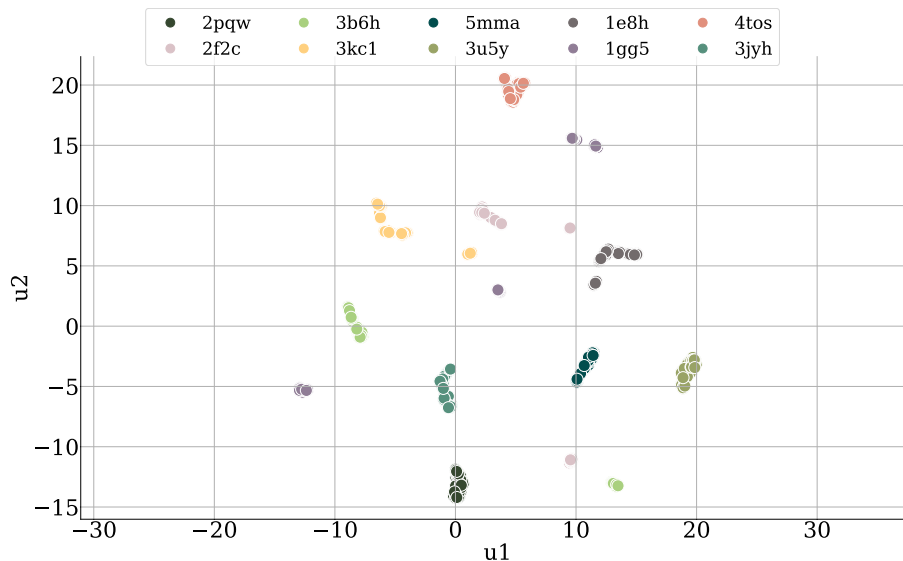


Figure 3. UMAP plot showing the 2d projections of the latent embeddings of 100 sampled ligands per target for ten randomly sampled test set targets.

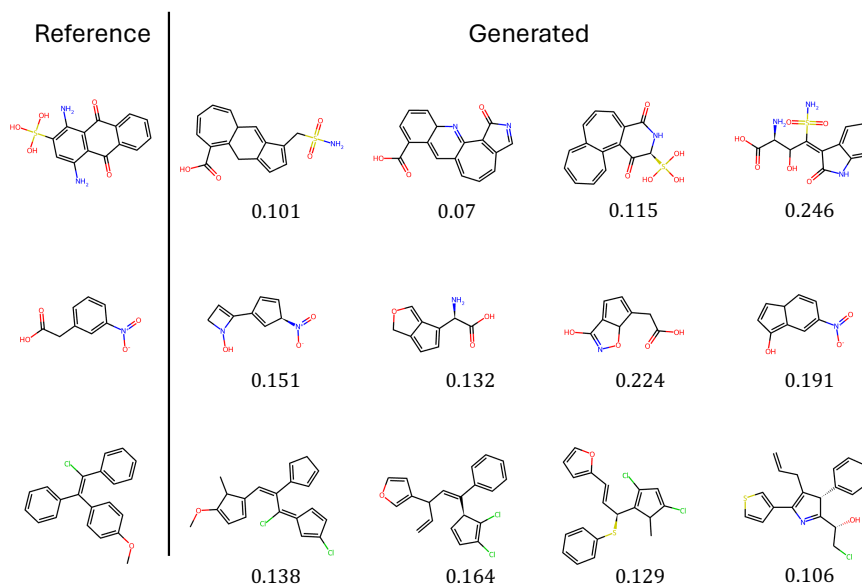


Figure 4. Reference molecules extracted from the CrossDocked test split (left) and four generated molecules sampled randomly with PoLiGenX. Below each generated ligand, we also show the chemical similarity to the reference ligand.

chemical similarity for conditional and unconditional sampling. We observe a mean shape similarity of 0.64 and 0.12 chemical similarity for EQGAT-diff. In contrast, PoLiGenX exhibits a significant increase in shape similarity with mean value of 0.87, but also generates a reasonably high diversity in samples with mean chemical similarity of 0.33.

To evaluate the expressiveness of the learned latent embeddings, Fig. 3 visualizes the UMAP projections of the latent embeddings. We sampled 100 ligands per receptor for ten

randomly selected targets of the CrossDocked test set. The resulting UMAP projections reveal that the latent embeddings effectively separate the ligands into distinct clusters specific to each target. This observation suggests that our latent model successfully captures the context of ligands in relation to their respective protein receptors.

Next, we compare molecules sampled conditionally from our model, PoLiGenX, with the reference test data, focusing on docking scores and chemical properties. As previously

Table 1. Docking performance on the CrossDocked test set and ligands generated using PoLiGenX. QuickVina2 is employed for docking. We report mean values across all targets with standard deviations given as subscripts. Drug-likeness is measured via RDKit’s QED value. Further, molecules are evaluated in terms of the octanol–water partition coefficient (logP), the molecular weight (MolWt) and the number of hydrogen acceptors and donors. Following Lipinski’s rule of five, we report the percentage of molecules that obey the respective rule. The last column gives the average of molecules fulfilling all rules.

Data	QVina2 (All) ↓	QVina2 (Top-10%) ↓	QED ↑	logP ↑	MolWt ↑	H-acceptors ↑	H-donors ↑	Lipinski ↑
CrossDocked test set	$-6.85 \pm 2.33$	-	$0.47 \pm 0.20$	0.79	0.85	0.84	0.8	$3.35 \pm 1.14$
PoLiGenX	<b><math>-7.21 \pm 2.22</math></b>	<b><math>-8.04 \pm 2.44</math></b>	<b><math>0.59 \pm 0.20</math></b>	<b>0.91</b>	<b>0.87</b>	<b>0.85</b>	<b>0.91</b>	<b><math>3.57 \pm 0.93</math></b>

outlined, the purpose of PoLiGenX is significantly different to recent *de novo* models, such as EQGAT-diff, hence we omit a comparison. Tab. 1 summarizes the results. We observe improved docking scores for generated samples compared to the CrossDocked test data, in particular within the top 10% of each target. Here, we reach a docking score of  $-8.04 \pm 2.44$  compared to  $-6.85 \pm 2.33$  for the test data. At the same time, the generated ligands per target show improvement in RDKit’s drug-likeness score (QED) and adherence to Lipinski’s Rule of Five. These are chemical features recognized from a medicinal chemistry perspective as guidelines to identify compounds likely to possess favorable bioavailability. Specifically, the octanol-water partition coefficient (logP) should be less than 5, molecular weight (MolWt) should be less than 500 Daltons, hydrogen bond acceptors (H-acceptors) less than 10 and hydrogen bond donors (H-donors) should be less than 5.

Fig. 4 depicts three randomly chosen test set ligands with four conditionally sampled and randomly selected ligands each. Judging by visual inspection, the topology is well preserved. We note that chemical similarity, especially based on fingerprints can change drastically if some chemical elements are interchanged. As shown in the bottom panel in Figure 2, PoLiGenX achieves a mean chemical similarity of around 0.33 while preserving shape similarity of 0.87 compared to the unconditional case with 0.12 and 0.64 for chemical and shape similarity, respectively.

The controllable generation of PoLiGenX can be further regulated by including a control parameter  $\lambda \in (0, 1]$  that scales the latent  $z$  when going into the diffusion model. That is, for small  $\lambda$  values approaching 0, PoLiGenX does not include any latent information and collapses to the unconditional EQGAT-diff and only leverages the pocket information as context. With  $\lambda$  interpolating between e.g. (0.5, 1.0), we observe that the mean chemical similarity for generated ligands with respect to the references also increases as depicted in Figure 5. We detail the influence of the latent variable  $z$  in combination with the scale parameter  $\lambda$  in the supplementary materials.

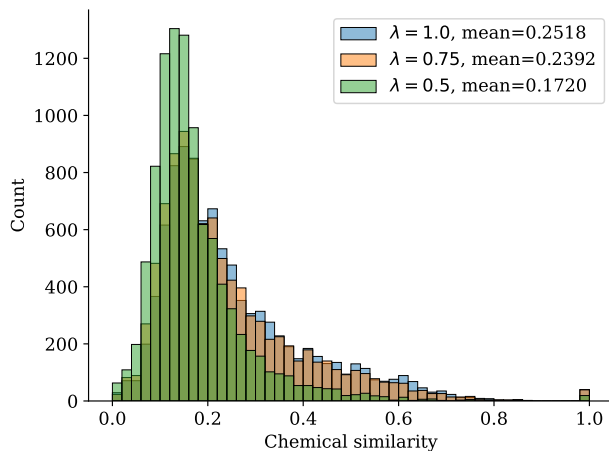


Figure 5. Density plot for chemical similarity of generated ligands from PoLiGenX with varying  $\lambda$  control parameter. With increasing  $\lambda$ , the latent  $z$  of reference/seed ligand  $M_0$  is preserved such that generated ligands exhibit higher chemical similarity to  $M_0$ .

## 5. Conclusions

We have developed PoLiGenX for controlled *de novo* ligand generation within a protein binding pocket. By incorporating a latent encoding from a seed molecule into the diffusion model, we ensure that the generated ligands preserve shape and also adhere to the structural constraints of the target protein binding site. The effectiveness of PoLiGenX is evidenced by improved docking scores compared to reference ligands. Additionally, the generated ligands conform to Lipinski’s Rule of Five, demonstrating their drug-likeness. Importantly, the model maintains chemical diversity, which is essential for exploring a broad range of chemical space and discovering novel therapeutic candidates. This integration of shape preservation, target specificity, and chemical diversity provides a powerful approach for the targeted generation of drug candidates, particularly useful in the hit expansion phase of drug discovery campaigns.

## References

- Adams, K. and Coley, C. W. Equivariant shape-conditioned generation of 3d molecules for ligand-based drug design. In *The Eleventh International Conference on Learning Representations*, 2023. URL <https://openreview.net/forum?id=4MbGnp4iPQ>.
- Austin, J., Johnson, D. D., Ho, J., Tarlow, D., and van den Berg, R. Structured denoising diffusion models in discrete state-spaces. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021. URL <https://openreview.net/forum?id=h7-XixPCAL>.
- Chen, Z., Peng, B., Parthasarathy, S., and Ning, X. Shape-conditioned 3d molecule generation via equivariant diffusion models, 2023. URL <https://arxiv.org/abs/2308.11890>.
- Francoeur, P. G., Masuda, T., Sunseri, J., Jia, A., Iovanisci, R. B., Snyder, I., and Koes, D. R. Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *Journal of Chemical Information and Modeling*, 60(9):4200–4215, Sep 2020. ISSN 1549-9596. doi: 10.1021/acs.jcim.0c00411. URL <https://doi.org/10.1021/acs.jcim.0c00411>.
- Gebauer, N., Gastegger, M., and Schütt, K. Symmetry-adapted generation of 3d point sets for the targeted discovery of molecules. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL [https://proceedings.neurips.cc/paper\\_files/paper/2019/file/a4d8e2a7e0d0c102339f97716d2fdfb6-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2019/file/a4d8e2a7e0d0c102339f97716d2fdfb6-Paper.pdf).
- Gebauer, N. W. A., Gastegger, M., Hessmann, S. S. P., Müller, K.-R., and Schütt, K. T. Inverse design of 3d molecular structures with conditional generative neural networks. *Nature Communications*, 13(1), feb 2022. doi: 10.1038/s41467-022-28526-y. URL <https://doi.org/10.1038/s41467-022-28526-y>.
- Guan, J., Qian, W. W., Peng, X., Su, Y., Peng, J., and Ma, J. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. In *The Eleventh International Conference on Learning Representations*, 2023. URL <https://openreview.net/forum?id=kJqXEPXMsE0>.
- Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020. URL [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf).
- Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3D. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S. (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8867–8887. PMLR, 17–23 Jul 2022. URL <https://proceedings.mlr.press/v162/hoogeboom22a.html>.
- Huang, X. and Belongie, S. Arbitrary style transfer in real-time with adaptive instance normalization. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 1510–1519, 2017. doi: 10.1109/ICCV.2017.167.
- Igashov, I., Stärk, H., Vignac, C., Satorras, V. G., Frossard, P., Welling, M., Bronstein, M. M., and Correia, B. E. Equivariant 3d-conditional diffusion models for molecular linker design. *ArXiv*, abs/2210.05274, 2022. URL <https://arxiv.org/abs/2210.05274>.
- Jing, B., Corso, G., Chang, J., Barzilay, R., and Jaakkola, T. S. Torsional diffusion for molecular conformer generation. In Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. (eds.), *Advances in Neural Information Processing Systems*, 2022. URL [https://openreview.net/forum?id=w6fj2r62r\\_H](https://openreview.net/forum?id=w6fj2r62r_H).
- Kingma, D. P., Salimans, T., Poole, B., and Ho, J. On density estimation with diffusion models. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021. URL <https://openreview.net/forum?id=2LdBqxclYv>.
- Le, T., Noe, F., and Clevert, D.-A. Representation learning on biomolecular structures using equivariant graph attention. In *The First Learning on Graphs Conference*, 2022. URL <https://openreview.net/forum?id=kv4xUo5Pu6>.
- Le, T., Cremer, J., Noé, F., Clevert, D.-A., and Schütt, K. Navigating the design space of equivariant diffusion-based generative models for de novo 3d molecule generation. In *The Twelfth International Conference on Learning Representations*, 2024. URL <https://openreview.net/forum?id=kzGuiRXZrQ>.
- Luo, S. and Hu, W. Diffusion probabilistic models for 3d point cloud generation. In *Proceedings of*

- the *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2837–2845, June 2021. URL [https://openaccess.thecvf.com/content/CVPR2021/html/Luo\\_Diffusion\\_Probabilistic\\_Models\\_for\\_3D\\_Point\\_Cloud\\_Generation\\_CVPR\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021/html/Luo_Diffusion_Probabilistic_Models_for_3D_Point_Cloud_Generation_CVPR_2021_paper.html).
- Luo, S., Guan, J., Ma, J., and Peng, J. A 3d generative model for structure-based drug design. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 6229–6239. Curran Associates, Inc., 2021. URL [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/314450613369e0ee72d0da7f6fee773c-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/314450613369e0ee72d0da7f6fee773c-Paper.pdf).
- Luo, Y. and Ji, S. An autoregressive flow model for 3d molecular geometry generation from scratch. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=C03Ajc-NS5W>.
- Peng, X., Luo, S., Guan, J., Xie, Q., Peng, J., and Ma, J. Pocket2Mol: Efficient molecular sampling based on 3D protein pockets. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S. (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 17644–17655. PMLR, 17–23 Jul 2022. URL <https://proceedings.mlr.press/v162/peng22b.html>.
- Schneuing, A., Du, Y., Harris, C., Jamasb, A., Igashov, I., Du, W., Blundell, T., Lió, P., Gomes, C., Welling, M., Bronstein, M., and Correia, B. Structure-based drug design with equivariant diffusion models, 2023. URL <https://arxiv.org/abs/2210.13695>.
- Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., and Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In Bach, F. and Blei, D. (eds.), *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2256–2265, Lille, France, 07–09 Jul 2015. PMLR. URL <https://proceedings.mlr.press/v37/sohl-dickstein15.html>.
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL <https://openreview.net/forum?id=PxtIG12RRHS>.
- Tolstikhin, I., Bousquet, O., Gelly, S., and Schoelkopf, B. Wasserstein auto-encoders. In *International Conference on Learning Representations*, 2018. URL <https://openreview.net/forum?id=HkL7n1-0b>.
- Winter, R., Bertolini, M., Le, T., Noe, F., and Clevert, D.-A. Unsupervised learning of group invariant and equivariant representations. In Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. (eds.), *Advances in Neural Information Processing Systems*, 2022. URL <https://openreview.net/forum?id=471pv23LDPr>.
- Xu, M., Yu, L., Song, Y., Shi, C., Ermon, S., and Tang, J. Geodiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=PzcvxEMzvQC>.
- Zeng, X., Vahdat, A., Williams, F., Gojic, Z., Litany, O., Fidler, S., and Kreis, K. LION: Latent point diffusion models for 3d shape generation. In Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. (eds.), *Advances in Neural Information Processing Systems*, 2022. URL <https://openreview.net/forum?id=tHK5ntjp-5K>.