

BindGPT: A Scalable Framework for 3D Molecular Design via Language Modeling and Reinforcement Learning

Artem Zholus^{1,3, 4,5*}, Maksim Kuznetsov¹, Roman Schutski², Rim Shayakhmetov¹, Daniil Polykovskiy¹, Sarath Chandar^{3,4, 5, 6} and Alex Zhavoronkov²

¹ Insilico Medicine Canada Inc.

² Insilico Medicine AI Limited

³ Mila – Quebec AI Institute

⁴ Polytechnique Montréal

⁵ Chandar Research Lab

⁶ CIFAR AI Chair

Abstract

Generating novel active molecules for a given protein is an extremely challenging task for generative models that requires an understanding of the complex physical interactions between the molecule and its environment. This paper presents a novel generative model, BindGPT, which uses a conceptually simple but powerful approach to create 3D molecules within the protein’s binding site. Our model produces molecular graphs and conformations jointly, eliminating the need for an extra graph reconstruction step. We pre-train BindGPT on a large-scale dataset and fine-tune it with reinforcement learning using scores from external simulation software. We demonstrate how a single pre-trained language model can serve at the same time as a 3D molecular generative model, a conformer generator conditioned on the molecular graph, and a pocket-conditioned 3D molecule generator. Notably, the model does not make any representational equivariance assumptions about the domain of generation. We show how such a simple conceptual approach combined with pre-training and scaling can perform on par or better than the current best-specialized diffusion models, language models, and graph neural networks while being two orders of magnitude cheaper to sample.

Website — <https://bindgpt.github.io/>

1 Introduction

The drug discovery landscape presents immense challenges and risks, demanding substantial investments of time and resources to design, test, and deliver new medicines to the market. Within this context, Computer-Aided Drug Design (CADD) (Yu and MacKerell 2017) stands as a pivotal methodology, harnessing software screenings and physical simulations to facilitate a more efficient exploration of the vast space of drug-like molecules, estimated to be around 10^{60} in size (Polishchuk, Madzhidov, and Varnek 2013). Deep learning advancements have revolutionized this exploration by leveraging neural generative models trained on extensive compound datasets. Notably, the textual representation of

molecular structures using SMILES (Weininger 1988) has enabled the utilization of Language Models for the generation of novel, drug-like molecular compounds (Bagal et al. 2022).

Recent research has demonstrated the capability of deep generative models to generate novel molecular compounds directly in 3D, with the flexibility to incorporate protein pocket and ligand subfragment conditions. Among these, diffusion models such as EDM (Hoogeboom et al. 2022) and DiffDock (Corso et al. 2023) initiate the generation process with an arbitrary spatial distribution of atoms and progressively refine their positions to yield physically viable molecular structures. Meanwhile, autoregressive models like Pocket2Mol (Peng et al. 2022) sequentially predict the type and location of each successive atom, building upon the existing molecular framework. Additionally, work by Flam-Shepherd and Aspuru-Guzik (2023) has highlighted the proficiency of language models in handling spatial representations of molecular and protein structures through formats like XYZ, CIF, and PDB. However, it’s noteworthy that most spatial molecular generators focus exclusively on atom types and locations. They depend on supplementary tools, such as OpenBabel (O’Boyle et al. 2011), for the critical task of bond reconstruction. This reliance can introduce vulnerabilities, as the precision required for atom placement means that minor positional adjustments can significantly alter reconstructed molecular bonds or even make the molecular graph disconnected.

In this work, we introduce a novel framework that applies language modeling to the domain of 3D molecular data represented by textual tokens. This entirely data-driven approach, devoid of any inductive biases at both the model and representation levels, capitalizes on the established GPT paradigm, integrating cutting-edge techniques to enhance the scalability of model training and inference. By adopting the language model pre-training paradigm, our framework showcases the ability to foster a powerful causal language model adept at navigating the complex space of 3D molecules. This proficiency is demonstrated through successful applications in downstream tasks, including learning the distribution of molecules without conditions and conformations for a molecular graph, and, crucially, by generating specific molecules

*Correspondence to: artem.zholus@mila.quebec, r.schutski@insilicomedicine.com
Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

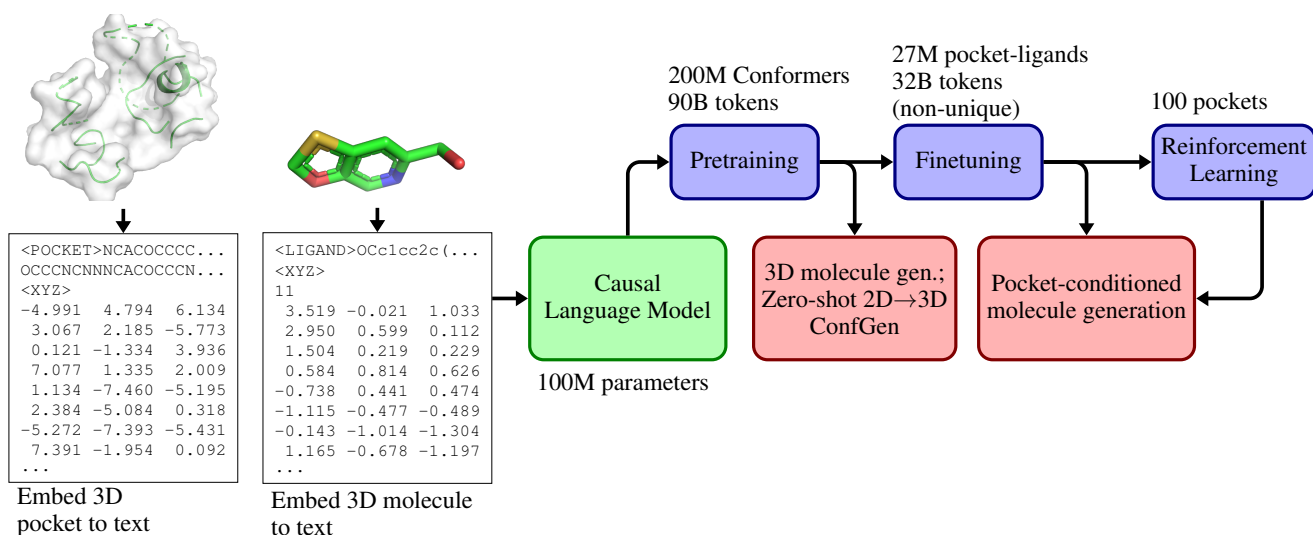


Figure 1: The pipeline of our 3D molecule pre-training and fine-tuning paradigm. During pre-training the model is trained on a mix of molecules and pockets in isolation. But at fine-tuning, we simply concatenate the pocket text representation and the molecule text representation for each pocket-ligand pair.

and conformations with targeted binding affinity to a protein. Our main contributions are the following:

- We introduce BindGPT, a Language Model for handling spatial molecular structures in text format. It uses structural SMILES and spatial XYZ formats to describe molecular graphs and atom locations, eliminating the dependency on external software for graph reconstruction.
- We propose a scalable pre-training and fine-tuning method for drug discovery in 3D that covers several 3D molecular generation tasks in a single paradigm.
- We show how BindGPT can create accurate and realistic 3D molecular structures both zero-shot and after fine-tuning, with the option to include molecular graphs or protein pocket descriptions as prompts. The method offers comparable generation quality to leading approaches with a speedup of up to 100x.
- Finally, we demonstrate the effectiveness of the Reinforcement Learning framework to fine-tune BindGPT with external feedback from docking software. We show that the resulting model can find structures with high binding scores for any given protein as a result of the RL fine-tuning.

2 Background

Molecule Generation Small drug-like molecules can be represented as 2D or 3D graphs with node and edge attributes. However, one of the most popular molecular representations in the machine learning community is SMILES (Weininger 1988), which can be seen as a compressed textual encoding of the Depth-First-Search applied to the molecular graph. Its simplicity and expressive power made it work very well with language models - even a simple LSTM (Hochreiter and Schmidhuber 1997) model can outperform graph-neural networks for the molecule generation task (Flam-Shepherd, Zhu, and Aspuru-Guzik 2022).

The biological function of small molecules arises through their binding to specific protein pockets. The spatial structure of the protein pocket is an essential domain knowledge to increase the efficiency of molecular generation in drug design tasks. With the increase of molecular structure datasets sizes (Francoeur et al. 2020; Hu et al. 2005) a plethora of pocket-conditioned generators emerged (Peng et al. 2022; Luo et al. 2021; Lin et al. 2022; Corso et al. 2023). The challenge with pocket-conditioned molecular generation arises from a relatively small size of existing 3D binding poses datasets, which motivated a heavy use of specialized architectures, like SE(3) equivariant neural networks (Hoogeboom et al. 2022).

Molecular Generative Models in 3D. Apart from Language Models, there are other types of models for molecule generation, including 3D-aware approaches. The first group of works uses Diffusion Models, which employ the denoising diffusion process (Ho, Jain, and Abbeel 2020) to learn to recover the data from noise. The second group of works relies on Graph Neural Networks (GNNs) to autoregressively build 2D or 3D molecular graphs. These approaches can be combined as GNNs can serve as efficient backbones for the diffusion process once they are node-equivariant (Niu et al. 2020) (to generate 2D graphs) or SE(3)-equivariant (Peng et al. 2023a) (to generate 3D graphs). SBDD (Luo et al. 2021) model uses autoregressive graph generation for pocket-conditioned molecule generation. TargetDiff (Schneuing et al. 2023) generalizes this model to use a diffusion-GNN for the same task. Pocket2Mol (Peng et al. 2022) uses an informed autoregressive sampling mechanism for efficient pocket-conditioned molecule generation. Another batch of works uses the aforementioned methods for unconditional molecule generation. EDM (Hoogeboom et al. 2022) proposes an E(3) equivariant diffusion model for molecule generation. MolDiff (Peng et al. 2023a) is a diffusion model that addresses the inconsistency problem between generated atoms and bonds connecting them.

Language Models for Drug Discovery. The language models show outstanding results in the drug discovery domain. The molecular structures can be easily represented in textual formats like SMILES (Flam-Shepherd, Zhu, and Aspuru-Guzik 2022), enabling the effective training of well-known language model architectures on large datasets of chemical entities. Recent studies reveal the potential of applying language models to address various challenges in drug discovery. For instance, LigGPT (Bagal et al. 2022) leverages the GPT (Radford et al. 2018) architecture to generate molecular structures given the conditions of molecular descriptors. MoLFormer (Katharopoulos et al. 2020) incorporates a billion-size chemical entities database to perform large-scale pre-train and further fine-tune to predict molecular properties. BARTSmiles (Chilingaryan et al. 2022) is developed atop of BART (Lewis et al. 2019) architecture, training a meaningful chemical representation and refining it for tasks such as chemical property prediction, chemical reaction prediction, and retrosynthesis.

However, the challenge of 3D molecule generation has received limited attention in the language model approach. Three notable studies in this area include the XYZ-transformer (Flam-Shepherd and Aspuru-Guzik 2023), Uni-Mol (Zhou et al. 2023), and Lingo3DMol (Feng et al. 2024). The XYZ-transformer leverages GPT to generate atom-wise descriptions of molecular and protein structures. Uni-Mol modifies BERT for large-scale pre-training on a large 3D structures dataset. In particular, Uni-Mol formulates molecular tasks as coordinates-to-coordinates mapping given molecular graph. **To the best of our knowledge, we propose the first approach that applies the modern decoder-only language modeling paradigm that uses pertaining, fine-tuning, and reinforcement learning to the 3D small molecule generation with several downstream applications.**¹

3 Method

The key idea of our method is utilizing an autoregressive token generation model, influenced by GPT-based models, to solve several 3D small molecule generation tasks in one simple yet flexible paradigm. The main principle in our approach is to formulate several 3D molecular design tasks as prompted generation of text. To achieve that, we position the tokens of a condition before the tokens of the object to generate. For instance, a prompt can be the protein pocket for the pocket-conditioned generation task or the 2D molecular structure for the conformation generation task.

3.1 Architecture and Text Format

In our work, we follow the decoder-only paradigm in Language Models and use the GPT-NeoX² architecture (Black

¹All of the previous generative approaches to 3D small molecules focus each of these tasks (molecule generation, conformer generation, pocket-conditioned molecule generation) alone, while we show how they can be efficiently solved within one pre-training paradigm.

²Despite other architectures existing today, our main focus was building a general framework of working with 3D molecules that can work with any Language Model architecture.

et al. 2022) that utilizes rotary position embeddings (Su et al. 2021). This technique allows for length generalization, which is required since the sequence lengths may vary significantly between the pre-training and fine-tuning stages.

Similarly to (Flam-Shepherd and Aspuru-Guzik 2023), we use the XYZ representation as a base format to describe the spatial atom allocation. The idea of XYZ format is to represent the atom type and its 3D coordinates within every line of text. The main drawback of this format is the lack of charge and connectivity information. One should use external software like RDKit or OpenBabel (O’Boyle et al. 2011) to reconstruct the molecular graph. It introduces instability since even small noise in atom positions can drastically change the reconstructed graph or even break it down (Peng et al. 2023b). To alleviate that, we propose to couple the XYZ format with the SMILES format. The latter can efficiently represent the molecular structure, while the former allows for describing atom positions. To align these two formats we enforce to have the same atom ordering in both. We also remove the atom symbol from the XYZ representation as it already was shown in SMILES. For proteins, there is no need to describe their connectivity, therefore we simply write atom names grouped by amino-acids.

A schematic example of the two kinds of model input is shown in Figure 2 and a very detailed example of concrete input sequences (including its tokenization) is shown in Appendix A. In particular, the sequence starts with the <LIGAND> token followed by a SMILES string, tokenized at the character level. Next, there goes the <XYZ> special token marking the end of SMILES and the beginning of the coordinate part of the string. The tokenization strategy uses 6 tokens per 3D position: we use one token for the integer part and one token for the fractional part of the number. When working with protein pockets, we use a similar strategy. Specifically, the sequence begins with the <POCKET> token followed by the sequence of atoms where each atom is a separate token. Since pockets can be hundreds of atoms large, we follow the AlphaFold’s (Jumper et al. 2021) approach and retain only the 3D coordinates of the Alpha-carbon atoms in the corresponding amino acids.

3.2 Pre-Training

In this work, we aim to leverage insights accumulated by the NLP community in the paradigm of Large Language Models: pre-training and fine-tuning, prompting, scaling, fine-tuning with Reinforcement Learning, tool use, etc. (Hoffmann et al. 2022; Radford et al. 2019). Since our model covers only a specialized domain of molecular tasks, it does not require trillion-scale diverse datasets for good performance as NLP tasks do. Thus, we use a large-scale but specialized dataset of 3D molecules and protein pockets. During pre-training, we use the model with 108M parameters consisting of 15 layers, 12 heads, and a hidden size of 768. We found this size of the model to be enough for the tasks we care about - generating molecules in 3D (See Appendix C for a justification of the size). Every sequence in the training batch is either a ligand sequence of tokens or a pocket sequence of tokens following the scheme described earlier. Since the dataset has much fewer pockets than ligands, for one epoch of training

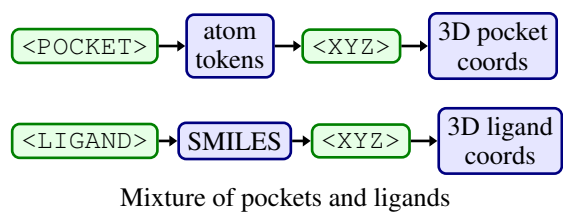


Figure 2: Data layout during the pre-training. Arrows show the token sequence order. Nodes such as `<POCKET>` show special tokens. Training is done on a mixture of pocket and ligand datasets. Concrete examples are shown in Appendix A.

on ligands, we do 5 epochs of training on proteins, that is, around 8% of all tokens seen by the model are pocket tokens. To speed up and stabilize pre-training, we use large batch training with 1.6M tokens per one training step. We found that having many tokens per batch is important for stable training in this task, even with smaller learning rates. A detailed description of the training implementation is provided in Appendix G. Despite the wide use of transformers in drug discovery, the majority of current works in this space do not use recent advancements of efficient Language Models pre-training: neither technical ones, such as Flash-attention (Dao 2023) or DeepSpeed (Rasley et al. 2020), nor the algorithmic ones, such as learning rate scaling. Our work aims to fill this gap by demonstrating the effectiveness of the pre-training for 3D drug discovery.

3.3 Fine-Tuning

Supervised Fine-Tuning (SFT) As a result of the pre-training, BindGPT gains an understanding of a broad chemical space. This comprehensive understanding enables us to efficiently narrow it down through the supervised fine-tuning on a specialized dataset. During the supervised fine-tuning phase, we continue model training on CrossDocked 2020 (Francoeur et al. 2020), which is a high-quality dataset containing aligned pocket-ligand pairs. Most of the prior methods subsample less than 1% of the best pocket-ligand pairs and they do not benefit from its diversity and scale. To obtain a bigger version of CrossDocked, we extract all intermediate ligand poses (with respect to the docking process), including the lower-quality ones. Despite quite large size, CrossDocked was created by docking 14k unique molecules into 3k pockets (Francoeur et al. 2020). This is why we observed a dramatic overfitting when training on the 1% version of CrossDocked and even on the full one. To alleviate that, we resort to two standard augmentation techniques used in drug discovery. First, we employ SMILES randomization (Bjerrum 2017), which can heavily randomize one molecule by yielding 100-1000 different SMILES strings (all corresponding to that molecule). Second, we randomly rotate the 3D coordinates of the protein pocket and of the ligand (with the same rotation matrix). This way our model learns to *understand structural and spatial properties of molecular binding beyond just token sequences*.

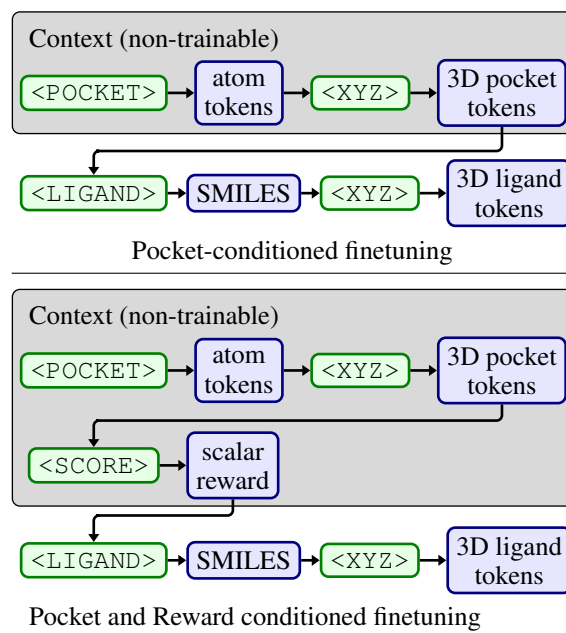


Figure 3: Data layout during the fine-tuning. Arrows show the token sequence order. Nodes such as `<POCKET>` show special tokens. Concrete examples are shown in Appendix A.

Since the pre-trained BindGPT is trained on both ligands (starting from the `<LIGAND>` token) and pockets (starting from the `<POCKET>` token), the information about the structure of both is learned by the model. In our fine-tuning setup, we represent each pocket-ligand pair as a string starting with the pocket string representation followed by the string representation of the ligand (See Section 3.1 for their description). Therefore, having learned them separately during pre-training, the fine-tuning exploits the independent knowledge of both pockets and ligands to learn a conditional dependency between them. In addition to that, since our version of CrossDocked contains both high and low score conformations, we test another version of the context where we condition on the pocket and binding energy score obtained from the CrossDocked dataset (which is originally computed through the docking software (Trott and Olson 2010; Eberhardt et al. 2021)). This way we can perform a variant of contrastive learning by learning the structure of good and bad examples. During an evaluation of the model, we can sample molecules conditioned on some desired value of the binding affinity. The input layout for both versions is shown in Figure 3.

Reinforcement Learning Despite the ubiquitous use of Reinforcement Learning (RL) for language models in Drug Discovery (see Section 2 and Appendix G), we did not find it being used within the pre-training paradigm of modern LLMs (Hoffmann et al. 2022; Ouyang et al. 2022). Our main motivation to use RL after the pre-training and fine-tuning stages is to use the knowledge distilled into the model from massive amounts of less structured data. **We believe, this is the first work performing reinforcement learning on**

	XYZ-TF		BindGPT (Ours)	
	No-H	H	No-H	H
Valid (\uparrow)	12.87%	17.86%	98.58%	77.33%
SA (\uparrow)	0.21	0.54	0.77	0.78
QED (\uparrow)	0.30	0.37	0.59	0.61
Lipinski (\uparrow)	4.79	4.82	4.86	4.91
RMSD (\downarrow)	-	-	0.89	3.44
Time,s (\downarrow)	165	394	13	156

Table 1: Generative metrics for the molecule generation task after the pre-training. (H) indicates that explicit hydrogens are generated with molecules. For XYZ-TF, the RMSD calculation algorithm failed to converge.

small molecules that utilizes knowledge from previous pre-training and supervised fine-tuning. Despite there being dozens of works doing RL with LMs on molecules, none of them do that within the LLM paradigm and none of them consider target-conditioned RL problem. In our opinion, the latter is primarily due to pocket-conditioned generation is not possible without large-scale pre-training as we show in the experimental section.

We apply the REINFORCE algorithm for further model fine-tuning. It allows using the feedback (called *reward*) from an external oracle to train the model to generate even better structures than those generated after the SFT stage. The resulting RL-fine-tuned model can generalize and produce high-affinity molecules even for the new pockets. In our procedure, on each training step, we generate the 3D structure of ligands for a batch of random protein pockets. Then, we compute the reward using an external docking software that estimates the binding energy between the pocket and the generated ligand. The final step involves updating the language model with the batch of prompts (pockets), responses (ligands), and rewards (binding energies). We initially tested PPO (Schulman et al. 2017) and REINVENT (Olivecrona et al. 2017), but found REINFORCE to be more stable for our project, which aligns with another recent finding in the field of RL applied to language models in NLP (Ahmadian et al. 2024). Also, it’s important to mention, that we apply the KL-penalty between the model’s initialized and current state to stabilize the procedure. Further details, such as specific hyperparameters, can be found in the Appendix B.

4 Results

In this section, we describe our experimental results. We start with a brief data description followed by the description of the three 3D molecular generative tasks: the 3D generative modeling of molecules and conformation generation given molecular graph (in Section 4.1 and Appendix H) and then pocket-conditioned generation (in Section 4.2, Appendix I).

For pre-training, we use a large 3D molecular dataset proposed by the authors of the Uni-Mol model (Zhou et al. 2023). The dataset contains 208M conformations for 12M molecules and 3.2M spatial structures of protein pockets. For finetuning in the pocket-conditioned generation task, we use the aforementioned CrossDocked dataset which contains aligned pocket-molecule pairs. Our filtration of the dataset has around

Group	Metrics	EDM	Mol Diff	BindGPT (Ours)
Druglikeness	QED (\uparrow)	0.558	0.668	0.616
	SA (\uparrow)	0.568	0.874	0.826
	Lipinski (\uparrow)	4.923	4.986	4.896
3D structures	JS. bond lengths (\downarrow)	0.246	0.365	0.029
	JS. bond angles (\downarrow)	0.282	0.155	0.075
	JS. dihedral angles (\downarrow)	0.328	0.162	0.098
Bonds	JS. num. bonds per atoms (\downarrow)	0.139	0.115	0.160
	JS. freq. bond types (\downarrow)	0.378	0.163	0.045
	JS. freq. bond pairs (\downarrow)	0.396	0.136	0.043
	JS. freq. bond triplets (\downarrow)	0.449	0.125	0.042
Rings	JS. num. rings (\downarrow)	0.106	0.062	0.094
	JS. num. n-sized rings (\downarrow)	0.107	0.092	0.023
	Num. Intersecting rings (\uparrow)	3.667	8.000	9.000
Speed	Time for 1000 valid molecules, s (\downarrow)	1.4e6	7500	200

Table 2: The qualities of the generated 3D molecules after pre-training on the Uni-Mol data followed by fine-tuning on GEOM-DRUGS. Note that the main metrics here are 3D ones: BindGPT outperforms baselines in nearly all of them.

27M pocket-ligand pairs covering a cross product of 14k molecules with 3k pockets (not all of the pairs are present, and for some of them, there is more than one pose, each with a different score). We also hold out a set of 100 pockets from the training data for evaluating the model performance. For the tasks of 3D molecule and 3D conformer generation (Section 4.1), to make comparisons with baselines more fair, we also fine-tune the model on the GEOM-DRUGS (Axelrod and Gómez-Bombarelli 2022) dataset, with drug-like molecules having high-quality 3D molecular conformations. This dataset contains 27M conformations for 300k molecules, and it serves as a standard benchmark for machine learning-based 3D molecular generators. Finally, we use the Platinum (Friedrich et al. 2017) dataset as a hold-out evaluation dataset to test our model and baselines on zero-shot conformer generation. The Platinum dataset contains the best-in-class experimentally validated conformations for testing conformer generation software.

4.1 Generative Modeling of 3D Molecules

Metrics. We provide the validity (\uparrow) of generated molecules and druglikeness metrics - SA (\uparrow), QED (\uparrow), and Lipinski (\uparrow) that are agnostic to 3D but measure how likely the molecule to be a drug. Also, we adopt a range of distribution metrics that were used for the MolDiff method (Peng et al. 2023b). Those metrics measure the discrepancy between true and modeled molecular distributions by computing the Jensen-Shannon divergences on the set of molecular properties and feature distributions. We compute RMSD (Root-Mean-Squared-Distance) (\downarrow) - which measures the quality of 3D structures by aligning the generated one with the one from RDKit (i.e., we regenerate conformer via RDKit) and computing the atomwise distance. Finally, we measure the time needed to generate 1K valid 3D molecules on one GPU. Note that this choice

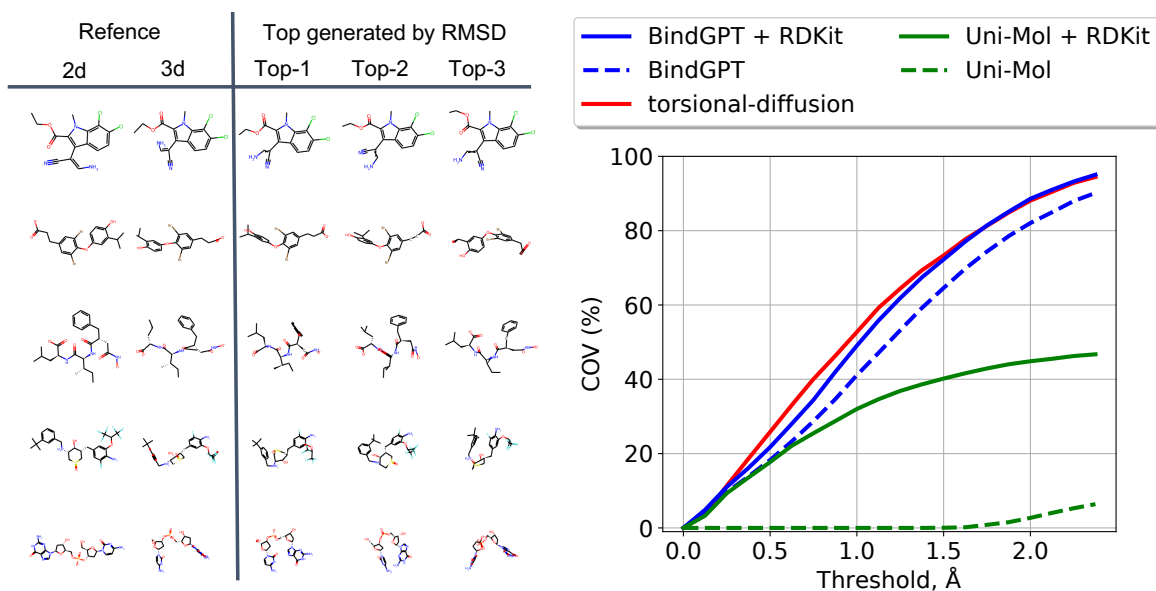


Figure 4: **(left)** Sampled conformations for reference molecules from the Platinum dataset (See Appendix F for a higher resolution image) **(right)** RMSD Coverage metric calculated on the Platinum dataset for the 3D conformation generation task.

of metrics is standard for this task (see Peng et al. (2023b) for a more detailed description of them). For the 3D conformation generation given molecule task, we compute the RMSD-coverage (\uparrow) metric. This is a standard performance metric for 3D conformer generation models (see e.g. Jing et al. (2022)). It is represented by the cumulative distribution function of RMSD between generated and reference conformers. The metric is a function of the threshold x : $P(\text{RMSD} < x)$. An ideal model should achieve the highest possible metric values at the lowest possible thresholds.

Baselines. For the molecule generation task, we consider the current best 3D generative models. EDM (Hoogeboom et al. 2022) and MolDiff (Peng et al. 2023b) are task-specialized diffusion models for 3D molecule generation. XYZ-Transformer (Flam-Shepherd and Aspuru-Guzik 2023) is another 3D molecular transformer that was proposed for small-scale data. Note that XYZ-TF is the only model capable of large-scale pre-training besides our model, so we pre-train only XYZ-TF and BindGPT on the Uni-Mol data. We also do the GEOM-DRUGS evaluation, where we report MolDiff and EDM trained on the full dataset and for BindGPT fine-tuned on the same version of it. For conformer generation, we compare BindGPT with the current state-of-the-art methods, Torsional Diffusion (Jing et al. 2022) and the Uni-Mol model (Zhou et al. 2023). The former is a specialized SE(3)-equivariant diffusion model capable of conformation generation only. The latter is a modified BERT (Devlin et al. 2019). As a coordinate-level encoder LM, the Uni-Mol model needs input coordinates to generate a conformation, which is why this model uses RDKit as a tool for initializing coordinates.

Results. The molecular generative modeling results are shown in Tables 2 and 1. First, the pre-trained BindGPT

model consistently outperforms the XYZ-TF baseline both without and with explicit hydrogens. The latter is a much more challenging task, and almost no baseline methods can do that (except EDM, which is not scalable) since reconstructing hydrogens can be done on a post-processing step, but explicit modeling of them makes the molecule size several times larger. BindGPT is the first model capable of modeling hydrogen explicitly at such a large scale. Also, XYZ-TF has a very low validity rate due to the need for graph reconstruction. Next, for the methods trained on the GEOM-DRUGS dataset, BindGPT (being fine-tuned on this data) shows state-of-the-art performance scores for nearly all distributional evaluation metrics. Even though BindGPT does not outperform MolDiff in Druglikeness, we argue that it does not have to - 3D distributional metrics evaluate how well the model learns 3D structures in the data, while druglikeness is not tied to any dataset and BindGPT shows the same values as after pre-training (see Table 1). For the conformation generation task, the current best baseline is Torsional Diffusion (TD) (Jing et al. 2022). We use the Platinum dataset to compare TD trained on GEOM-DRUGS with Uni-Mol-BERT and BindGPT, both of which are pre-trained and fine-tuned on the same data. Figure 4(b) shows the results for zero-shot evaluations on Platinum. Surprisingly, Uni-Mol fails to generalize to this new dataset (even assisted by RDKit), which we think is because of its structural diversity. BindGPT, in contrast, is capable of matching the performance of TD when assisted by the RDKit tool and having a small gap when not. All the above results demonstrate the generalizability of our model - none of the baselines is able to solve this wide range of tasks at this level of quality.

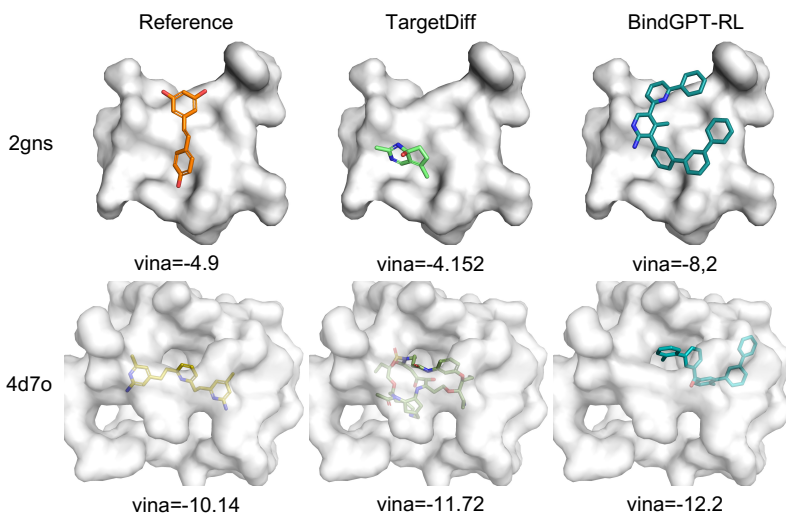


Figure 5: Examples of binding poses and Vina scores (\downarrow) for 2gns and 4d7o pockets.

Method	Vina score (\downarrow)	SA (\uparrow)	QED (\uparrow)	Lipinski (\uparrow)
Pocket2Mol	-7.15 ± 4.89	0.75 ± 0.12	0.57 ± 0.15	4.88 ± 0.37
TargetDiff	-7.80 ± 3.61	0.58 ± 0.12	0.48 ± 0.19	4.51 ± 0.85
BindGPT-FT (Ours)	-5.44 ± 2.09	0.78 ± 0.10	0.50 ± 0.17	4.72 ± 0.70
BindGPT-RFT (Ours)	-7.24 ± 1.68	0.74 ± 0.11	0.48 ± 0.22	4.32 ± 1.25
BindGPT-RL (Ours)	-8.60 ± 1.90	0.84 ± 0.05	0.43 ± 0.17	4.81 ± 0.52

Table 3: Generative metrics for the pocket-conditioned generation task. Note that the reward function optimizes the raw vina score (see Appendix B.3). This is done for a fair comparison with baselines: other metrics like SA or QED could easily be included into the reward, but baselines only optimize vina score via sampling constraints so we compare them with RL optimization.

4.2 Pocket-Conditioned Molecule Generation

Metrics. The main metrics for this task include the measure of ligand-pocket affinity; and druglikeness of the ligand. The first one is the binding energy (\downarrow) computed by the QVINA (Alhossary et al. 2015) docking software, while the second one comprises the druglikeness metrics (SA (\uparrow), QED (\uparrow), and Lipinsky (\uparrow)). For each baseline, we report the time required to generate 100 valid molecules for one pocket.

Baselines. Apart from the BindGPT model, we include the baselines such as 3D diffusion model (TargetDiff (Guan et al. 2023)) and autoregressive Graph Neural Network (Pocket2Mol (Peng et al. 2022)). Note that none of the baselines perform large-scale pre-training. Instead, they resort to heavy inductive biases to efficiently learn from small data.

Results. Performance of our approach is summarized in Table 3. We depict the performance of three versions of BindGPT. First, BindGPT-FT is a model fine-tuned on the complete CrossDocked data (the data layout as described in Figure 3, top), i.e. both good and bad binding pairs. This model serves as an initialization for the RL model. Second, BindGPT-RFT is the model fine-tuned on CrossDocked with the reward in the context. To get higher affinity molecules from that model, we condition the model on ran-

dom binding energy values within $[-12, -10]$, which are the best scores observed by the model (in around 0.1% of examples). Finally, the BindGPT-RL model is trained with RL (see Section 3.3 and Appendix B.3) from the BindGPT-FT initialization. Our main conclusion is that the RL fine-tuned model can learn to search the space of binding molecules much more efficiently and significantly outperforms all the previous best baselines in terms of binding energy.

5 Discussion and Conclusion

We presented BindGPT, a framework for training capable language models that can generate 3D molecules. Through a series of studies on a range of different 3D molecular tasks, we demonstrate the generality of our approach as it can solve each of them by matching or surpassing the baselines. Our method does not have any inductive biases about the domain acting as a data-driven approach. Unlike all the baselines which have inductive biases, our method solves each downstream task without any such assumptions. The task of particular interest in our work is pocket-based molecule generation, where our model outperforms all the baselines by a large margin. We show that the large-scale pre-training, fine-tuning, and RL can be used in 3D drug discovery.

Acknowledgements

Sarath Chandar is supported by the Canada CIFAR AI Chairs program, the Canada Research Chair in Lifelong Machine Learning, and the NSERC Discovery Grant.

References

- Ahmadian, A.; Cremer, C.; Gallé, M.; Fadaee, M.; Kreutzer, J.; Pietquin, O.; Üstün, A.; and Hooker, S. 2024. Back to Basics: Revisiting REINFORCE Style Optimization for Learning from Human Feedback in LLMs. arXiv:2402.14740.
- Alhossary, A.; Handoko, S. D.; Mu, Y.; and Kwok, C.-K. 2015. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, 31(13): 2214–2216.
- Axelrod, S.; and Gómez-Bombarelli, R. 2022. GEOM, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1): 185.
- Bagal, V.; Aggarwal, R.; Vinod, P. K.; and Priyakumar, U. D. 2022. MolGPT: Molecular Generation Using a Transformer-Decoder Model. *Journal of Chemical Information and Modeling*, 62(9): 2064–2076.
- Bjerrum, E. J. 2017. SMILES Enumeration as Data Augmentation for Neural Network Modeling of Molecules. *CoRR*, abs/1703.07076.
- Black, S.; Biderman, S.; Hallahan, E.; Anthony, Q.; Gao, L.; Golding, L.; He, H.; Leahy, C.; McDonell, K.; Phang, J.; Pieler, M.; Prashanth, U. S.; Purohit, S.; Reynolds, L.; Tow, J.; Wang, B.; and Weinbach, S. 2022. GPT-NeoX-20B: An Open-Source Autoregressive Language Model. In *Proceedings of the ACL Workshop on Challenges & Perspectives in Creating Large Language Models*.
- Chilingaryan, G.; Tamoyan, H.; Tevosyan, A.; Babayan, N.; Khondkaryan, L.; Hambardzumyan, K.; Navoyan, Z.; Khachatryan, H.; and Aghajanyan, A. 2022. BARTSmiles: Generative Masked Language Models for Molecular Representations. arXiv:2211.16349.
- Corso, G.; Stärk, H.; Jing, B.; Barzilay, R.; and Jaakkola, T. S. 2023. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. In *The Eleventh International Conference on Learning Representations*.
- Dao, T. 2023. FlashAttention-2: Faster Attention with Better Parallelism and Work Partitioning.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Burstein, J.; Doran, C.; and Solorio, T., eds., *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. Minneapolis, Minnesota: Association for Computational Linguistics.
- Eberhardt, J.; Santos-Martins, D.; Tillack, A. F.; and Forli, S. 2021. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, 61(8): 3891–3898.
- Feng, W.; Wang, L.; Lin, Z.; Zhu, Y.; Wang, H.; Dong, J.; Bai, R.; Wang, H.; Zhou, J.; Peng, W.; Huang, B.; and Zhou, W. 2024. Generation of 3D molecules in pockets via a language model. *Nature Machine Intelligence*.
- Flam-Shepherd, D.; and Aspuru-Guzik, A. 2023. Language models can generate molecules, materials, and protein binding sites directly in three dimensions as XYZ, CIF, and PDB files. arXiv:2305.05708.
- Flam-Shepherd, D.; Zhu, K.; and Aspuru-Guzik, A. 2022. Language models can learn complex molecular distributions. *Nature Communications*, 13(1).
- Francoeur, P. G.; Masuda, T.; Sunseri, J.; Jia, A.; Iovanisci, R. B.; Snyder, I.; and Koes, D. R. 2020. 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.
- Friedrich, N.-O.; Meyder, A.; de Bruyn Kops, C.; Sommer, K.; Flachsenberg, F.; Rarey, M.; and Kirchmair, J. 2017. High-Quality Dataset of Protein-Bound Ligand Conformations and Its Application to Benchmarking Conformer Ensemble Generators. *Journal of Chemical Information and Modeling*, 57(3): 529–539.
- Guan, J.; Qian, W. W.; Peng, X.; Su, Y.; Peng, J.; and Ma, J. 2023. 3D Equivariant Diffusion for Target-Aware Molecule Generation and Affinity Prediction. In *The Eleventh International Conference on Learning Representations*.
- Ho, J.; Jain, A.; and Abbeel, P. 2020. 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, 6840–6851. Curran Associates, Inc.
- Hochreiter, S.; and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, 9(8): 1735–1780.
- Hoffmann, J.; Borgeaud, S.; Mensch, A.; Buchatskaya, E.; Cai, T.; Rutherford, E.; de Las Casas, D.; Hendricks, L. A.; Welbl, J.; Clark, A.; Hennigan, T.; Noland, E.; Millican, K.; van den Driessche, G.; Damoc, B.; Guy, A.; Osindero, S.; Simonyan, K.; Elsen, E.; Rae, J. W.; Vinyals, O.; and Sifre, L. 2022. Training Compute-Optimal Large Language Models. arXiv:2203.15556.
- Hoogeboom, E.; Satorras, V. G.; Vignac, C.; and Welling, M. 2022. 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*, 8867–8887. PMLR.
- Hu, L.; Benson, M. L.; Smith, R. D.; Lerner, M. G.; and Carlson, H. A. 2005. Binding MOAD (mother of all databases). *Proteins: Structure, Function, and Bioinformatics*, 60(3): 333–340.
- Jing, B.; Corso, G.; Chang, J.; Barzilay, R.; and Jaakkola, T. S. 2022. Torsional Diffusion for Molecular Conformer Generation. In Oh, A. H.; Agarwal, A.; Belgrave, D.; and Cho, K., eds., *Advances in Neural Information Processing Systems*.
- Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873): 583–589.

- Katharopoulos, A.; Vyas, A.; Pappas, N.; and Fleuret, F. 2020. Transformers are RNNs: Fast Autoregressive Transformers with Linear Attention. In *Proceedings of the International Conference on Machine Learning (ICML)*.
- Lewis, M.; Liu, Y.; Goyal, N.; Ghazvininejad, M.; Mohamed, A.; Levy, O.; Stoyanov, V.; and Zettlemoyer, L. 2019. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. arXiv:1910.13461.
- Lin, H.; Huang, Y.; Liu, M.; Li, X.; Ji, S.; and Li, S. Z. 2022. DiffBP: Generative Diffusion of 3D Molecules for Target Protein Binding. arxiv:2211.11214.
- Luo, S.; Guan, J.; Ma, J.; and Peng, J. 2021. 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, 6229–6239. Curran Associates, Inc.
- Niu, C.; Song, Y.; Song, J.; Zhao, S.; Grover, A.; and Ermon, S. 2020. Permutation Invariant Graph Generation via Score-Based Generative Modeling. arXiv:2003.00638.
- O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; and Hutchison, G. R. 2011. Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(1): 33.
- Olivecrona, M.; Blaschke, T.; Engkvist, O.; and Chen, H. 2017. Molecular De Novo Design through Deep Reinforcement Learning. arXiv:1704.07555.
- Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C. L.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; Schulman, J.; Hilton, J.; Kelton, F.; Miller, L.; Simens, M.; Askell, A.; Welinder, P.; Christiano, P.; Leike, J.; and Lowe, R. 2022. Training language models to follow instructions with human feedback. arXiv:2203.02155.
- Peng, X.; Guan, J.; Liu, Q.; and Ma, J. 2023a. MolDiff: Addressing the Atom-Bond Inconsistency Problem in 3D Molecule Diffusion Generation. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 27611–27629. PMLR.
- Peng, X.; Guan, J.; Liu, Q.; and Ma, J. 2023b. MolDiff: Addressing the Atom-Bond Inconsistency Problem in 3D Molecule Diffusion Generation. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 27611–27629. PMLR.
- Peng, X.; Luo, S.; Guan, J.; Xie, Q.; Peng, J.; and Ma, J. 2022. Pocket2Mol: Efficient Molecular Sampling Based on 3D Protein Pockets. In *International Conference on Machine Learning*.
- Polishchuk, P. G.; Madzhidov, T. I.; and Varnek, A. 2013. Estimation of the size of drug-like chemical space based on GDB-17 data. *J Comput Aided Mol Des*, 27(8): 675–679.
- Radford, A.; Narasimhan, K.; Salimans, T.; and Sutskever, I. 2018. Improving language understanding by generative pre-training.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; and Sutskever, I. 2019. Language Models are Unsupervised Multitask Learners.
- Rasley, J.; Rajbhandari, S.; Ruwase, O.; and He, Y. 2020. DeepSpeed: System Optimizations Enable Training Deep Learning Models with Over 100 Billion Parameters. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD ’20*, 3505–3506. New York, NY, USA: Association for Computing Machinery. ISBN 9781450379984.
- 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. 2023. Structure-based Drug Design with Equivariant Diffusion Models. arXiv:2210.13695.
- Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal Policy Optimization Algorithms. arXiv:1707.06347.
- Su, J.; Lu, Y.; Pan, S.; Wen, B.; and Liu, Y. 2021. RoFormer: Enhanced Transformer with Rotary Position Embedding. *CoRR*, abs/2104.09864.
- Trott, O.; and Olson, A. J. 2010. 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.
- Weininger, D. 1988. 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.
- Yu, W.; and MacKerell, A. D. 2017. *Computer-Aided Drug Design Methods*, 85–106. New York, NY: Springer New York. ISBN 978-1-4939-6634-9.
- Zhou, G.; Gao, Z.; Ding, Q.; Zheng, H.; Xu, H.; Wei, Z.; Zhang, L.; and Ke, G. 2023. Uni-Mol: A Universal 3D Molecular Representation Learning Framework. In *The Eleventh International Conference on Learning Representations*.