Generative Models for Structure-Based Drug Design

Angelos Nikitaras [1]

Abstract

The integration of generative models into structure-based drug design (SBDD) has the potential to revolutionize the field, offering unprecedented opportunities for the rapid discovery of novel therapeutics. This paper provides a comprehensive review of recent advances in generative models for SBDD, highlighting the key challenges in the field. We discuss state-of-the-art methods for *de novo* molecular design, focusing on novel 3D approaches, such as diffusion-based models. We also address the limitations of current methods and propose potential directions for future research.

1. Introduction

Structure-based drug design (SBDD) represents a powerful approach for the design and optimization of novel drugs by exploiting the three-dimensional structure of biological targets (Anderson, 2003). The traditional drug discovery pipeline involves the screening of large chemical libraries to identify compounds that bind to a specific target, followed by optimization of the lead compounds through iterative rounds of synthesis and testing. This process is time-consuming and expensive, with the cost of developing a new drug estimated to be close to 800 million USD (Batool et al., 2019). In addition, the space of drug candidates is gigantic, with approximately 10^{60} molecules, while the number of substances that have been synthesized and tested are only a tiny fraction of this number (Polishchuk et al., 2013). This renders the traditional methods insufficient for exploring the vast chemical space, necessitating the development of new computational approaches.

A promising direction for accelerating the drug discovery process is the utilization of generative artificial intelligence, and in particular, Deep Generative Models (Bond-Taylor et al., 2021). These models leverage neural networks to learn the underlying distribution of data and generate new samples from it, with remarkable success in a wide range of applications, such as image generation and text synthesis (Podell et al., 2023; Achiam et al., 2023). In the context of SBDD, these models try to capture the conditional distribution of the ligands given the protein structure, enabling efficient sampling of novel drug candidates with high binding affinity to the specific target (Peng et al., 2022). There are important challenges associated with this task, including the choice of a suitable representation for molecular structures and the design of appropriate generative models which respect the underlying symmetries of the physical system (Du et al., 2022; Hoogeboom et al., 2022). Addressing these challenges is an active area of research, with recent advances showing promising results (Guan et al., 2024).

This paper aims to provide a comprehensive review of the recent developments in generative models for SBDD. We first introduce the theoretical backbone of Diffusion models and Graph Neural Networks (GNNs), two key components of the most recent works in the field. We then provide a detailed analysis of the different methods for *de novo* molecular design, concluding with a discussion of the major challenges and potential future directions.

2. Background

2.1. Denoising Diffusion Probabilistic Models

Denoising Diffusion Probabilistic Models are a class of latent variable models that have gained significant attention in recent years (Ho et al., 2020). They consist of two main components: a forward diffusion process and a reverse denoising process. During the forward diffusion process, noise is gradually injected to the data, transforming it into pure isotropic Gaussian noise at the final step. Conversely, the reverse denoising process, typically parametrized by a neural network, aims to reconstruct the original data from the noisy samples.

More specifically, let $\boldsymbol{x}_0 \in \mathbb{R}^d$ be a sample from the real distribution $p(\boldsymbol{x})$. Then, the forward diffusion process produces a sequence of noisy samples $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$, according to the following conditional distribution:

[1]Department of Computer Science, Technical University of Munich, Munich, Germany. Correspondence to: Angelos Nikitaras <angelos.nikitaras@tum.de>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

$$q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_t; \sqrt{1-\beta_t}\,\boldsymbol{x}_{t-1}, \beta_t \boldsymbol{I}), \quad (1)$$

where $\beta_t \in (0, 1)$. The data distribution at any intermediate step t can be computed in closed form as:

$$q(\boldsymbol{x}_t|\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_t; \sqrt{\bar{\alpha}}\,\boldsymbol{x}_0, (1-\bar{\alpha})\boldsymbol{I}), \quad (2)$$

where $a_t = 1 - \beta_t$ and $\bar{\alpha} = \prod_{i=1}^{t} a_i$.

Using Bayes theorem, we can also obtain the reverse conditional probability conditioned on $\boldsymbol{x}_0$:

$$q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_{t-1}; \tilde{\boldsymbol{\mu}}_t(\boldsymbol{x}_t, \boldsymbol{x}_0, \tilde{\beta}_t), \tilde{\beta}_t \boldsymbol{I}), \quad (3)$$

where $\tilde{\mu}_t$ and $\tilde{\beta}_t$ can be analytically computed.

Ultimately, the goal is to approximate the reverse process with a probability distribution $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$, where θ are the parameters to learn. Given that we are dealing with a latent variable model, where $x_1, \ldots, x_T$ represent the latent variables, we can optimize the parameters θ by maximizing the Evidence Lower Bound (ELBO) of the data distribution $p(\mathbf{x})$:

$$\log p(\mathbf{x}) \geq -\mathbb{E}_{q(\mathbf{x}_{1:T}|\mathbf{x}_0)} \left(\log \frac{q(\mathbf{x}_{1:T}|\mathbf{x}_0)}{p_\theta(\mathbf{x}_{0:T})} \right) \quad (4)$$

$$= KL\left[q(\mathbf{x}_T|\mathbf{x}_0)||p_\theta(\mathbf{x}_T)\right] - \mathbb{E}_q[log\ p_\theta(\mathbf{x}_0|\mathbf{x}_1)] \quad (5)$$

$$+ \sum_{t=2}^{T} KL[q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)||p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)] \quad (6)$$

The first term is constant, while the second term is typically approximated with a separate model (Ho et al., 2020). To approximate the third term, the distribution p_θ is modeled as a Gaussian distribution parametrized by a neural network. In practice though, it has been proven more effective to let the neural network predict the noise $\boldsymbol{\epsilon}_\theta$, using the reparametrization trick (Kingma & Welling, 2013).

2.2. Graph Neural Networks

In SBDD, Graph Neural Networks (GNNs) are a natural choice for modeling molecular structures. For example, diffusion based models typically employ GNNs to parametrize the noise predictor $\boldsymbol{\epsilon}_\theta$ (Schneuing et al., 2022). Here, we consider in particular the so called message-passing GNNs, which update the node representations by aggregating information from neighboring nodes (Gilmer et al., 2017). Given a graph $G = (V, E)$, where V is the set of nodes and E is the set of edges, we define the node representation update rule as follows:

$$\mathbf{m}_{ij} = \phi_e(\mathbf{h}_i^l, \mathbf{h}_j^l, \mathbf{e}_{ij}), \quad (7)$$

$$\mathbf{m}_i = \sum_{j \in N(i)} \mathbf{m}_{ij}, \quad (8)$$

$$\mathbf{h}_i^{l+1} = \phi_v(\mathbf{h}_i^l, \mathbf{m}_i), \quad (9)$$

where $\mathbf{h}_i^l$ is the representation of node i at layer l, $\mathbf{e}_{ij}$ is the edge feature between nodes i and j, ϕ_e is the edge update function, ϕ_v is the node update function, and $N(i)$ is the set of neighbors of node i. Most commonly, the update functions are Multi-Layer Perceptrons (MLPs) with learnable parameters.

The geometric nature of GNNs makes them particularly well-suited for SBDD, since they can take into account symmetry considerations which are not easily captured by the traditional MLPs (Bronstein et al., 2017). In this context, the notions of invariance and equivariance play a crucial role. Without delving into the mathematical details, a model is said to be invariant to a transformation if the output of the model remains the same under the transformation. On the other hand, a model is said to be equivariant to a transformation if the output of the model transforms in the same way as the input under the transformation. To give an example, assume a charge distribution, with electric dipole moment $\mathbf{p}$ and energy E, as shown in Figure 1. Rotating the charge distribution leads to the same energy, but the dipole moment rotates, according to the rotation of the charge distribution. In this case, the energy is invariant under the rotation, while the dipole moment is equivariant.

For molecular structures, we are mainly interested in equivariance with respect to rotations, translations, and permutations of atoms. The general GNN architecture is inherently equivariant to permutations, but extra care is needed to ensure equivariance with respect to rotations and translations (Hoogeboom et al., 2022).

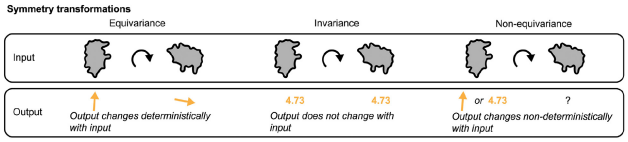

Figure 1. Illustration of the concepts of invariance and equivariance. An invariant quantity remains the same under certain transformations, while an equivariant quantity transforms in the same way as the input (Isert et al., 2023).

3. Generative models for SBDD

The main goal of *de novo* SBDD is the design of small molecules (*ligands*) that specifically bind to the cavity of

a target protein (*receptor*), thereby modulating its activity. The protein is composed of a sequence of amino acids, and its 3D structure is typically known either from experimental methods, such as X-ray crystallography, or from computational methods, such as protein structure prediction algorithms (Jumper et al., 2021). In contrast, the ligand is a small organic molecule, which can form a stable complex with the protein, by interacting with specific residues in the binding pocket.

To leverage the power of generative models for SBDD, a crucial first step is representing the molecular structures in a suitable format. While our focus will be on novel 3D methods, we will also briefly discuss sequence-based methods, which were the first to be applied to SBDD.

3D methods

Modern approaches in SBDD represent both the ligand and the protein in 3D space. A pioneering example of this is liGAN (Ragoza et al., 2022), which utilizes a voxel-based representation for both the ligand and receptor, where each voxel contains the atomic density of the corresponding atom, akin to 3D images. The model uses a conditional variational autoencoder to generate atomic density grids, which are then converted to 3D structures using a separate atom fitting algorithm. Despite its innovative approach, liGAN faces several limitations, such as the lack of rotational equivariance and poor scalability, due to the cubic complexity with respect to the pocket size.

To overcome these limitations, subsequent works consider the ligand and the receptor as 3D graphs or point clouds, harnessing the power of GNNs for context encoding. For instance, Luo et al. (2021) proposed a model which estimates the distribution of atom's occurrences in the 3D space. Equivariance is achieved through a rotationally and translationally invariant GNN-based context encoder, which learns a latent representation of the atoms belonging to the binding site. To sample from the learned distribution, they formulated an autoregressive algorithm that generates the atoms sequentially. Similarly, GraphBP (Liu et al., 2022) utilized the framework of autoregressive normalizing flows, to generate molecules atom by atom (Papamakarios et al., 2021). In this case, in every atom generation step, a previous atom was used as a local reference frame, to ensure the equivariance of the generation process. Despite the promising results of these methods, the major shortcoming of the autoregressive generation scheme is that it induces an artificial ordering of the atoms. This may lead to unrealistic structures, since the model may not be able to capture the global context of the ligand.

Recently, diffusion-based models have emerged as a powerful alternative for generating molecular structures in a non-autoregressive manner (Hoogeboom et al., 2022). In the context of SBDD, Guan et al. (2023) proposed TargetDiff, a diffusion-based model that generates ligands in 3D space, by diffusing the atom positions and atom types independently. The noise was injected only into the ligand atoms, since ligand generation was conditioned on the protein structure. To enforce the invariance of the likelihood with respect to translations and rotations, they shifted the center of mass of the protein to the origin and parametrized the reverse diffusion process with equivariant GNNs (Satorras et al., 2021). In addition, they showed that the latent representation learned by TargetDiff, can be used for binding affinity estimation of the generated ligands. Schneuing et al. (2022) introduced DiffSBDD, a similar model to TargetDiff, but they modeled additionally the joint distribution of the ligand and the receptor, by injecting noise into both structures. Taking cues from the problem of image inpainting, they developed a diffusion model which learns to fill in the "missing" ligand atoms into the binding pocket, using the joint distribution (Lugmayr et al., 2022). It is important to note that both TargetDiff and DiffSBDD generate a point cloud representation of the ligand, and a post-processing bond inference step is required to obtain the molecular structure. Moreover, bond information is not explicitly taken into account. Alternatively, DecompDiff (Guan et al., 2024) decomposes the ligand into fragments, consisting of the so-called *scaffold* and *arms*. In drug design, a scaffold is the core structure of a molecule, while arms are the chemical groups attached to the scaffold that can be modified to fine-tune its properties. By introducing informative priors of the decomposed arms and scaffold, it restricts the chemical space of the generated ligands, leading to more realistic structures. In addition, not only the atom positions are generated, but also the bonds between the atoms, leading to unambiguous molecular structures.

Sequence-based methods

The first attempts at generating molecular structures heavily relied on Simple Molecular Input Line Entry System (SMILES) notation, which represents molecules as strings of characters, a text-based notation that efficiently encodes the connectivity and atom types within a molecule (Weininger, 1988). The main advantage of using such a representation is that text generation models can be directly applied to generate SMILES strings. Xu et al. (2021) proposed a conditional RNN model, which generated SMILES strings conditioned on different descriptors of the target protein, using an architecture resembling conditional autoencoders (Kotsias et al., 2020). In a different direction, Ma et al. (2021) combined RNNs with Monte Carlo Tree Search (MCTS), utilizing docking scores as rewards to guide the generation process. More recently, inspired by the success of transformer models in natural language processing, Wang et al. (2023) proposed cMolGPT, a variant of Generative

Pre-trained Transformers that leverages the attention mechanism to capture long-range dependencies in the molecular structure (Radford et al., 2018; Vaswani et al., 2017).

Evaluation and Datasets

The evaluation of generative models for SBDD is a challenging task, with researchers employing diverse metrics tailored to the specific molecular representation. Common metrics include: (i) *Validity* is the percentage of generated molecules that are chemically valid. (ii) *Diversity* assesses the variety of the generated molecules. (iii) *Vina score* is a scoring function that estimates the binding affinity of the generated molecules to the target protein (Trott & Olson, 2010). Lower scores indicate stronger binding affinity. It is often calculated after re-docking the ligand to the target protein. (iv) *QED score* is an empirical measure quantifying the drug-likeness (Bickerton et al., 2012). (v) *SA score* is the synthetic accessibility score, which captures the difficulty of drug synthesis. Its values range from 0 to 1, with higher values indicating easier synthesis.

For 3D methods, the most commonly used dataset is the CrossDocked dataset, a massive collection of protein-ligand complexes (Francoeur et al., 2020). This dataset was constructed by first gathering experimentally verified protein-ligand complexes and then cross-docking the ligands with structurally similar, non-cognate proteins. In practice though, filtering steps are required to ensure validity and diversity of the structures. Another common choice is the Binding MOAD dataset, an experimentally determined dataset of protein-ligand complexes (Hu et al., 2005). Despite the high quality of the dataset, it has a limited number of protein-ligand complexes compared to the CrossDocked dataset.

4. Review

Starting with 1D methods, their inherent limitation lies in their inability to capture crucial three-dimensional conformational information. However, for chemists and biologists, 1D and 2D representations are often more interpretable, giving them direct insight into the chemical structure of the ligands. Notably, a gap in the literature exists concerning a comprehensive evaluation framework that facilitates a direct comparison between 1D and 3D methods. The choice of representation often remains a matter of preference, highlighting the need for a nuanced discussion on the trade-offs between these approaches.

Regarding 3D methods, Table 1 summarizes the results of the aforementioned models (Guan et al., 2024). Note that the Vina score is calculated *after* re-docking the ligand to the target protein. We observe that DecompDiff achieves the best Vina score, highlighting the effectiveness of incorporating chemical priors during ligand generation. Overall, diffusion models appear more adept at generating ligands with high binding affinity, a critical factor in SBDD. This could be attributed to their non-autoregressive nature, which facilitates an efficient exploration of the chemical space. In other metrics, the models show comparable performance, with Luo et al. (2021) demonstrating the highest QED and SA scores, and GraphBP achieving the highest diversity. While property-related metrics like QED and SA are valuable for initial assessment, their primary role lies in filtering out compounds that fall outside acceptable parameter ranges. In that regard, they serve as a rough guide, rather than a definitive measure of the quality of the generated molecules.

Method	Vina Score	QED	SA	Diversity
liGAN	-6.33	0.39	0.59	0.66
Luo et al	-6.75	**0.51**	**0.63**	0.70
GraphBP	-4.80	0.43	0.49	**0.79**
TargetDiff	-7.80	0.48	0.58	0.72
DecompDiff	**-8.39**	0.45	0.61	0.68

Table 1. Comparison of the average *Vina* score, *QED* score, *SA* score and *Diversity* of the generated ligands for different methods. All models are trained on the CrossDocked dataset. The best score for each metric is highlighted in bold.

Despite the promising results, there are several limitations of the current methods, inhibiting their adoption in real-world drug discovery pipelines. First, the lack of a common and comprehensive evaluation framework reflecting the diverse aspects of SBDD, makes it difficult to compare the performance of different models. In addition, current evaluations rely heavily on scoring functions, neglecting experimental validation, which is the ultimate criterion for the success of a drug candidate. On top of that, scoring functions are used in data generation too, making the methods even more dependent on their accuracy. Last, it is essential for the models to incorporate extensive chemical and pharmacological knowledge, to restrict the chemical space and enhance synthesis feasibility. The optimization problem is inherently multi-objective, and focusing only on binding affinity is not sufficient to replace the already established methods.

5. Conclusion

In this review, we have explored the recent advances in generative models for SBDD, focusing on the advantages and limitations of the current methods. Looking ahead, we anticipate a closer collaboration between machine learning researchers and domain experts, facilitating the development of models with tangible impact on drug discovery. The potential of generative models for SBDD is immense, and there is still much to be explored in this exciting field.

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

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

Batool, M., Ahmad, B., and Choi, S. A structure-based drug discovery paradigm. *International journal of molecular sciences*, 20(11):2783, 2019.

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., and Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012.

Bond-Taylor, S., Leach, A., Long, Y., and Willcocks, C. G. Deep generative modelling: A comparative review of vaes, gans, normalizing flows, energy-based and autoregressive models. *IEEE transactions on pattern analysis and machine intelligence*, 44(11):7327–7347, 2021.

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4): 18–42, 2017.

Du, Y., Fu, T., Sun, J., and Liu, S. Molgensurvey: A systematic survey in machine learning models for molecule design. *arXiv preprint arXiv:2203.14500*, 2022.

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

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *International conference on machine learning*, pp. 1263–1272. PMLR, 2017.

Guan, J., Qian, W. W., Peng, X., Su, Y., Peng, J., and Ma, J. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. *arXiv preprint arXiv:2303.03543*, 2023.

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

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

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

Hu, L., Benson, M. L., Smith, R. D., Lerner, M. G., and Carlson, H. A. Binding moad (mother of all databases). *Proteins: Structure, Function, and Bioinformatics*, 60(3): 333–340, 2005.

Isert, C., Atz, K., and Schneider, G. Structure-based drug design with geometric deep learning. *Current Opinion in Structural Biology*, 79:102548, 2023.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

Kingma, D. P. and Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Kotsias, P.-C., Arús-Pous, J., Chen, H., Engkvist, O., Tyrchan, C., and Bjerrum, E. J. Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. *Nature Machine Intelligence*, 2(5): 254–265, 2020.

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

Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., and Van Gool, L. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11461–11471, 2022.

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

Ma, B., Terayama, K., Matsumoto, S., Isaka, Y., Sasakura, Y., Iwata, H., Araki, M., and Okuno, Y. Structure-based de novo molecular generator combined with artificial intelligence and docking simulations. *Journal of Chemical Information and Modeling*, 61(7):3304–3313, 2021.

Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., and Lakshminarayanan, B. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021.

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

Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Muller, J., Penna, J., and Rombach, R. Sdxl: Improving latent diffusion models for high-resolution image synthesis. *arXiv preprint arXiv:2307.01952*, 2023.

Polishchuk, P., Madzhidov, T., and Varnek, A. Estimation of the size of drug-like chemical space based on gdb-17 data. *Journal of Computer-Aided Molecular Design*, 27: 675–679, 2013. doi: 10.1007/s10822-013-9672-4.

Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., et al. Improving language understanding by generative pre-training. 2018.

Ragoza, M., Masuda, T., and Koes, D. R. Generating 3d molecules conditional on receptor binding sites with deep generative models. *Chem. Sci.*, 13:2701–2713, 2022. doi: 10.1039/D1SC05976A. URL http://dx.doi.org/10.1039/D1SC05976A.

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

Schneuing, A., Du, Y., Harris, C., Jamasb, A., Igashov, I., Du, W., Blundell, T., Lió, P., Gomes, C., Welling, M., et al. Structure-based drug design with equivariant diffusion models. *arXiv preprint arXiv:2210.13695*, 2022.

Tang, Y., Moretti, R., and Meiler, J. Recent advances in automated structure-based de novo drug design. *Journal of Chemical Information and Modeling*, 64(6):1794–1805, 2024.

Tong, X., Liu, X., Tan, X., Li, X., Jiang, J., Xiong, Z., Xu, T., Jiang, H., Qiao, N., and Zheng, M. Generative models for de novo drug design. *Journal of Medicinal Chemistry*, 64(19):14011–14027, 2021.

Trott, O. and Olson, A. J. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Wang, Y., Zhao, H., Sciabola, S., and Wang, W. cmolgpt: A conditional generative pre-trained transformer for target-specific de novo molecular generation. *Molecules*, 28 (11):4430, 2023.

Weininger, D. 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, 1988.

Xie, W., Wang, F., Li, Y., Lai, L., and Pei, J. Advances and challenges in de novo drug design using three-dimensional deep generative models. *Journal of Chemical Information and Modeling*, 62(10):2269–2279, 2022.

Xu, M., Ran, T., and Chen, H. De novo molecule design through the molecular generative model conditioned by 3d information of protein binding sites. *Journal of Chemical Information and Modeling*, 61(7):3240–3254, 2021.

Zhang, Z., Yan, J., Liu, Q., and Che, E. A systematic survey in geometric deep learning for structure-based drug design. *arXiv preprint arXiv:2306.11768*, 2023.