Briefings in Bioinformatics, 22(6), 2021, 1–11

https://doi.org/10.1093/bib/bbab344
Review

Molecular design in drug discovery: a comprehensive review of deep generative models

Yu Cheng, Yongshun Gong, Yuansheng Liu, Bosheng Song and Quan Zou

Corresponding authors: Yuansheng Liu, College of Information Science and Engineering, Hunan University, 2 Lushan S Rd, Yuelu District, 410086 Changsha, China. E-mail: yuanshengliu@hnu.edu.cn; Bosheng Song. E-mail: boshengsong@hnu.edu.cn

Abstract

Deep generative models have been an upsurge in the deep learning community since they were proposed. These models are designed for generating new synthetic data including images, videos and texts by fitting the data approximate distributions. In the last few years, deep generative models have shown superior performance in drug discovery especially *de novo* molecular design. In this study, deep generative models are reviewed to witness the recent advances of *de novo* molecular design for drug discovery. In addition, we divide those models into two categories based on molecular representations in silico. Then these two classical types of models are reported in detail and discussed about both pros and cons. We also indicate the current challenges in deep generative models for *de novo* molecular design. *De novo* molecular design automatically is promising but a long road to be explored.

Key words: deep generative model; deep learning; de novo drug design; molecular design

Introduction

In the long-term struggle between humans and diseases, especially the recent pandemic of coronavirus disease 2019 (COVID-19), drugs are playing an increasingly significant role. However, the drug discovery process has been confronted with obstacles, which requires a great deal of manpower, material and financial resource. For example, the drug development cycle from preclinical target screening to final marketing takes an average of at least 13.5 years [1]. And developing a new drug for a pharmaceutical company is a costly tour of about 1.8 billion with a high failure rate. The challenging process of drug discovery is derived from the large and discrete searching space of chemical molecules [2]. Specifically, the scale of possible structures of drug-like compounds is between 10^{23} and 10^{60}, but a small proportion of them about 10^{8} are therapeutically relevant [3, 4]. Traditional methods like high-throughput screening [4] are inefficient as the required amounts of resources and small number of hit compounds are not balanced. Big data and high-performance computing capabilities have allowed artificial intelligence to surpass traditional brute strength [5]. With the widespread application of deep learning, it is naturally considered as the potential way for drug discovery. Deep learning has been utilized in drug discovery and development, providing a new direction in pharmaceutical science [6]. Some related applications are shown in Figure 1.

Deep learning, whose prototype was the perceptron known as neural networks for pattern recognition [7], aimed at learning the latent distribution and representation of data. The concept of deep learning was formally proposed for solving the vanishing gradient problem by Hinton *et al.* [8] in 2006. Then in the ImageNet image recognition competition, the team led by Hinton used the AlexNet model [9] that made a sensation for eliminating vanishing gradient via the 'ReLU' activation function. In 2016, the triumph of AlphaGo [10] proved that deep learning was promising in surpassing humans. Up to now, deep learning has

Yu Cheng is currently a graduate student at Hunan University. Her research interest is bioinformatics.
Yongshun Gong is an associate professor at Shandong University. His research interests machine learning and spatiotemporal data mining.
Yuansheng Liu is an assistant professor at Hunan University. His research interests are bioinformatics and compression algorithms.
Bosheng Song is an associate professor at Hunan University. His research interests include biocomputing and bioinformatics.
Quan Zou is a professor at University of Electronic Science and Technology of China. He is a senior member of IEEE and ACM. He majors in bioinformatics, machine learning and algorithms.
Submitted: 1 July 2021; **Received (in revised form):** 19 July 2021

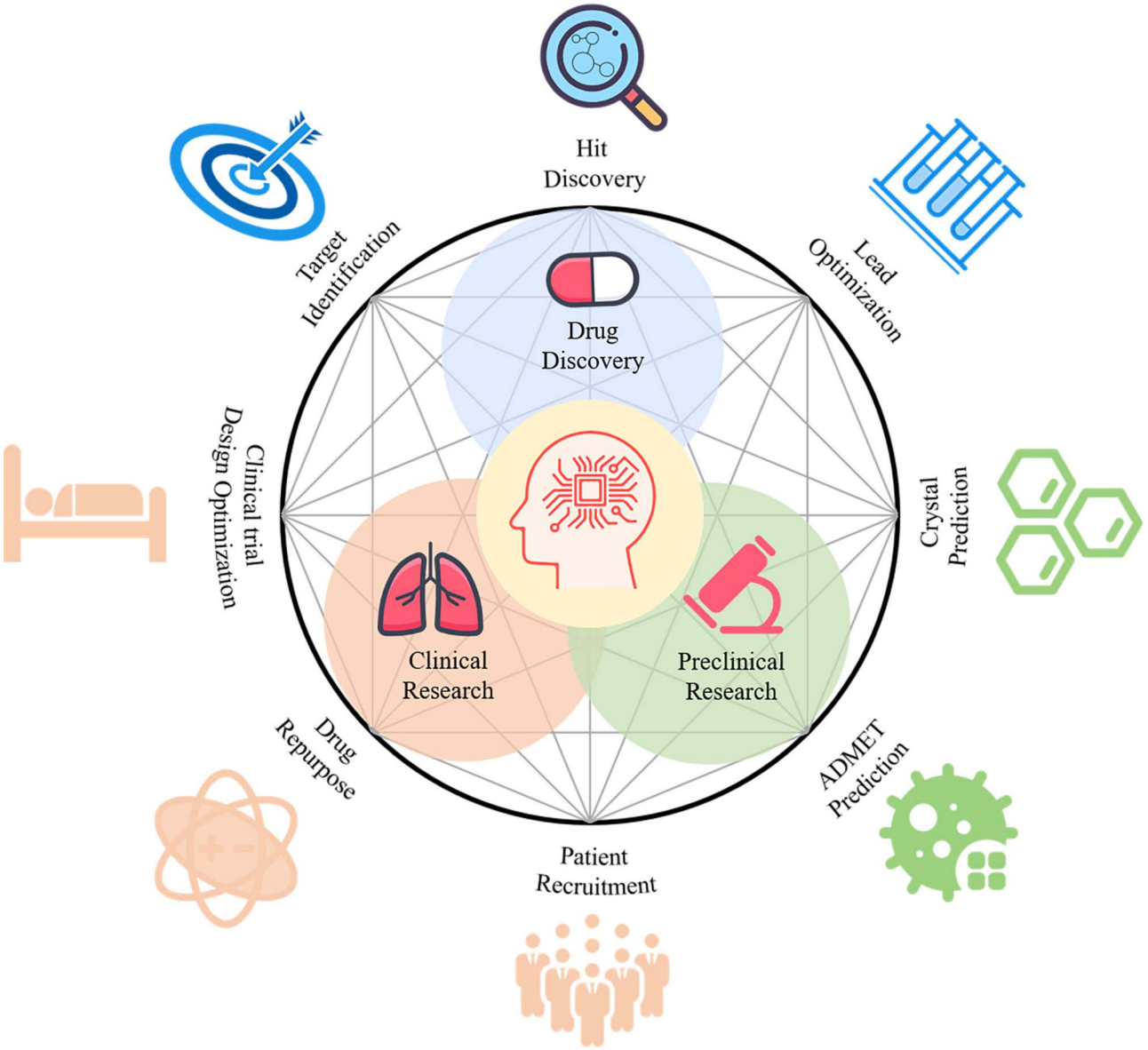

Figure 1. The applications of deep learning in drug discovery and development.

been applied successfully to computer vision [11, 12], natural language processing [13, 14], and some other fields [15, 16]. Deep neural networks are divided into discriminative models and generative models. The discriminative models which reflect the difference between heterogeneous data are to find the optimal classification [7]. The generative models, modeling the prior probability, represent the similarity of congener data.

Deep generative models develop rapidly as generating new synthetic data from given samples, including images [17], text [18] and video [19]. The representations of molecules in silico are similar to texts in natural language processing and graphs in social networks. Hence, it is natural to extend such models for *de novo* molecular design in drug discovery [20]. Different from using discriminative models to screen databases and classify molecules as active or inactive, deep generative models design new molecules with target properties from scratch. The desire for generating molecules automatically has been mentioned in the past by Gómez-Bombarelli et al. [21]. And in recent years, plenty of deep generative models have been devoted to boosting the de novo molecular design, which predominantly has followed two strategies based on the representations of molecules in silico. The first strategy focuses on a sequence representation—simplified molecular input line entry system (SMILES) [22], which utilizes deep generative models and text to generate moleculars. An alternative is to encode molecular into graphs [23] that learn to aggregate information (e.g., bond features and atoms). As a consequence, we categorize these typical models into two categories, i.e. SMILES-based and graph-based models.

In this review, we mainly focus on deep generative models for molecular generation in drug discovery. We first introduce the representation methods of molecules and conclude the prevalent databases. We show the pros and cons of different representations. As for generative models, we emphasize the recent advances based on different representations in the *de*

novo molecular design domain. The objective evaluation and the comparison of state-of-the-art models facilitate the selection and improvement for readers. However, there are still some challenges about data and deep learning methods in drug discovery. Therefore, we enumerate current challenges that have been observed in the field to promote the development of new research. Here, the survey is meant to accelerate drug discovery through the sharing and comparison of deep generative models, finally reduce the cost and time with the intervention of silico models.

Molecular representation and dataset in molecular generation

Molecular representations

The past few decades have observed the arising of keen interest in computational methods especially with the emergence of deep learning. The appearance of such techniques opened the door to the computer-aided drug design. The major challenges here were how to recognize and store molecules accurately by computers and be acceptable for chemists. A flurry of molecular representations have been designed in the last few years owing to the rapid development of computers [24]. Here we introduce two common representations used in *de novo* molecular design, including SMILES and graphs.

SMILES-based representation

Sequence-based representations mainly use linear strings to express compounds and they can easily be memorized and processed by computer systems. One-dimensional linear representation currently includes SMILES and international chemical identifier (InCHI) [25]. SMILES is an ASCII string that uses a mapping algorithm from molecular graphs to text, where simplifies chemical structure with strict grammars. The example of the SMILES form of a molecule is shown in Figure 2A. The conversion from molecular structure to texts makes SMILES easy to be processed by computers, convenient for chemists and easy to use for training machine learning models [26]. The first strategy for automatic molecular design is to use SMILES-based deep generative models and convert this representation into one-hot vectors. Both the good sides, there are also the disadvantages of SMILES: (1) SMILES fails to capture the molecular structural similarity. A small change between two similar structures may cause the SMILES strings to be greatly different, which the latent space learned from generative models is not smooth. (2) The SMILES string is non-unique, a molecule can be encoded into multiple SMILES representations. These problems have been solved more or less in current work. Aiming at the shortages of canonical SMILES in the generative models, there are many studies in producing the variants of the SMILES notation and improving the models, more will be shown specifically in the section of SMILES-based models..

Graph-based representation

SMILES is generated from the graph-based representation of molecules [27]. And structural formulas are often used to represent molecules in chemistry. Thus, a more intuitive way to depict molecular structures is the molecular graph. The example of the graph form of a molecule is shown in Figure 2B. Each molecule can be represented as an undirected graph $\mathcal{G}$ where the nodes set $\mathcal{V}$ and the edge set $\mathcal{E}$ are composed of atoms v_i, and bonds (v_i, v_j) respectively. Specifically, each atom type (carbon, hydrogen, oxygen., etc.) can be encoded into T dimensional one hot vector x_i, and the bond type (single, double, triple and aromatic bond) can be represented as $y \in \{1, \ldots, Y\}$. With the hot research trend of deep learning on the graph [28, 29], training deep generative models based on the molecular graph has emerged in a matter of just a few years and now many works make this be one of the most prominent fields. The more details of graph-based works will be described below.

Dataset for drug discovery

The training of models in machine learning is based on the data, hence we focus on the datasets involved in *de novo* molecular design here. Specifically, we divide the datasets involved in the typical molecular generative models into the following categories.

The first is the comprehensive databases, which usually contain diverse information such as biological activity, chemical structure and physical properties, including ZINC [30, 31], ChEMBL [32], PubChem [33] and DrugBank [34, 35] appeared in higher frequency. In particular, the data fields of drug in DrugBank can be linked to other databases like PubChem [36]. The second one is the merged databases. These databases are chemical datasets by combining and screening existing databases not only for generating molecules, but also for the validation of various machine learning methods as the benchmark. The MOSES platform [37] screens the ZINC by some rules and divides the final data set into three groups, training set, test set and scaffold set to ensure the diversity of molecules. There are also task-specific databases, which are used in other tasks related to drug discovery, such as L1000 CMap [38] with gene expression profiles, CEPDB [39] for learning potential structures of photovoltaics and so on. The last type, chemical space datasets, contains compounds of specific atom composition in a way similar to enumerating chemical space. For instance, quantum machine (QM) [40, 41] extracted from GDB [42, 43], containing molecules composed of CHONF and their quantum chemical properties. Table 1 shows the specific description of the datasets which are commonly used for *de novo* molecular design, including the number of compounds contained in these datasets up to now, released years, links, etc.

Deep molecular generative models

In recent years, *de novo* molecular design, a concept of generating molecules with desired from scratch, can be implemented by either professional experts or machines. Due to the development of generative models, molecular generation not only decreases the searching space of chemical molecules but also time consumption for drug discovery compared with humans. Here, we overview some typical molecular generative models based on two classical representations in the following and summary the timeline of them in Figure 3.

SMILES-based models

Powerful deep learning techniques have driven the development of generative models. After training on realistic data, generative models are able to produce new synthetic data that are similar to given samples. A central question in deep generative models to be solved is how to capture the unknown data distribution and reveal the internal hidden structures. One of the ways is to learn

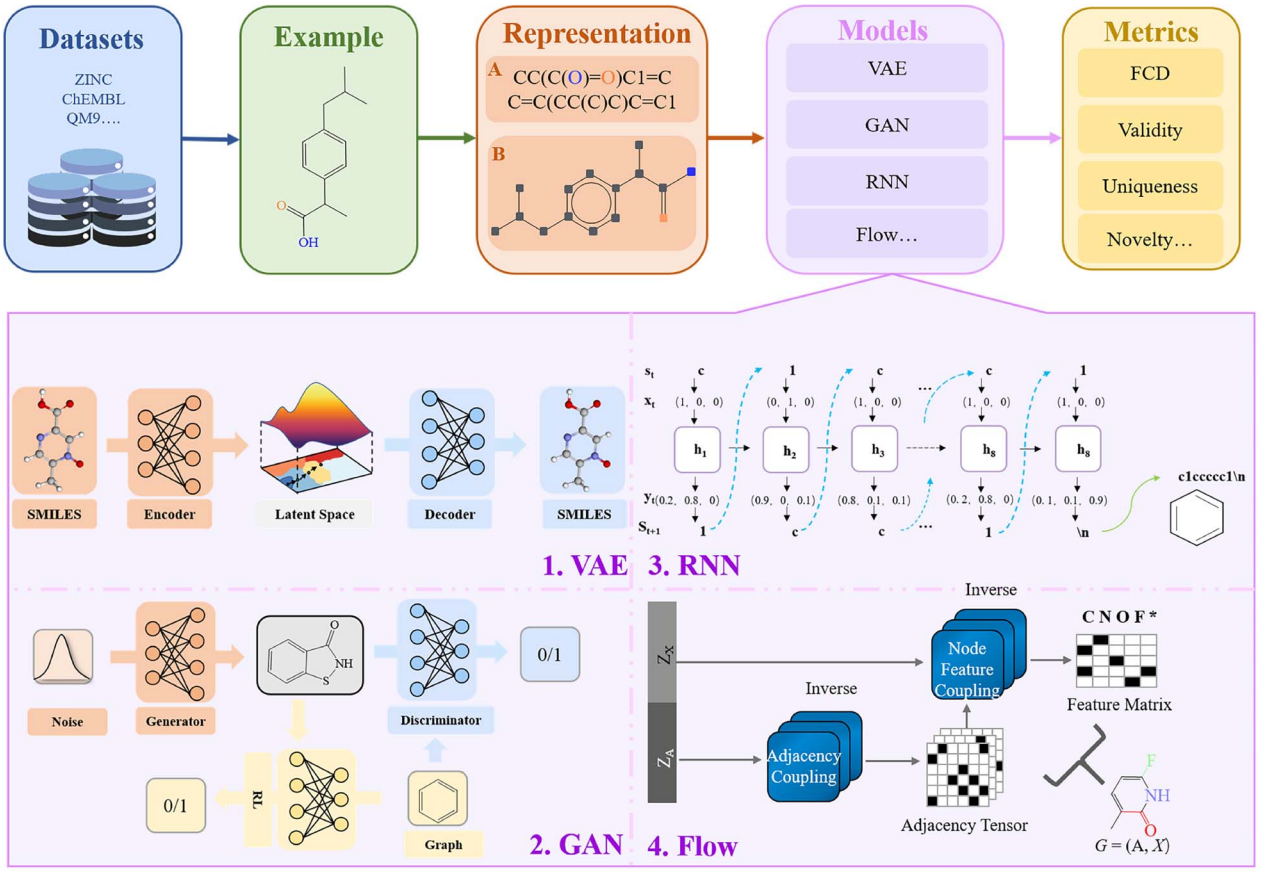

Figure 2. The process of de novo molecular design for drug discovery.

Table 1. Datasets of de novo molecular design for drug discovery

Datasets	Description	Compound	Link	Year	Reference
ZINC	A free virtual scanning dataset of commercial compounds	> 750000000	http://zinc15.docking.org/	2004	[30, 31]
ChEMBL	Manage and edit biologically active molecules with drug-like properties	1961462	http://zinc15.docking.org/	2009	[32]
QM9	Small organic molecules with maximum *nine* atoms in *four* different types	134k	http://quantum-machine.org/datasets/∖#qm9	2014	[40, 41]
PubChem	Unique structures with the largest and free chemical information	103278272	https://pubchem.ncbi.nlm.nih.gov/	2004	[33]
ExCAPE-DB	Aggregate PubChem and ChEMBL	997992	https://solr.ideaconsult.net/search/excape/	2016	[44]
GDB-13	Chemical universe database	977468314	http://gdb.unibe.ch/downloads/	2009	[42]
GDB-17	Chemical universe database	50000000	http://gdb.unibe.ch/downloads/	2012	[43]
MEGx	Natural products from plants and microorganisms	> 4200	https://ac-discovery.com/purified-natural-product-screening-compounds/	2017	/
DrugBank	With bioinformatics and cheminformatics resource	13643	https://www.drugbank.ca/	2006	[34, 35]
MOSES	A benchmarking dataset	4591276	https://github.com/molecularsets/moses	2018	[32]
CEPDB	The Harvard Clean Energy Project Database	2300000	http://cepdb.molecularspace.org	2011	[37]
L1000	Contains mainly induced gene expression profiles	32855	https://clue.io/	2017	[38]

the data representations which can be easily modeled [45]. In the field of *de novo* molecular design, a good representation also is capable of being converted back into molecules readily. In terms of the simple characteristics of SMILES, it has been proven easier to learn for deep learning. And the sequence-based methods can be further divided into variational encoder (VAE) [46], generative adversarial networks (GANs) [47] and recurrent neural networks (RNNs) [48] based models.

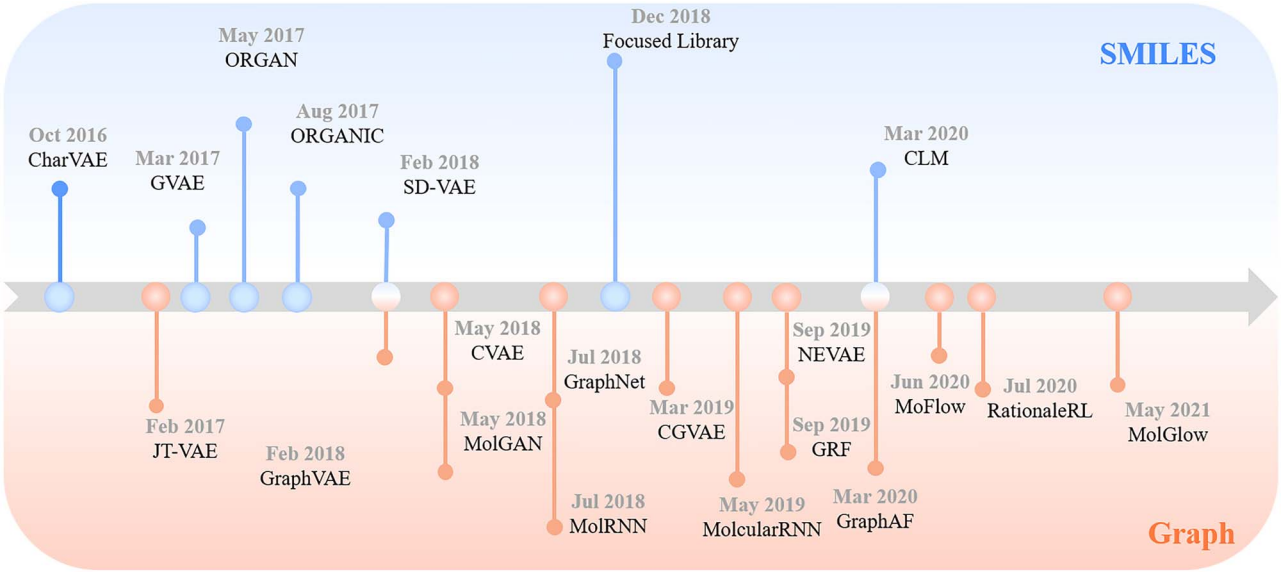

Figure 3. The timeline of deep generative models for molecular design.

- ***VAE-based generative models***

VAE generally contains an encoder and a decoder, which the encoder maps discrete data to a continuous latent space[46]. Further, in order to perform unconstrained optimization for specific properties, the decoder is responsible for reconstructing from the latent vector to SMILES with chemical validity. VAE-based models aim at maximizing the evidence lower bound (ELBO) of the likelihood with Kullback-Leibler divergence. Notably, the latent space of VAE for molecular generation is potentially operated such as controlling the specific properties and the training process is stable. However, reconstructing the training sets limits the ability of exploring in unknown chemical space.

Gómez-Bombarelli *et al.* [21] first proposed the character VAE composed of encoder, decoder and predictor (refer Figure 2.1). First, the kernel density estimation was used to learn to capture the relevant features of the molecules. Then continuous latent spaces were learned on dimensions, optimizing the specific properties of the molecules, allowed the use of powerful gradient-based to efficiently guide the search. Adding the joint training task of multi-layer perceptron and encoder guaranteed the prediction ability of molecular properties. A trick here was using Gaussian process to reach the points with target attributes. The contribution could be described as a new method for exploring the molecular space in which no prior knowledge was required to manually construct a compound library. Simultaneously, the model captured the character characteristics of the molecule and showed good predictive ability. We take the view that not all the generated one in this model can be converted back to original space owing to the non-uniqueness of SMILES. For this situation, one approach is to give the model explicit restrictions about how to produce valid molecules. For instance, GVAE [49] incorporated the grammar production rules of SMILES into models. It indicated the discrete data could be directly represented as the parse tree by using context-free grammar. The decoder generated valid outputs by learning these rules in order. Taking parse trees into account enabled the model extended to other text representation learning without context. Later, Dai *et al.* [50] argued that GVAE was lack of semantics and structural information such as the generated ring bonds must be close. However, adding the extra structural constraints in GVAE may cause the unnecessary waste of computing and time. Inspired of the attribute grammar, Dai *et al.* [50] proposed to introduce the stochastic lazy links into attribute grammars which achieved on-the-fly generated guidance for both syntax and semantics check.

- ***GANs-based generative models***

During the past 5 years, case studies using GANs towards the generation of novel molecules with specific desired properties have made milestone progress, especially the combination of GAN and reinforcement learning [51]. GAN includes a generator that imitates the real samples and a discriminator that distinguishes the output of the generator from the actual sample to the greatest extent, while the generator is a liar for the discriminator. The ultimate goal of GANs is to make the discriminator unable to judge whether the output of the generator is the fake. Due to the unstable training of GAN, some variants were proposed like wassertein GAN (WGAN) [52]. WGAN incorporates the Earth-Mover (EM) distance, which reflects the minimum cost under optimal planning to get a smoother gradient. WGAN not only alleviates the problem of unstable training, but also evaluates reliably generative models to avoid mode collapse.

Building on SeqGAN [53], called objective-reinforced generative adversarial networks (ORGAN) [51], was proposed where added the expert-based rewards under the framework of a WGAN [52]. The combined rewards from the discriminator and domain-specific objectives were extended to the training process that the generator was trained as an agent (refer Figure 2.2). There was also a penalized item which avoided mode collapse. ORGANIC [54], a promotion of ORGAN for inverse-design chemistry, implemented the molecular biased generation towards specific properties.

Besides using GANs associated with reinforcement learning, many studies followed the mix of GANs and autoencoder to alleviate the instability of GANs. For instance, Prykhodko *et al.* [55] proposed the latentGAN which combined an autoencoder and a GAN. Previous experiments showed that different randomized SMILES representations of the same molecule were encoded into identical latent vectors which moderated the overfitting due to

the canonical SMILES. Compared with the proposed GAN-based models, this architecture was comprised of a heteroencoder, that converted pairs of different representations of SMILES into latent vectors, and the decoder, the inputs of which were trained on the generator and discriminator. It is apparent that SMILES is not the input of the generator and discriminator, but instead latent vectors. Another example was the stacked GAN conditioned on the gene expressions signatures that combines with GVAE. Efficient approaches considering the ligand-target interaction to generate molecules with desired biological activity were lacking before. Lucio *et al.* [56] employed the transcriptomic data as conditions that obtained the active-like molecules for desired targets.

- ***RNNs-based generative models***

VAE models with no extra constraint have a high probability to induce invalid molecules. However, language models extract the information automatically at grammar and semantic levels. RNNs are connected models which are able to capture the dynamics of sequences via cycled units in the network of nodes. Consequently, the models can easily process the input and output that consists of sequences. In recent years, some improvement on the network architectures like long-short term memory (LSTM) [57] and gated recurrent unit (GRU) [58] have been proposed due to the difficult training of RNNs. LSTM, adding the memory cell that replaces conventional units, solves difficulties with training encountered by RNNs. And the simplicity of GRU is more suitable for building larger networks due to the smaller amount of parameters.

Since the sequence representation of SMILES, the analogy of natural language processing tasks and molecular generation is feasible [27]. For RNNs, the features obtained from large molecular datasets can be transferred to produce molecules with activity on demand in small ones, so that Segler et al. [59] generated focused molecule libraries by retraining the model (refer Figure 2.3). Sampling from the large-scale datasets ensured the diversity of molecules and fine-turning increased the focused properties. The drawback of the model was lack of interpretability. Besides, Zheng *et al.* [60] built a quasi-biogenic compound library including stereo-chemical properties. In addition to adding transfer learning, the computational model proposed by Moret *et al.* [61] called chemical language model (CLM), was used to design novel molecules in a designated area of the chemical space by combining three preceding optimized methods (data augmentation, temperature sampling and transfer learning). Linking bioactive synthetic compounds with natural products provides source of inspiration for drug discovery and the result expands the application scope of CLM in a small data regime. Notably, conditional generative models have been recommended, which utilized additional information to guide the molecular design. For example, molecular descriptors values were incorporated into the RNNs-based models [62], which were more focused that the traditional methods.

As mentioned above, the canonical SMILES form has the lower capacity to create large chemical spaces of valid and semantics structures than randomized SMILES. [63] sampled with replacement 2 billion times and explored three different variants of the SMILES notation (canonical, randomized and deep) from GDB-13 to prove that hypothesis. In addition, different cell architectures (LSTM [57], GRU [58]) and training set sizes (1000000, 10000 and 1000) were the factors that affected the performance. Experiment showed that LSTM on 1 million randomized SMILES achieved state-of-the-art performance. Moreover, traditional RNNs always generate molecules in a forward manner (left to right). Motivated by the bidirectional RNN [64], non-univocity and non-directionality of SMILES, bidirectional molecule designed by alternate learning (BIMODAL) [65] was proposed. BIMODAL referred the neural autoregressive distribution estimator (NADE) [66], which reconstructed missing information by reading the preceding and subsequent tokens in both directions, and synchronous forward backward RNN (FB-RNN) [67] which generated SMILES forward and backward. BIMODAL predicted the sequence alternately by forward at odd positions and backward at even positions. Due to the limitation of two generating directions in BIMODAL, Arós-ous *et al.* [68] reported a model that generated molecules from a given scaffold without an assigned attachment. Specifically, the model made use of a slicing algorithm to obtain scaffold sets with randomized SMILES representations. And then partially built molecules were decorated in one attachment once or more than one at a time.

Since SMILES is regarded as string of texts, a large number of models in natural language processing are able to be extended to the field of *de novo* molecular design. In future research, for example, we can consider the molecular generation for desired properties as a translation, which can translate from the specific target language (protein sequence) to the SMILES language. Notably, despite the surge of SMILES-based models in recent years, there are still some burning problems. Not only is it facing the issues of validity, but the unstructured nature of SMILES makes two similar molecules be quite different with a high probability. And it is expensive to force the validity constraint to incorporate into the decoders, which requires for designing a novel representation with more structural information.

Graph-based models

Deep molecular generative models based on graphs have been a hot trend in the graph research with a prospect for drug discovery. In recent three years, there are many surprisingly effective works in the field of molecular graph generation. Considering the success of VAE models on SMILES, architectures based on VAE with molecular graph design were later developed. Gómez-Bombarelli *et al.* [21] believed that graph-based representation methods should be further explored. Moreover, with the popularity of graph neural networks, graph-based models also play a dominate role in de novo molecular design.

- ***VAE-based generative models***

One of the most representative work is junction tree variational autoencoder (JT-VAE) [69]. JT-VAE assembled the building blocks from the substructures of molecules. The substructures included rings, functional groups and atoms by decomposing the molecules from training sets. In contrast to generating graphs node by node before, the entire process was divided into two phases, first represented the valid scaffolds and their arrangement as the trees and then integrated the whole trees into the graph by adding edges between intersecting components. JT-VAE outperformed the proposed models including CVAE [21], GVAE [49], SD-VAE [50] and GraphVAE [70] in molecular reconstruction and the octanol-water partition coefficients logP score, at mean whiles, JT-VAE reached 100% in generating valid molecules. The results positively advocated the model for graph-based de novo molecular design, with JT-VAE showing superior results to the previous methods for most of the criteria in the tested conditions. However, this design had three critical limitations. First, it was more difficult to grapple in properties optimization with JT-VAE because two molecules with identical junction tree might correspond with markedly different attributes. Second, leaving

node order permutation out of consideration during generation procedure caused time-consuming. The ultimate sequence under some possible node permutation might be mapped into the same graph. Third, less than 20 atoms in a substructure were not practical due to the complexity of drug molecules in realistic. Later, in [71], the authors regarded the molecular optimization task as graph-to-graph translation which aimed to learn a multi-model mapping between two domains.

Recently, Jin *et al.* [72] keeps on the research of molecular generation by employing substructures to enrich the properties of molecules. However, integrating multiple properties into one molecule is a challenge, there are two reasons: (1) lacking of the molecules fits all the constraints such as potency, safety and desired properties in realistic and (2) bad performance of successful ratio and novelty when adding all the property constraints. To address these drawbacks, substructure sets with respect to single property, called molecular rationale, were assembled into multi-rationale vocabularies of interest and completed by graph generation model. Four properties including c-Jun N-terminal kinase 3 (JNK3) and glycogen synthase kinase 3β (GSK-3β), quantitative estimate of drug-likeness (QED) [73] and synthetic accessibility (SA) [74], aiming at generating Alzheimers disease dual inhibitors with drug-likeness and synthetically accessibility, reached a 100% successful rate.

Correlations between edges often are decomposed into discrete sequenced representations during the generation trace. CGVAE [75] was incorporated two types of correlations including some known rules like valency rules as hard one and ring strain (adverse to cycles) as soft one. CGVAE was kept correlative to convert into SMILES with validity semantically. Notably, molecules with similar properties practically differed in the size of nodes and edges while generative models for molecular graphs like GraphVAE [70] did not take this into account. Samanta *et al.* [76] have proposed NEVAE aggregated features by distinct hops number. Concurrently, graphs generated from NEVAE took on the permutation invariance at node level and meditated the impact of spatial position on properties. Specially, NEVAE was integrated potential energy with the stability of atoms implemented by Gaussian toolbox.

- ***GANs-based generative models***

Albeit wide application of GANs in some areas, the developments of GANs in generating molecular graphs are tender and delicate. To our best knowledge, GANs are inclined to spawn mode collapse [77]. Since averting likelihood-based loss functions, GAN sends molecular optimization hard stable. It is mansized to balance adversarial training and property constraints. De Cao *et al.* [78] proposed GANs-based graph generative models, called MolGAN, utilized the likelihood free method and avoided the expensive graph matching procedure. Similar to ORGAN, MolGAN applied the policy gradient reward and WGAN with penalty gradient for small molecular graphs.

Cycle-consistent generative adversarial network (CycleGAN) [79] addressed different domains translation without paired input-output examples so that it is potential to optimize molecules from sets without desired properties. Mol-CycleGAN [80] was designed to implement structural transformations and molecular optimization under the embedding of JT-VAE. Datasets were divided into two parts, one was not equipped with target properties (namely halogen moieties, number of aromatic rings, active) for training and the other was opposite for testing. Notably, bioisosteres are groups or substituents that possess similar chemical properties and produce a wide range of similar or opposite biological activities. Bioisosteres replacement like CF3 and CN is one of the methods in modern molecular design. Specifically, the framework of Mol-CycleGAN could be described as follow: (1) datasets: sets X (e.g., with CN group) and Y (e.g., with CF3 group); (2) mapping: $G : X \rightarrow Y$ and $F : Y \rightarrow X$) cycle consistency loss: encouraged $F(G(x)) \approx x$ and $G(F(y)) \approx y$) metrics: optimized values of a given property, such as *logP*. Mol-CycleGAN has shed new light on molecular optimization under the realistic situation of devoid paired samples. Inspired by Mol-CycleGAN, we can regard molecular optimization as the problem of machine translation or graph translation.

- ***RNNs-based generative models***

Generative networks based on RNNs model the graph generation as a sequential process and make auto-regressive decisions while they generate graphs. GraphNet [81], the first RNNs-based model on arbitrary graph, was on the framework of the message-passing neural networks (MPNN) [82]. The essence of GraphNet was to add a new atom or bond into the existing graph. More concretely, (1) choosing to add an atom or not, (2) computing the probabilities over the existing graph to determine if adds a new edge, (3) calculating the probabilities which one node in graph to connect. In addition, Li *et al.* [83] explored MolMP and Mol-RNN based on graph convolutional networks (GCN) [84] which was similar with the generation of GraphNet, which generated molecules by iteratively adding nodes and edges to the existing subgraphs. Converting the extra constraints into available conditional codes that did not require reinforcement learning provided higher flexibility and outputs the molecules with more diversity.

GraphRNN [85], a hierarchical model on the graph and edge levels, aimed at capturing the joint probability of nodes and edges. The process of graph generation was regarded as sequences of adjacency vectors under different node orderings. GraphRNN was equipped with the scalability by introducing a breadth-first search (BFS) node ordering. Subsequently, Some works for molecular generation were extended from GraphRNN. For example, MolecularRNN [86] was added nodes and edges feature vectors associated with them based on GraphRNN. The model was inserted valency-based rejection sampling to ensure valid molecular with 100% rate. It incarnated the traits and advantages of merging pre-training on large datasets and tuning with policy gradient algorithm. MolecularRNN outperformed JT-VAE, GCPN and ORGAN in the penalized *logP* coefficient and QED [73] with reinforcement learning.

- ***Flow-based generative models***

Another series of methods, flow-based generative models [45, 87, 88], have been applied for image generation and have recently begun to obtain attention in the molecular generation community. With the help of normalizing flow, the flow-based generation models explicitly learn the data distribution which are consist of invertible transformations. The flow takes an initial variable as input and converts it into a variable with an isotropic Gaussian by repeatedly using the change of variable rule, which is similar to the inference procedure in an encoder of VAE [89]. Non-linear independent components estimation (NICE) [45] was the first normalizing flow architecture which showed satisfying performance on the mixed national institute of standards and technology (MNIST) database and was applied for inpainting. It just roughly stacked fully connected layers so that flow-based models needed to be explored further. During the follow-up work, RealNVP [87] and Glow [88] yielded unusually brilliant results and became strong performers in the field of generative models.

According to what we know, prior works put forward five models to generate molecular graphs, GraphNVP [90], graph residual flow (GRF) [91], GraphAF [92], MoFlow [93] and MolGrow [94] included (refer Figure 2.4). GraphNVP [90], the first flow-based molecular graph generation model, improved the uniqueness of molecules. Compared with GraphNVP, GRF [91] reached almost equivalent performance while the number of parameters was reduced. Unfortunately, those two one-shot models displayed poor performance in generating valid molecules. Enlightened by the autoregressive and few flow-based models, a flow-based autoregressive sequential model called GraphAF [92] was proposed. GraphAF outperformed the contemporary state-of-art model graph convolutional policy network (GCPN) [36] and generated 100% valid molecules by incorporating valency checking. As a one-shot manner, MoFlow [93] was broken many state-of-the-art results that generated bonds by a variant of Glow and atoms with a given bond through a new graph conditional flow. Moreover, the author proposed a new validity correction procedure by deleting the bond of the last order recursively that maintained the largest valid components. Recently, MolGrow [94] showed great results constrained optimization of properties by using latent variables of the model. MolGrow recursively splitted a node into two from a single one to generate molecular structures, which could be regarded as plug-and-play modules. And the model achieved better performance while learning on a fixed atom ordering.

Overall, based on the generation process, the existing graph-based models can roughly classify into two types, one is the sequential iterative process, the other is one-shot generation. Specifically, they can divide into atom-by-atom, subgraph-based (fragment) models. In order to reduce the number of predicting the edge and train under the possible node permutation, some models such as RationaleRL, MolecularRNN was adopted in a BFS manner. Masking in sequence-based models is introduced for maintaining local structural and functional properties like NEVAE, CGVAE. Graph-based generative models now play a dominant role in the molecular design due to the advantages of graphs and the development of graph neural networks, however, some challenges still remain. To our best knowledge, with the increases of nodes size, the total calculation will grow up the square of nodes number at least, which is difficult to acquire the precise likelihood. Hence, the problem of node ordering should be better solved, which is beneficial for generating molecules with high quality.

Challenges

Data. *De novo* molecular design is facing the common failing in artificial intelligence, including the representation, quality and scarcity of data. The training of deep neural networks always relies on sufficient data namely data-driven. Therefore, constructing more satisfying datasets in the field of molecular generation is also a hot-potato to solve. In addition, data for some desired targets are scarce. For this, some models [61] choose to pre-train on the large dataset and then fine-tuned to generate molecules for the specific targets. We take the view that incorporating multi-omics data can make up for the insufficiency of data scarcity in the future. Further, designing a representation with enriched information for molecules is also a challenge. No doubt that sequence-based representations are simpler, but they ignore the structure information to some extent. Moreover, graph-based methods have been widely used, nevertheless, incorporating 3D information into graph-based models is still lacking. Combining 3D information with appropriate structure-based models in a simple manner is the Achilles' heel and it will be an interesting venue for the future work [27]. Last but not least, learning molecules under the representation of images may be a feasible orientation due to the mature of computer vision.

Models. Most of current models for molecular generation draw lessons from existing methods in computer vision and natural language processing that do not develop novel models from the perspective of this field. While molecules imitate the representation of images and texts, the generation of images and texts is fault-tolerant. And molecules for drug discovery are extremely strict with validity. From this aspect, designing unique models and appropriate representations belongs to molecules are warranted. To some extent, such models can also be able to be extended to the problems in other fields. And in the future, we are also excited about developing a hierarchical model, which can generate molecular with desired properties in a coarse-to-fine manner. A hierarchical model is beneficial for extracting different information when incorporating multi-omics data. As early mentioned, some generative models themselves exist some challenges to face. For example, although the flow-based models reconstruct samples perfectly, the cost of computation is still not as friendly as other generative models. In light of that, reducing expensive cost of flow-based models is the next action to optimize. In addition, the explainability of generative models for molecular design is equally worth being researched.

Metrics. Most of models employ the evaluation metrics from various aspects as following. Bickerton *et al.* [73] utilized the concept of desirability called the QED to measure drug-likeness. And Fréchet ChemNet Distance (FCD) [95] is a measure of distribution between training sets and generated molecules. That logP is a particular descriptor estimates the octanol–water partition coefficient. A variant of logP, called penalized logP [49], takes synthetic accessibility and ring sizes into account as penalty. MOSES [37], a benchmarking platform, contains a standardized dataset, a set of indicators and multiple baselines for comparing molecular generation models. However, there are several tasks for which these models generate hard synthetic molecules and provide synthetic routes difficultly despite performing well on common benchmarks. And as a matter of fact, these systematic metrics are a far cry from industry to discovery drugs, namely the generated molecules do not meet the requirement for the practical use. How to balance and unify two metrics systems for discovering drug in a faster and effective fashion runs tough at present. And designing the metrics for the practical use and combining with experiments will allow a major step towards molecular generation.

Conclusion

In this review, we have done our utmost to report different stages of molecular generation evolutionary path and highlight recent advances of research. Both of sequence-based and graph-based generative models have their own merits. The way in which molecular generative models are developed plays an important role for drug discovery and mirrors the evolution of deep neural networks in cross realm. Although substantial progress has been made, there is still large room for improving the performance of existing generative models and ameliorating the metrics of synthetic accessibility. These promotions of technologies and computing power promise to further advance the qualities of generating molecules with well-designed drug-like properties and make further efforts to accelerate the de novo drug design

in a fully automated fashion. And these advances of molecular generation also herald a promising future of related problems such as retrosynthesis. With the development of friendly and easy-to-use automat tools, collaborative work of chemists and computer technicians will promote drug discovery further in the future.

Key Points

- The ability of *de novo* molecular design based on deep generative models is of great scientific interest and practical importance for drug discovery.
- There are several publicly available drug-related datasets for training deep generative models in molecular design.
- There are still challenges on generative models for de novo molecular design, including data, model design and evaluation metrics.

Funding

The work was supported by National Natural Science Foundation of China (61972138), the Fundamental Research Funds for the Central Universities (531118010355, 531118010626), Hunan Provincial Natural Science Foundation of China (2020JJ4215) and the Key Research and Development Program of Changsha (kq2004016).

References

1. Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 2010; **9**(3): 203–14.
2. Mullard A. The drug-maker's guide to the galaxy. *Nature News* 2017; **549**(7673): 445.
3. Polishchuk PG, Madzhidov TI, Varnek A. Estimation of the size of drug-like chemical space based on GDB-17 data. *J Comput Aided Mol Des* 2013; **27**(8): 675–9.
4. Hert J, Irwin JJ, Laggner C, et al. Quantifying biogenic bias in screening libraries. *Nat Chem Biol* 2009; **5**(7): 479–83.
5. Rifaioglu AS, Atas H, Martin MJ, et al. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. *Brief Bioinform* 2019; **20**(5): 1878–912.
6. Jing Y, Bian Y, Hu Z, et al. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. *AAPS J* 2018; **20**(3): 1–10.
7. Shrestha A, Mahmood A. Review of deep learning algorithms and architectures. *IEEE Access* 2019; **7**:53040–65.
8. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science* 2006; **313**(5786): 504–7.
9. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 2012; **25**:1097–105.
10. Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge. *Nature* 2017; **550**(7676): 354–9.
11. Lin J, Pang Y, Xia Y, et al. TuiGAN: Learning versatile image-to-image translation with two unpaired images. In: *European Conference on Computer Vision*. Springer, 2020, 18–35.
12. Chen X, Duan Y, Houthooft R, et al. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. Curran Associates Inc, 2016, 2180–8.
13. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc, 2017.
14. Hsu ST, Moon C, Jones P, et al. An interpretable generative adversarial approach to classification of latent entity relations in unstructured sentences. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.
15. Huang X, Qi J, Yu S, et al. Mala: Cross-domain dialogue generation with action learning. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 7977–84.
16. Sheng N, Cui H, Zhang T, et al. Attentional multi-level representation encoding based on convolutional and variance autoencoders for lncRNA–disease association prediction. *Brief Bioinform* 2021; **22**(3):bbaa067.
17. Shao H, Yao S, Sun D, et al. ControlVAE: Controllable variational autoencoder. In: *International Conference on Machine Learning*. PMLR, 2020, 8655–64.
18. Song L, Wang A, Jinsong S, et al. Structural information preserving for graph-to-text generation. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2020, 7987–98.
19. Balaji Y, Min MR, Bai B, et al. Conditional GAN with discriminative filter generation for text-to-video synthesis. In: *International Joint Conference on Artificial Intelligence*, 2019, 1995–2001.
20. Xue D, Gong Y, Yang Z, et al. Advances and challenges in deep generative models for de novo molecule generation. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2019; **9**(3):e1395.
21. Gómez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science* 2018; **4**(2): 268–76.
22. Weininger D. SMILES, a chemical language and information system. 1. introduction to methodology and encoding rules. *J Chem Inf Comput Sci* 1988; **28**(1): 31–6.
23. Xia X, Hu J, Wang Y, et al. Graph-based generative models for de novo drug design. *Drug Discov Today Technol* 2020.
24. David L, Thakkar A, Mercado R, et al. Molecular representations in AI-driven drug discovery: a review and practical guide. *J Chem* 2020; **12**(1): 1–22.
25. Heller S, McNaught A, Stein S, et al. InChI - the worldwide chemical structure identifier standard. *J Chem* 2013; **5**(1): 1–9.
26. Elton DC, Boukouvalas Z, Fuge MD, et al. Deep learning for molecular design-a review of the state of the art. *Molecular Systems Design & Engineering* 2019; **4**(4): 828–49.
27. Schwalbe-Koda D, Gómez-Bombarelli R. Generative models for automatic chemical design. In: *Machine Learning Meets Quantum Physics*. Springer, 2020, 445–67.
28. Faez F, Ommi Y, Baghshah MS, et al. Deep graph generators: A survey arXiv preprint arXiv:2012.15544. 2020.
29. Jin S, Zeng X, Xia F, et al. Application of deep learning methods in biological networks. *Brief Bioinform* 2021; **22**(2): 1902–17.
30. Irwin JJ, Sterling T, Mysinger MM, et al. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 2012; **52**(7): 1757–68.
31. Sterling T, Irwin JJ. ZINC 15–ligand discovery for everyone. *J Chem Inf Model* 2015; **55**(11): 2324–37.

32. Gaulton A, Bellis LJ, Bento AP, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 2012; **40**(D1): D1100–7.
33. Kim S, Thiessen PA, Bolton EE, et al. PubChem substance and compound databases. *Nucleic Acids Res* 2016; **44**(D1): D1202–13.
34. Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 2006; **34**(suppl_1).
35. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018; **46**(D1): D1074–82.
36. Sun M, Zhao S, Gilvary C, et al. Graph convolutional networks for computational drug development and discovery. *Brief Bioinform* 2020; **21**(3): 919–35.
37. Polykovskiy D, Zhebrak A, Sanchez-Lengeling B, et al. Molecular sets (MOSES): a benchmarking platform for molecular generation models. *Front Pharmacol* 2020; **11**.
38. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017; **171**(6): 1437–52.
39. Hachmann J, Olivares-Amaya R, Atahan-Evrenk S, et al. The harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid. *The Journal of Physical Chemistry Letters* 2011; **2**(17): 2241–51.
40. Schütt KT, Arbabzadah F, Chmiela S, et al. Quantum-chemical insights from deep tensor neural networks. *Nat Commun* 2017; **8**(1): 1–8.
41. Chmiela S, Tkatchenko A, Sauceda HE, et al. Machine learning of accurate energy-conserving molecular force fields. *Sci Adv* 2017; **3**(5):e1603015.
42. Lorenz C Blum and Jean-Louis Reymond. 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13. *J Am Chem Soc*, **131**(25): 8732–3, 2009.
43. Ruddigkeit L, Van Deursen R, Blum LC, et al. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *Journal of Chemical iInformation and Modeling* 2012; **52**(11): 2864–75.
44. Jiangming Sun, Nina Jeliazkova, Vladimir Chupakhin, Jose-Felipe Golib-Dzib, Ola Engkvist, Lars Carlsson, Jörg Wegner, Hugo Ceulemans, Ivan Georgiev, Vedrin Jeliazkov, et al. ExCAPE-DB: an integrated large scale dataset facilitating big data analysis in chemogenomics. *Journal of Cheminformatics*, **9**(1): 1–9, 2017.
45. Dinh YBL, Krueger D. NICE: non-linear independent components estimation. In: *International Conference on Learning Representations*, 2015.
46. Kingma DP, Welling M. Auto-encoding variational bayes. In: *International Conference on Learning Representations*, 2014.
47. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014.
48. Irsoy O, Cardie C. Deep recursive neural networks for compositionality in language. *Advances in Neural Information Processing Systems* 2014; **27**:2096–104.
49. Kusner MJ, Paige B, Hernández-Lobato JM. Grammar variational autoencoder. In: *International Conference on Machine Learning*. PMLR, 2017, 1945–54.
50. Dai H, Tian Y, Dai B, et al. Syntax-directed variational autoencoder for molecule generation. In: *International Conference on Learning Representations*, 2018.
51. Guimaraes GL, Sanchez-Lengeling B, Outeiral C, et al. Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models arXiv preprint arXiv:1705.10843. 2017.
52. Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: *International Conference on Machine Learning*. PMLR, 2017, 214–23.
53. Yu L, Zhang W, Wang J, et al. Sequence generative adversarial nets with policy gradient. In: *AAAI conference on Artificial Intelligence*, Vol. **490**, 2017.
54. Sanchez-Lengeling B, Outeiral C, Guimaraes GL, et al. Optimizing distributions over molecular space. An objective-reinforced generative adversarial network for inverse-design chemistry (ORGANIC) ChemRxiv. 2017;2017.
55. Prykhodko O, Johansson SV, Kotsias P-C, et al. A de novo molecular generation method using latent vector based generative adversarial network. *J Chem* 2019; **11**(1): 1–13.
56. Méndez-Lucio O, Baillif B, Clevert D-A, et al. De novo generation of hit-like molecules from gene expression signatures using artificial intelligence. *Nat Commun* 2020; **11**(1): 1–10.
57. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997; **9**(8): 1735–80.
58. Cho K, vanMerriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 2014, 1724–34. Association for Computational Linguistics.
59. Segler MHS, Kogej T, Tyrchan C, et al. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Central Science* 2018; **4**(1): 120–31.
60. Zheng S, Yan X, Qiong G, et al. QBMG: quasi-biogenic molecule generator with deep recurrent neural network. *J Chem* 2019; **11**(1): 1–12.
61. Moret M, Friedrich L, Grisoni F, et al. Generative molecular design in low data regimes. *Nature Machine Intelligence* 2020; **2**(3): 171–80.
62. Kotsias P-C, Arús-Pous J, Chen H, et al. Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. *Nature Machine Intelligence* 2020; **2**(5): 254–65.
63. Arús-Pous J, Johansson SV, Prykhodko O, et al. Randomized smiles strings improve the quality of molecular generative models. *J Chem* 2019; **11**(1): 1–13.
64. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* 1997; **45**(11): 2673–81.
65. Grisoni F, Moret M, Lingwood R, et al. Bidirectional molecule generation with recurrent neural networks. *J Chem Inf Model* 2020; **60**(3): 1175–83.
66. Berglund M, Raiko T, Honkala M, et al. Bidirectional recurrent neural networks as generative models. In: *Advances in Neural Information Processing Systems*, 2015, 856–64.
67. Mou L, Yan R, Li G, et al. Backward and forward language modeling for constrained sentence generation arXiv preprint arXiv:1512.06612. 2015.
68. Arús-Pous J, Patronov A, Bjerrum EJ, et al. SMILES-based deep generative scaffold decorator for de-novo drug design. *J Chem* 2020; **12**:1–18.
69. Jin W, Barzilay R, Jaakkola T. Junction tree variational autoencoder for molecular graph generation. In: *International Conference on Machine Learning*. PMLR, 2018, 2323–32.

70. Simonovsky M, Komodakis N. Graphvae: Towards generation of small graphs using variational autoencoders. In: *International Conference on Artificial Neural Networks*. Springer, 2018, 412–22.
71. Jin W, Yang K, Barzilay R, et al. Learning multimodal graph-to-graph translation for molecule optimization. In: *International Conference on Learning Representations*, 2019.
72. Jin W, Barzilay R, Jaakkola T. Multi-objective molecule generation using interpretable substructures. In: *International Conference on Machine Learning*. PMLR, 2020, 4849–59.
73. Bickerton GR, Paolini GV, Besnard J, et al. Quantifying the chemical beauty of drugs. *Nat Chem* 2012; **4**(2): 90–8.
74. Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Chem* 2009; **1**(1): 1–11.
75. Liu Q, Allamanis M, Brockschmidt M, et al. Constrained graph variational autoencoders for molecule design. In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Curran Associates Inc, 2018, 7806–15.
76. Samanta B, De A, Jana G, et al. NEVAE: A deep generative model for molecular graphs. In: *Journal of Machine Learning Research*, 2020.
77. Liang G, Zhou Y. A review: Generative adversarial networks. In: *2019 14th IEEE Conference on Industrial Electronics and Applications*. IEEE, 2019, 505–10.
78. De Cao N, Kipf T. MolGAN: An implicit generative model for small molecular graphs. In: *ICML 2018 workshop on Theoretical Foundations and Applications of Deep Generative Models*, 2018.
79. Zhu J-Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2017, 2223–32.
80. Maziarka Ł, Pocha A, Kaczmarczyk J, et al. Mol-CycleGAN: a generative model for molecular optimization. *J Chem* 2020; **12**(1): 1–18.
81. Li Y, Vinyals O, Dyer C, et al. Learning deep generative models of graphs. In: *International Conference on Learning Representations*, 2018.
82. Gilmer J, Schoenholz SS, Riley PF, et al. Neural message passing for quantum chemistry. In: *International Conference on Machine Learning*. PMLR, 2017, 1263–72.
83. Li Y, Zhang L, Liu Z. Multi-objective de novo drug design with conditional graph generative model. *J Chem* 2018; **10**(1): 1–24.
84. Wu Z, Ramsundar B, Feinberg EN, et al. MoleculeNet: a benchmark for molecular machine learning. *Chem Sci* 2018; **9**(2): 513–30.
85. You J, Ying R, Ren X, et al. GraphRNN: Generating realistic graphs with deep auto-regressive models. In: *International Conference on Machine Learning*. PMLR, 2018, 5708–17.
86. Popova M, Shvets M, Oliva J, et al. MolecularRNN: Generating realistic molecular graphs with optimized properties arXiv preprint arXiv:1905.13372. 2019.
87. Dinh L, Sohl-Dickstein J, Bengio S. Density estimation using real NVP. In: *International Conference on Learning Representations*, 2017.
88. Durk, Kingma P, Dhariwal P. Glow: Generative flow with invertible 1x1 convolutions. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2018.
89. Sun H, Mehta R, Zhou HH, et al. Dual-Glow: Conditional flow-based generative model for modality transfer. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, 10611–20.
90. Madhawa K, Ishiguro K, Nakago K, et al. GraphNVP: An invertible flow model for generating molecular graphs arXiv preprint arXiv:1905.11600. 2019.
91. Honda S, Akita H, Ishiguro K, et al. Graph residual flow for molecular graph generation arXiv preprint arXiv:1909.13521. 2019.
92. Shi C, Xu M, Zhu Z, et al. GraphAF: a flow-based autoregressive model for molecular graph generation. In: *International Conference on Learning Representations*, 2020.
93. Zang C, Wang F. MoFlow: an invertible flow model for generating molecular graphs. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, 617–26.
94. Kuznetsov M, Polykovskiy D. MolGrow: A graph normalizing flow for hierarchical molecular generation. *Proceedings of the AAAI Conference on Artificial Intelligence* 2021; **35**(9): 8226–34.
95. Preuer K, Renz P, Unterthiner T, et al. Fréchet ChemNet distance: A metric for generative models for molecules in drug discovery. *J Chem Inf Model* 2018; **58**(9): 1736–41.