

# VGAE-MCTS: a New Molecular Generative Model combining Variational Graph Auto-Encoder and Monte Carlo Tree Search

*Hiroaki Iwata<sup>1\*</sup>, Taichi Nakai<sup>1</sup>, Takuto Koyama<sup>1</sup>, Shigeyuki Matsumoto<sup>1</sup>, Ryosuke Kojima<sup>1</sup>, and  
Yasushi Okuno<sup>1,2\*</sup>*

1 Graduate School of Medicine, Kyoto University, 53 Shogoin-kawaharacho, Sakyo-ku, Kyoto  
606-8507, Japan

2 HPC- and AI-driven Drug Development Platform Division, RIKEN Center for  
Computational Science, Kobe 650-0047 Hyogo, Japan

\*E-mail: iwata.hiroaki.3r@kyoto-u.ac.jp; okuno.yasushi.4c@kyoto-u.ac.jp

## KEYWORDS

Molecular generation, Deep learning, Reinforcement learning, Chemical space, Physicochemical property, Drug discovery, Material design

## **Abstract**

Molecular generation is crucial for advancing drug discovery, material design, and chemical exploration. It expedites the search for new drug candidates, facilitates tailored material creation, and enhances our understanding of molecular diversity. By employing artificial intelligence techniques, such as molecular generative models based on molecular graphs, researchers have tackled the challenge of efficiently molecules with desired properties. We proposed a new molecular generative model combining deep learning and reinforcement learning evaluated the validity, novelty, and optimized physicochemical properties of the generated molecules. Importantly, the model explored uncharted regions of chemical space, allowing for the efficient discovery and design of new molecules. This innovative approach has significant potential to revolutionize drug discovery, material science, and chemical research for accelerating scientific innovation. By leveraging advanced techniques and exploring previously unexplored chemical spaces, this study offers promising prospects for the efficient discovery and design of new molecules in the field of drug development.

## Introduction

Molecular generation is highly significant for its applications in novel drug discovery, material design, and the exploration of chemical space. It enables the efficient search and identification of new drug candidates, speeding up the process of drug development<sup>1</sup>. In material design, it allows for the creation of materials with tailored properties, contributing to advancements in various industries<sup>2</sup>. Furthermore, molecular generation aids in the systematic exploration of chemical space, uncovering novel compounds with unique properties and expanding our understanding of molecular diversity<sup>3</sup>. Overall, it has the potential to revolutionize the fields of drug discovery, material science, and chemical research for accelerating scientific innovation<sup>2, 4</sup>.

Whereas it has been estimated as there are theoretically more than  $10^{60}$  small organic molecule chemical structures<sup>5</sup>, the number of molecules actually explored in drug discovery is limited to about  $10^8$  at most<sup>6</sup>. To efficiently propose new molecules with desirable physicochemical properties from a wide chemical space, an artificial intelligent (AI) technique called molecular generative model has been studied in recent years<sup>4, 7</sup>. As an input of the AI model, the chemical structure is represented in two ways: SMILES<sup>8</sup> and a molecular graph<sup>9</sup>. In general, molecular graphs are more robust and precise to represent the molecular features than SMILES, because graph representations can capture the molecular similarity and can consider chemical checks, such as protecting the number of valence electrons, unlike SMILES representations<sup>10, 11</sup>. According to these advantages, much cheminformatics research has actually reported to work well with chemical structures represented as molecular graphs<sup>10-12</sup>. We focused on a graph representation for molecules in this study.

The two main learning methods for molecular generative models are deep learning and reinforcement learning<sup>10, 13-15</sup>. Since deep learning-based models, which learn molecular features

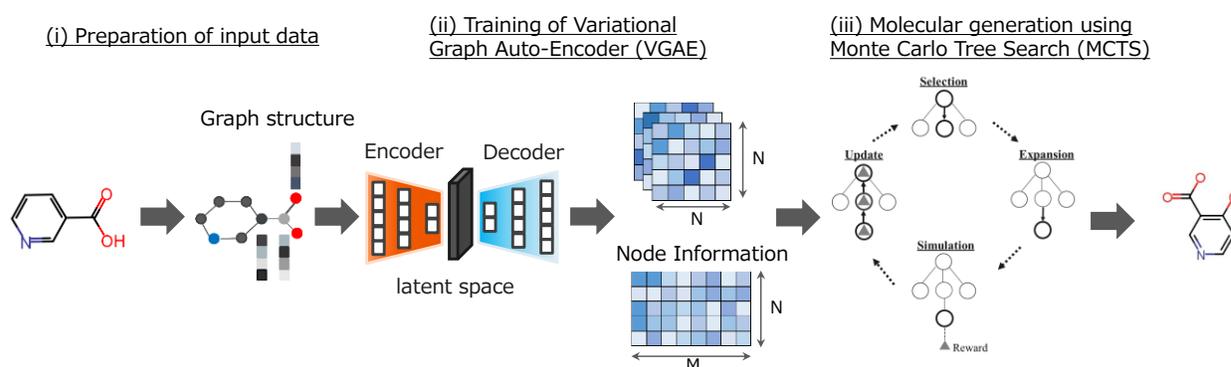
of known compounds<sup>10, 13</sup>, tend to generate molecules similar to the learned compounds, ability of generating structurally new compounds is fundamentally limited<sup>16</sup>. On the other hand, reinforcement learning-based models, which learn molecular features from scratch without prior learning of known compounds, is superior in generating molecules with structures distinct from known compounds<sup>14, 15</sup>. However, the generated molecules fundamentally lack drug-like properties due to the algorithm exploring chemical space distinct from the existing compounds. The molecules produced by the two types of the molecular generative model would be located far apart from each other in the chemical space, indicating that there is a region beyond the reach of exploration between the two groups of molecules.

In this study, we proposed a new molecular generative model that can explore chemical spaces unreachable by previous molecular generative models and discover new molecules with drug-like properties by combining deep learning and reinforcement learning based on a molecular graph representation. Specifically, the proposed method uses chemical features which learned physicochemical properties of known compounds using the Variational Graph Auto-Encoder (VGAE)<sup>17</sup> and generates molecules with desirable properties through reinforce learning with Monte Carlo Tree Search (MCTS)<sup>18</sup>. Evaluation of the generated molecules demonstrated that the validity and novelty of the chemical and the optimization of physicochemical properties was equivalent to or better than the previous methods. Furthermore, investigating the chemical structure diversity showed that the generated molecules are distributed in chemical space that was not well explored by the previous methods. The proposed method is expected to be useful for efficiently discovering and designing new molecules in the drug development.

## Results

### Proposed method for molecular design

We have developed a new molecular design model that combines a deep learning model, Variational Graph Auto-Encoder (VGAE), and a reinforcement learning model, Monte Carlo Tree Search (MCTS). Our developed model (called VGAE-MCTS) is divided into three parts: a part for preparation of input data, a part for training of VGAE, and a part for molecular generation using MCTS (Figure 1). Details of each part are described in Materials and Methods section.



**Figure 1. Workflow of our proposed method (VGAE-MCTS).** VGAE-MCTS is consisted of three parts: (i) Converting the molecules of the training data into feature maps (preparation of input data), (ii) Learning the distribution of molecules in the training data using VGAE (training of Variational Graph Auto-Encoder (VGAE)), and (iii) Generating molecules by connecting atoms and bonds one by one based on the feature map output from the learned VGAE decoder using MCTS (molecular generation using Monte Carlo Tree Search (MCTS)).

### Proposed method for molecular design

The basic performance of molecular generation using VGAE-MCTS, namely validity, uniqueness, novelty, Kullback-Leibler divergence (KL divergence), and Fréchet ChemNet Distance (FCD) was evaluated with Distribution-Learning Benchmarks in the GuacaMol framework<sup>19</sup> (Table 1). We compared the performance of VGAE-MCTS among the previous models that are Graph MCTS<sup>20</sup> and VGAE<sup>21</sup>.

The molecules generated by VGAE-MCTS showed scores of 1.000 for validity, uniqueness, and novelty. In other words, all molecules generated were valence electron counts protected, no duplications, and novel molecules that were not present in the training dataset. The results for these three types of scores were comparable or better than the previous models. The KL divergence score for VGAE-MCTS was 0.659. This is the highest result compared to the previous models, Graph MCTS and VGAE. The details of KL divergence scores for VGAE-MCTS are shown in Table S1. The FCD score for VGAE-MCTS was 0.009. FCD compares the similarity of the distribution of predicted bioactivity values between the generated molecules and compounds from the ChEMBL database. Similar to the previous models, VGAE-MCTS also had a low FCD score.

Table 1. Benchmarking Results using GuacaMol distribution-learning benchmarks

	GraphMCTS*	VGAE*	VGAE-MCTS
validity	1.000	0.830	1.000
uniqueness	1.000	0.944	1.000
novelty	0.994	1.000	1.000
KL divergence	0.522	0.554	0.659
FCD	0.015	0.016	0.009

\* Values of GraphMCTS and VGAE are taken from Table 4 in O. Mahmood, et al.<sup>22</sup>

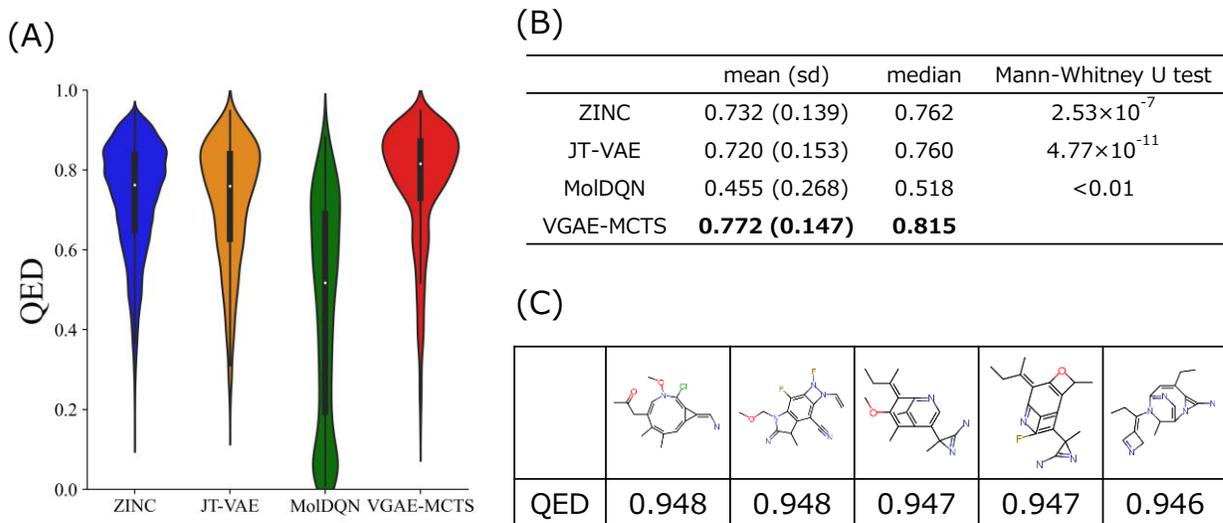
### **Optimizing physicochemical properties of drug discovery**

The ability of VGAE-MCTS to generate molecules was evaluated when the physicochemical properties were optimized. The physicochemical properties to be optimized are the Quantitative Estimate of Drug-likeness (QED)<sup>23</sup> and penalized logP. QED is a quantitative measure of drug-likeness and an evaluation of the ability to optimize single properties. Penalized logP is an index that combines three physicochemical properties: liposolubility, synthetic accessibility score, and penalty for large rings, and is an evaluation of the ability to optimize multi-properties. Both indices are commonly used physicochemical properties of drug discovery in the evaluation of molecular generation models<sup>11, 14</sup>. These indices range from 0 to 1, with molecules closer to 1 indicating that they are better molecules for drug discovery. We compared the performance of VGAE-MCTS among the prior study models that are JT-VAE<sup>10</sup> and MolDQN<sup>14</sup>.

### **QED optimization**

Molecules generated by optimizing QED score by VGAE-MCTS were compared with molecules from the ZINC dataset used as training data. In addition, comparisons were also performed with molecules generated by JT-VAE and MolDQN. The distribution of QED scores for each method is shown in Figure 2(A), and the mean, standard variance, and median statistics were shown in Figure 2(B). Examples of molecules generated by VGAE-MCTS were also shown in Figure 2(C). The QED scores of the molecules generated by VGAE-MCTS (mean: 0.772, median: 0.815) was clearly higher than those in the ZINC dataset (mean: 0.732, median: 0.762) (Mann-Whitney U test:  $P=2.53 \times 10^{-7}$ ). This result suggests that VGAE-MCTS is able to expand

molecules toward better QED that is a physicochemical property in MCTS-based search. VGAE-MCTS was also able to generate higher QED scoring molecules compared to previous methods, JT-VAE (mean: 0.720, median: 0.760) and MolDQN (mean: 0.455, median: 0.518) (Mann-Whitney U test:  $P=4.77\times 10^{-11}$ ,  $P<0.01$ ).

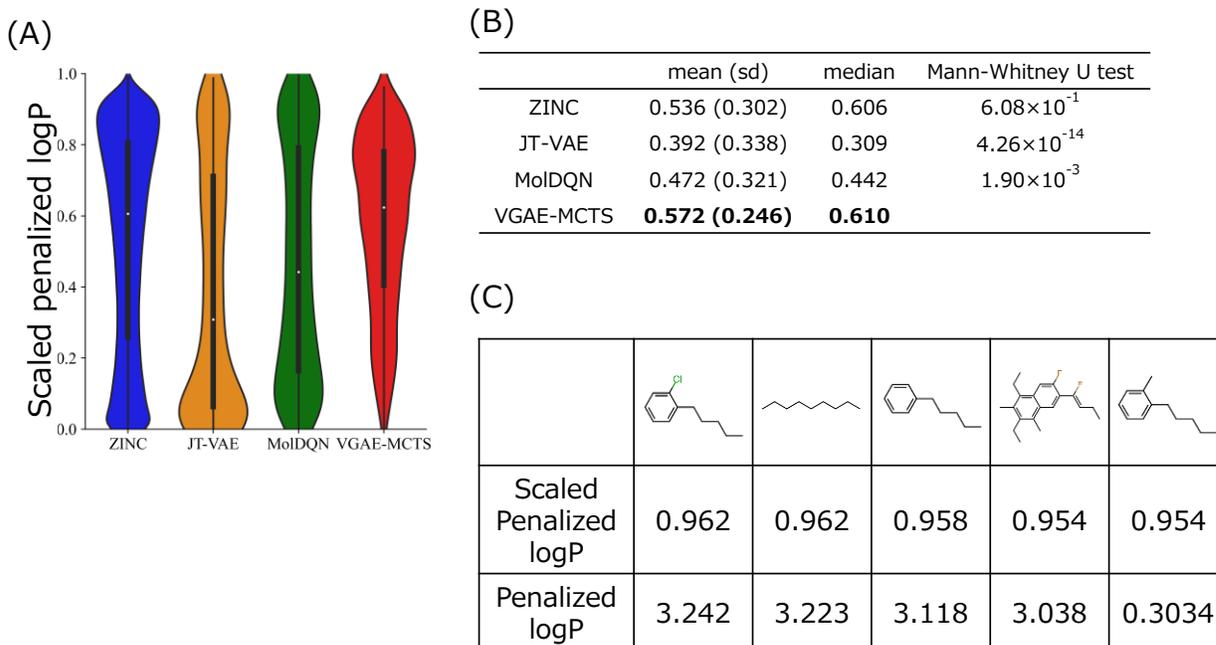


**Figure 2. Results of QED-optimized generated molecules.** (A) The vertical axis is the QED value from 0 to 1, the horizontal axis is the molecules of the ZINC dataset, VGAE-MCTS, and previous models. The white dots represent the mean values, and the bulge represents the density. (B) The mean (standard deviation), and median QED values for each molecular group are shown. (C) Chemical structures of the top 5 molecules generated by VGAE-MCTS and their QED values are displayed.

### Penalized logP optimization

As with the QED optimization, we compared the molecules generated by optimizing penalized logP using VGAE-MCTS with the molecules in the ZINC dataset and molecules generated using

JT-VAE and MolDQN (Figure 3(A), (B), and (C)). The penalized logP of the molecules generated by VGAE-MCTS (mean: 0.536, median: 0.606) was higher than those in the ZINC dataset (mean: 0.572, median: 0.610) (Mann-Whitney U test:  $P=6.08\times 10^{-1}$ ). We found that the molecules generated by VGAE-MCTS had a smaller percentage of low penalized logP values than the molecules in the ZINC dataset and the molecules generated by the previous models. In other words, this suggests that VGAE-MCTS avoids expanding molecules toward the lower penalized logP in the molecular generation using MCTS. VGAE-MCTS was also able to generate molecules with higher penalized logP compared to previous methods, JT-VAE (mean: 0.392, median: 0.309) and MolDQN (mean: 0.472, median: 0.442) (Mann-Whitney U test:  $P=4.26\times 10^{-14}$ ,  $P=1.90\times 10^{-3}$ ).

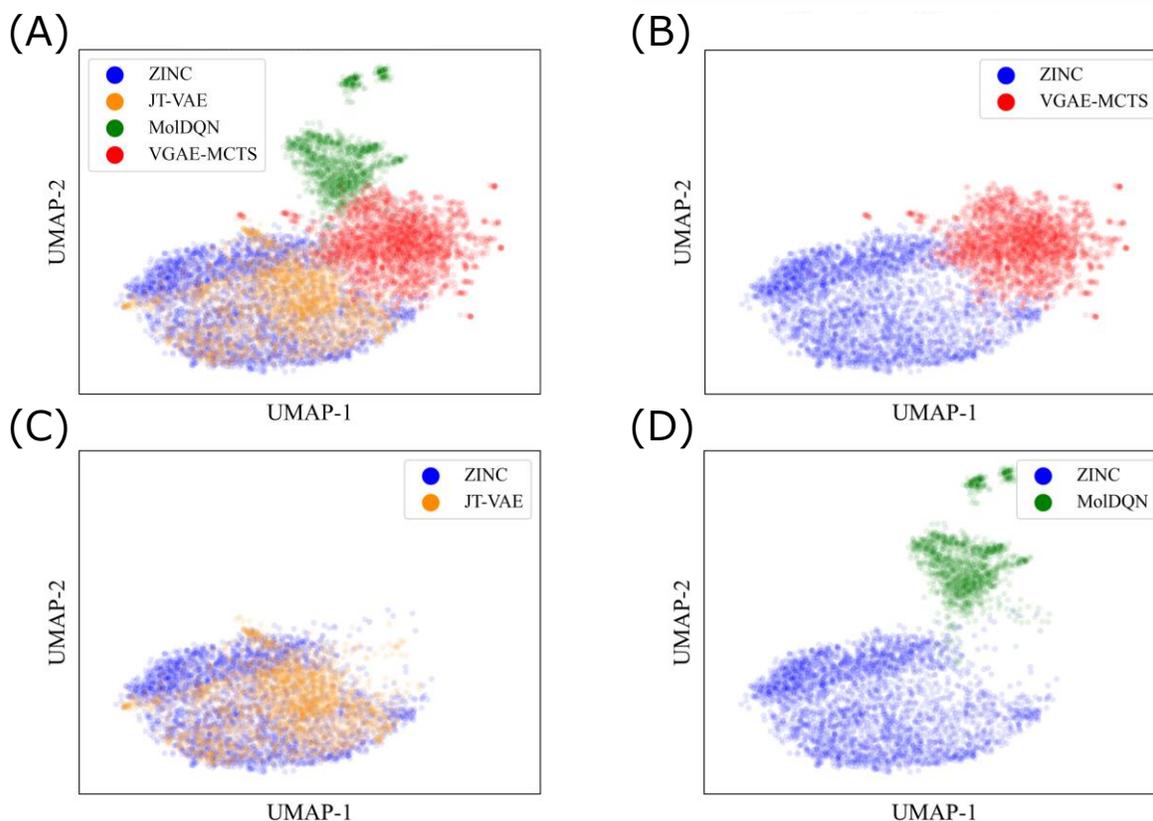


**Figure 3. Results of penalized logP-optimized generated molecules.** (A) The vertical axis is the scaled penalized logP value from 0 to 1, the horizontal axis is the molecules of the ZINC dataset, VGAE-MCTS, and previous models. The white dots represent the mean

scaled penalized logP values and the bulge represents the density. (B) The mean (standard deviation), and median scaled penalized logP value for each molecular group are shown. (C) Chemical structures of the top 5 molecules generated by VGAE-MCTS and their scaled penalized logP and penalized logP values are displayed.

### **Visualizing chemical space of generated molecules**

We evaluated whether the molecules generated by VGAE-MCTS were able to expand the chemical space from the training molecules. Specifically, molecules generated by each of the QED-optimized models (JT-VAE, MolDQN, and VGAE-MCTS) and molecules from the ZINC data set were mapped onto the chemical space (Figure 4(A)). First, the molecules generated by VGAE-MCTS were plotted in a slightly different chemical space than the molecules in the ZINC dataset (Figure 4(B)). In other words, the molecules generated by VGAE-MCTS have a new chemical structure that is slightly different from the training molecules. On the other hand, the molecules generated by JT-VAE were plotted in almost the same chemical space as the molecules in the ZINC dataset (Figure 4(C)). In other words, the molecules generated were very chemically similar to those in the ZINC data set. Molecules generated by MolDQN, which is trained without using the dataset, were plotted in a chemical space that was significantly different from the molecules in the ZINC dataset (Figure 4(D)).



**Figure 4. Visualization of QED-optimized generated molecules.** Molecules generated by optimizing QED are plotted in two dimensions using ECFP descriptors. Molecules from the ZINC training data are shown in blue. The molecules generated by JT-VAE are shown in orange, MolDQN in green, and VGAE-MCTS in red. (A) Distribution of molecules in ZINC training data and molecules generated by the three methods. (B) Distribution of molecules in ZINC training data and molecules generated by VGAE-MCTS. (C) Distribution of molecules in ZINC training data and molecules generated by JT-VAE. (D) Distribution of molecules in ZINC training data and molecules generated by MolDQN.

## Discussion

In this study, our proposed molecular generation model, VGAE-MCTS, was developed by combining VGAE, a deep learning model, and MCTS, a reinforcement learning model, to be able to explore chemical spaces that could not be explored by previous models.

First, the basic performance of VGAE-MCTS in generating molecules was evaluated with the Distribution-Learning Benchmarks in the GuacaMol framework. The molecules generated by VGAE-MCTS had a validity of 100%. This result is due to the fact that the chemical structure is represented as a molecular graph and the MCTS creates molecules by connecting atoms and bonds while protecting the number of valence electrons. The uniqueness and novelty scores of VGAE-MCTS were also higher. These results are thought to be due to the fact that VGAE-MCTS is able to output a wide variety of molecules because the more atoms that make up a molecule, the more molecules are candidates for expansion, and the type of atoms selected is stochastic.

Next, molecular generation was performed using VGAE-MCTS to optimize for each of the two types of physicochemical properties values, QED and penalized logP, and in both cases, the accuracy was confirmed to be equal to or better than that of previous studies. The molecules generated by QED optimization are more drug-like than those generated by the models in the previous methods, indicating that VGAE-MCTS is a valuable method for use in drug discovery. Penalized logP is composed of a combination of the three physicochemical properties of logP, SA score, and RingPenalty, and multi-property optimization was relatively successful in VGAE-MCTS. This suggests that VGAE-MCTS can be used to search for molecules considering multiple physicochemical properties. VGAE-MCTS can be expected to be used in practical drug discovery process where multiple conditions are optimized.

Finally, we evaluated whether the molecules generated by VGAE-MCTS were able to expand the chemical space from the training data. Because the ZINC dataset is registered for drug-like compounds, many molecules have a large QED, a quantitative measure of drug-likeness (approximately 92% of the molecules in the ZINC have a  $QED \geq 0.5$ ). Therefore, we can evaluate whether the generated molecules have expanded their chemical space using the chemical space of drug-like molecules in the ZINC as a reference. Figure 4 shows that the molecules generated by VGAE-MCTS were different in structure from those in the ZINC data set. In other words, the molecules generated by VGAE-MCTS showed a chemical spatial spread in the form of derivatives from molecules in the ZINC data set. On the other hand, molecules generated by the deep learning-based JT-VAE showed little chemical spatial spread. The molecules generated by the reinforcement learning-based MolDQN were found to be located in a completely different chemical space than the molecules in ZINC data set, i.e., they were not drug-like molecules. The above confirms that the molecules generated by VGAE-MCTS were located in a different part of the chemical space than those generated by other methods, and that the molecules generated by VGAE-MCTS were in a part of the chemical space that has not been found in previous methods.

We developed a new molecular generation model, VGAE-MCTS, which combines VGAE, a deep learning model, and MCTS, a reinforcement learning model, to explore chemical spaces that could not be explored by the models in previous studies. VGAE-MCTS showed comparable or better performance than existing models in the GuacaMol benchmark. We also showed that the performance of the optimization of the physicochemical properties, QED and penalized logP, is comparable or better than previous studies. In addition, to assess the diversity of chemical structures generated, we evaluated the distribution of molecules generated by VGAE-MCTS and

several previous models in chemical space. The results indicate that the molecules generated by VGAE-MCTS are distributed in areas that were not well explored by the molecules generated by the previous models. Based on these results, it is expected that our proposed VGAE-MCTS will be able to propose molecules that may have been out of the scope of exploration so far, which will be useful for drug development.

## Materials and Methods

### Data

Two compound datasets were prepared for the training of VGAE in the proposed method. The first dataset was compounds obtained from ChEMBL<sup>24</sup> for the evaluation of basic molecular generating capability (validity, uniqueness, novelty, KL divergence, and FCD) using GuacaMol's Distribution-Learning Benchmarks<sup>19</sup>. The total number of compounds obtained from ChEMBL was 1,352,672, which were divided into 1,273,104 for training and 79,568 for validation. As the second dataset, compounds were obtained from ZINC<sup>25</sup>, where drug-like compounds are registered, in order to evaluate the capability to optimize the physicochemical properties of molecular generation. The number of compounds obtained from ZINC was 249,456, which were divided into 199,565 for training and 49,891 for validation.

### Proposed Methods

**Preparation of input data.** The molecules of the training dataset are represented in a feature map, which is the data format for input to VGAE. The graph structure representation of a molecule is to convert it into a vector, with atoms represented by nodes and bonds by edges. In the conversion of the molecules to vector representation, node features and edge features were computed using RDKit<sup>26</sup>. Details of node and edge features are shown in Tables S2 and S3, respectively. The features of the nodes were concatenated to create a feature map of the nodes. The features of the edges were also concatenated to create a feature map of the edges.

**Training of Variational Graph Auto-Encoder (VGAE).** VGAE was used to generate feature maps for use in MCTS. VGAE consists of an encoder part (Figure S1 left) and a decoder part (Figure S1 right)<sup>27</sup>. The loss function of the VGAE consists of a regularization term calculated

by the Kullback-Leibler (KL) divergence between the normal distribution with mean 0 and variance 1 and the distribution of the encoder's output, and the sum of the reconstruction error calculated based on the input data and the data output by the decoder. For the training of VGAE in our proposed method, a latent space of 64 dimensions, a learning rate of 0.001, and a batch size of 64 were used.

**Molecular generation using Monte Carlo Tree Search (MCTS).** Molecules are generated by connecting atoms and bonds one by one in MCTS based on the feature map output by the learned VGAE. Latent variables are randomly selected from the latent space of the learned VGAE. The selected latent variables are passed through a decoder to output a feature map of edges, which represents the probability of existence of atom-atom edges, and a feature map of nodes, which represents the features of atoms.

MCTS generates molecules using the feature maps output from the trained VGAE. In MCTS, the following 1) Selection, 2) Expansion, 3) Simulation, and 4) Update are considered one search and repeated for the number of times specified by the user. When the search is completed for the number of times specified by the user on one feature map, the molecule with the best physicochemical property value at each depth of MCTS is output for numerators below the user-specified depth (`minimum_depth`). Then, the search is moves on to a next feature map.

- 1) Selection: Select one node that has the smallest value in the following equation.

$$-\frac{s}{n} - c \sqrt{\frac{\ln N}{n}} \quad \cdot \cdot \cdot (4)$$

where  $s$  is the score of the node,  $n$  is the number of times the node has been visited,  $c$  is the search coefficient ( $c = 1.5$  in our case), and  $N$  is the number of times the parent node has been visited. Equation (4) corresponds to Equation of the Upper Confidence Bound 1

(UCB1) <sup>28</sup>, which is well known in reinforcement learning. At this time, the depth is increased by one with the selected node.

- 2) Expansion: Bonding and atom addition are performed based on the candidate edges for the selected node (molecule).
- 3) Simulation: Roll out the molecules to which bonds and atoms were added in the Expansion part.
- 4) Update: The node is updated with a reward based on the physical properties of the molecule after the rollout.

The threshold and the number of searching for candidate edge extraction for this model were set to 0.10 and 8,000. For `minimum_depth`, we set it to 21 for the GuacaMol benchmark measurement, 17 for the QED optimization, and 6 for the penalized logP optimization.

In our model, “aromatic force cycle mode” is introduced to facilitate the formation of aromatic rings, which are important for drug discovery. The “aromatic force cycle mode” has the following procedures 1) to 3).

- 1) The feature map output by the VGAE is used to determine if it is possible to form aromatic rings of the specified size. In this study, aromatic ring sizes were set to 5- and 6-membered rings.
- 2) If it is determined that aromatic rings can be formed, aromatic rings are generated at the beginning of MCTS.
- 3) Increase the reward value of nodes for molecules with aromatic rings by 0.5 to make them more likely to be selected than nodes for molecules without aromatic rings during MCTS Selection.

In addition, to generate realistic molecules, the proposed model introduces two filters, a Steric strain filter<sup>11</sup> and a filter to make it difficult to create a ring structures larger than 7-membered rings. If a node was trapped by at least one of these two filters, our method made it less likely to be selected as a node to be searched by increasing the MCTS reward value by a factor of 10.

### **Performance evaluation of molecular generation**

**GuacaMol benchmarks.** Distribution-Learning Benchmarks within the GuacaMol framework were used to evaluate five indicators: validity, uniqueness, novelty, Kullback-Leibler (KL) divergence, Fréchet ChemNet Distance (FCD). The scores all range from 0 to 1, with the better the score, the closer the value to 1.

**Optimization of physicochemical property.** The Quantitative Estimate of Drug-likeness (QED)<sup>23</sup> and penalized logP were set as the physicochemical properties to be optimized. QED is a quantitative measure of drug-likeness<sup>23</sup>. QED is a quantitative measure of drug-likeness and ranges from 0 to 1, with values closer to 1 indicating that the molecule is more drug-like. When optimizing QED, the  $1 - QED$  score was used as the reward function for MCTS.

Penalized logP is a measure that combines three physicochemical properties: lipophilicity (logP), ease of synthesis (SA score), and penalty for large rings (RingPenalty). The formula used in the penalized logP optimization is defined below<sup>10,29</sup>.

$$\text{penalized logP} = \text{logP}(m) - \text{SA}(m) - \text{cycle}(m) \quad (5)$$

$$\text{Score} = \text{sigmoid}(\text{penalized logP}) \quad (6)$$

where  $m$  denotes the numerator. Using equation (6), penalized logP was converted to a range of 0 to 1 to be the score for penalized logP optimization; molecules closer to 1 indicate better

molecules. When optimizing penalized logP, the  $1 - \text{sigmoid}(\text{penalized logP})$  score was used as the reward function for MCTS.

In the evaluation of the optimization of the physicochemical properties, 3,000 molecules were randomly selected from the ZINC data set and 3,000 molecules were selected in the order in which they were generated by each molecule generation model. The distribution of physicochemical properties for the generated molecules was then calculated and evaluated.

**Statistical analysis.** The Mann-Whitney U test<sup>30</sup> was used to test for differences in the distribution of physicochemical property values between methods for molecules generated by optimizing QED and penalized logP. In addition, the significance probability p-values were corrected by Bonferroni's correction<sup>31</sup>. The molecules used for the test were 500 randomly selected from the 3,000 molecules generated for each method.

**Visualizing chemical space.** Molecules from the ZINC dataset used as training data, and molecules generated by optimizing QED with each method, were mapped to chemical space. Molecules were mapped in two dimensional space by calculating 2048 dimensional ECFP descriptors<sup>32</sup> with a diameter of 4 in RDKit, and then performing dimensionality reduction using UMAP<sup>33</sup> in ChemPlot<sup>34</sup>.

## **Supporting Information.**

Table S1. The detail of KL divergence; Table S2. Node features of a graph representation; Table S2. Edge features of a graph representation; Figure S1. Model structures of Encoder and Decoder. Encoder: Considering the entire molecular structure by the GCN model, feature maps are represented by compressing information as low-dimensional vectors and embedding them in the latent space. Decoder: From the distribution in the latent space, the information of the molecules compressed into a low-dimensional vector is recovered in the form of a feature map.

## AUTHOR INFORMATION

### **Corresponding Authors**

\*Hiroaki Iwata - Graduate School of Medicine, Kyoto University, 53 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan. E-mail: [iwata.hiroaki.3r@kyoto-u.ac.jp](mailto:iwata.hiroaki.3r@kyoto-u.ac.jp)

\*Yasushi Okuno - Graduate School of Medicine, Kyoto University, 53 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan. E-mail: [okuno.yasushi.4c@kyoto-u.ac.jp](mailto:okuno.yasushi.4c@kyoto-u.ac.jp)

### **Author Contributions**

H.I. and T.N. contributed equally to this work.

### **Funding Sources**

This research was conducted in “Development of a Next-generation Drug Discovery AI through Industry-academia Collaboration (DAIIA)” supported by Japan Agency for Medical Research and Development (AMED) under Grant Number JP22nk0101111.

## Data and Software Availability

The data that support the findings of this study are available on request from the corresponding author.

## REFERENCES

- (1) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **2016**, *47*, 20-33. DOI: 10.1016/j.jhealeco.2016.01.012 From NLM Medline. Vijayan, R. S. K.; Kihlberg, J.; Cross, J. B.; Poongavanam, V. Enhancing preclinical drug discovery with artificial intelligence. *Drug discovery today* **2022**, *27* (4), 967-984. DOI: 10.1016/j.drudis.2021.11.023 From NLM Medline.
- (2) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559* (7715), 547-555. DOI: 10.1038/s41586-018-0337-2 From NLM PubMed-not-MEDLINE.
- (3) Dobson, C. M. Chemical space and biology. *Nature* **2004**, *432* (7019), 824-828.
- (4) Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **2018**, *361* (6400), 360-365.
- (5) Schneider, G.; Fechner, U. Computer-based de novo design of drug-like molecules. *Nature reviews. Drug discovery* **2005**, *4* (8), 649-663. DOI: 10.1038/nrd1799.
- (6) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound databases. *Nucleic Acids Res.* **2016**, *44* (D1), D1202-1213. DOI: 10.1093/nar/gkv951. Walters, W. P. Virtual Chemical Libraries. *J. Med. Chem.* **2019**, *62* (3), 1116-1124. DOI: 10.1021/acs.jmedchem.8b01048 From NLM Medline.
- (7) Meyers, J.; Fabian, B.; Brown, N. De novo molecular design and generative models. *Drug discovery today* **2021**, *26* (11), 2707-2715. DOI: 10.1016/j.drudis.2021.05.019 From NLM Medline.
- (8) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31-36.
- (9) Kireev, D. B. Chemnet: A novel neural network based method for graph/property mapping. *J. Chem. Inf. Comput. Sci.* **1995**, *35* (2), 175-180. Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. A neural device for searching direct correlations between structures and properties of chemical compounds. *J. Chem. Inf. Comput. Sci.* **1997**, *37* (4), 715-721.
- (10) Jin, W.; Barzilay, R.; Jaakkola, T. Junction Tree Variational Autoencoder for Molecular Graph Generation. In Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research; 2018.
- (11) You, J.; Liu, B.; Ying, Z.; Pande, V.; Leskovec, J. Graph convolutional policy network for goal-directed molecular graph generation. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.
- (12) Kearnes, S.; McCloskey, K.; Berndl, M.; Pande, V.; Riley, P. Molecular graph convolutions: moving beyond fingerprints. *J. Comput. Aided Mol. Des.* **2016**, *30*, 595-608. Duvenaud, D. K.; Maclaurin, D.; Iparraguirre, J.; Bombarell, R.; Hirzel, T.; Aspuru-Guzik, A.; Adams, R. P. Convolutional networks on graphs for learning molecular fingerprints. *Adv. Neural*

- Inf. Process. Syst.* **2015**, 28. Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.; Pande, V. MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.* **2018**, 9 (2), 513-530. DOI: 10.1039/c7sc02664a.
- (13) Guimaraes, G. L.; Sanchez-Lengeling, B.; Outeiral, C.; Farias, P. L. C.; Aspuru-Guzik, A. Objective-reinforced generative adversarial networks (organ) for sequence generation models. *arXiv preprint arXiv:1705.10843* **2017**.
- (14) Zhou, Z.; Kearnes, S.; Li, L.; Zare, R. N.; Riley, P. Optimization of Molecules via Deep Reinforcement Learning. *Sci. Rep.* **2019**, 9 (1), 10752. DOI: 10.1038/s41598-019-47148-x From NLM PubMed-not-MEDLINE.
- (15) Rajasekar, A. A.; Raman, K.; Ravindran, B. Goal directed molecule generation using monte carlo tree search. *arXiv preprint arXiv:2010.16399* **2020**.
- (16) Santana, M. V. S.; Silva-Jr, F. P. De novo design and bioactivity prediction of SARS-CoV-2 main protease inhibitors using recurrent neural network-based transfer learning. *BMC Chem* **2021**, 15 (1), 8. DOI: 10.1186/s13065-021-00737-2 From NLM PubMed-not-MEDLINE. Wang, W.; Wang, Y.; Zhao, H.; Sciabola, S. A Transformer-based Generative Model for De Novo Molecular Design. *arXiv preprint arXiv:2210.08749* **2022**.
- (17) Kipf, T. N.; Welling, M. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308* **2016**.
- (18) Browne, C. B.; Powley, E.; Whitehouse, D.; Lucas, S. M.; Cowling, P. I.; Rohlfshagen, P.; Tavener, S.; Perez, D.; Samothrakis, S.; Colton, S. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games* **2012**, 4 (1), 1-43.
- (19) Brown, N.; Fiscato, M.; Segler, M. H. S.; Vaucher, A. C. GuacaMol: Benchmarking Models for de Novo Molecular Design. *J. Chem. Inf. Model.* **2019**, 59 (3), 1096-1108. DOI: 10.1021/acs.jcim.8b00839 From NLM Medline.
- (20) Jensen, J. H. A graph-based genetic algorithm and generative model/Monte Carlo tree search for the exploration of chemical space. *Chem. Sci.* **2019**, 10 (12), 3567-3572. DOI: 10.1039/c8sc05372c From NLM PubMed-not-MEDLINE.
- (21) Kwon, Y.; Yoo, J.; Choi, Y. S.; Son, W. J.; Lee, D.; Kang, S. Efficient learning of non-autoregressive graph variational autoencoders for molecular graph generation. *J. Cheminform.* **2019**, 11 (1), 70. DOI: 10.1186/s13321-019-0396-x From NLM PubMed-not-MEDLINE.
- (22) Mahmood, O.; Mansimov, E.; Bonneau, R.; Cho, K. Masked graph modeling for molecule generation. *Nat Commun* **2021**, 12 (1), 3156. DOI: 10.1038/s41467-021-23415-2 From NLM PubMed-not-MEDLINE.
- (23) Bickerton, G. R.; Paolini, G. V.; Besnard, J.; Muresan, S.; Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nat. Chem.* **2012**, 4 (2), 90-98. DOI: 10.1038/nchem.1243 From NLM Medline.
- (24) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magarinos, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Maranon, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **2019**, 47 (D1), D930-D940. DOI: 10.1093/nar/gky1075 From NLM Medline.
- (25) Sterling, T.; Irwin, J. J. ZINC 15--Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, 55 (11), 2324-2337. DOI: 10.1021/acs.jcim.5b00559 From NLM Medline.
- (26) Landrum, G. RDKit: open-source cheminformatics from machine learning to chemical registration. In *Abstracts of Papers of the American Chemical Society*, 2019; AMER CHEMICAL SOC 1155 16TH ST, NW, WASHINGTON, DC 20036 USA: Vol. 258.

- (27) Kojima, R.; Ishida, S.; Ohta, M.; Iwata, H.; Honma, T.; Okuno, Y. kGCN: a graph-based deep learning framework for chemical structures. *J. Cheminform.* **2020**, *12* (1). DOI: 10.1186/s13321-020-00435-6.
- (28) Auer, P.; Cesa-Bianchi, N.; Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine learning* **2002**, *47*, 235-256.
- (29) Kusner, M. J.; Paige, B.; Hernández-Lobato, J. M. Grammar variational autoencoder. In *International conference on machine learning*, 2017; PMLR: pp 1945-1954.
- (30) Mann, H. B.; Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* **1947**, 50-60.
- (31) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* **1995**, *57* (1), 289-300.
- (32) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742-754.
- (33) McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* **2018**.
- (34) Cihan Sorkun, M.; Mullaj, D.; Koelman, J. V. A.; Er, S. ChemPlot, a Python library for chemical space visualization. *Chemistry - Methods* **2022**, *2* (7), e202200005.