Structure-based *de novo* molecular generator combined with artificial intelligence and docking simulations

Biao Ma^{†,‡}, *Kei Terayama*^{§,¶,*}, *Shigeyuki Matsumoto*[§], *Yuta Isaka*^{†,‡}, *Yoko Sasakura*^{†,‡}, *Hiroaki Iwata*[§], *Mitsugu Araki*[§], *Yasushi Okuno*^{†,‡,§,*}

[†]Center for Cluster Development and Coordination, Foundation for Biomedical Research and Innovation at Kobe, 1-5-2, Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

[‡]RIKEN Center for Computational Science, 7-1-26 Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

[§]Graduate School of Medicine, Kyoto University, 53 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan

[¶]Graduate School of Medical Life Science, Yokohama City University, 1-7-29, Suehiro-cho, Tsurumi-ku, 230-0045 Kanagawa, Japan.

KEYWORDS

De novo molecular design, Monte Carlo tree search, Recurrent neural network, Docking simulation, Structure-based drug design

ABSTRACT. Recently, molecular generation models based on deep learning have attracted significant attention in drug discovery. However, most existing molecular generation models have a serious limitation in the context of drug design wherein they do not sufficiently consider the effect of the three-dimensional (3D) structure of the target protein in the generation process. In this study, we developed a new deep learning-based molecular generator, SBMolGen, that integrates a recurrent neural network, a Monte Carlo tree search, and docking simulations. The results of an evaluation using four target proteins (two kinases and two G protein-coupled receptors) showed that the generated molecules had a better binding affinity score (docking score) than the known active compounds, and they possessed a broader chemical space distribution. SBMolGen not only generates novel binding active molecules but also presents 3D docking poses with target proteins, which will be useful in subsequent drug design.

1. INTRODUCTION

Structure-based drug design (SBDD) has emerged as an essential tool for accelerating cost-efficient lead compound discovery and optimization; indeed, numerous drugs are available in the market that have been designed by SBDD.^{1,2} Furthermore, structure-based virtual screening, the first step of SBDD, is used as an alternative computational approach to experimental high-throughput screening (HTS) because HTS is time consuming and expensive. However, existing screening approaches have a critical limitation in that the chemical space of the available

compound library is not large enough, despite that an estimated 10^{23} – 10^{60} synthesizable compound exists.³

To address this limitation, many molecular generative methods have been developed to design drug-like molecules; in particular, deep learning-based methods have recently attracted attention as *de novo* generative models.^{4,5} For example, the first neural network-based molecular generator using a variational autoencoder (VAE)⁶ and the first recurrent neural network (RNN) generative model⁷ for *de novo* molecular generation have been reported. These deep learning-based generative models are considerably effective in terms of expanding the chemical space. However, they generate only drug-like molecules, after which further structural optimization is required for actual drug development. In previous studies,^{8,9} following the generation of candidate molecules using the generative model, promising candidates were selected through docking simulations with the target protein. However, these approaches merely link the docking simulation after the calculation of the generative model. Therefore, the final selected compound is not always an optimum solution for a three-dimensional (3D) binding affinity and conformation with the target protein. Recently, Xu et al.¹⁰ attempted to generate molecules by combining the VAE model with the docking score. However, this deep learning generative model is simply a greedy search of the latent space of VAEs based on the docking score, and it can result in a longer search time before reaching the optimal molecule. Thus, although many molecular generative models have been developed to date, no deep learning-based generative model exists that can optimize the 3D binding affinity and conformation with the target biomolecule.

In this paper, we report the development of a structure-based molecular generator (SBMolGen) based on deep learning, Monte Carlo tree search (MCTS),¹¹ and docking simulations. By

combining RNN and MCTS, SBMolGen can efficiently generate novel molecules while optimizing the molecular generation process. By incorporating docking simulations into the generation process, molecules can be generated based on the binding affinity and conformation of the target protein. To evaluate the usability of SBMolGen, we perform molecular generation experiments on four different target proteins. Further, the binding affinity scores (docking scores) and chemical space distributions of the generated compounds are evaluated.

2. MATERIALS AND METHODS

2.1. SBMolGen

SBMolGen is a Python library for *de novo* molecular generation that is combined with ChemTS¹² and molecular docking based on rDock¹³. The SBMolGen scheme is shown in Figure 1.

ChemTS searches for promising targets (molecules) by growing a search tree through repetition of the four steps of MCTS: selection, expansion, simulation (rollout), and back-propagation. The details of the implementation are described in a previous report.¹² For the expansion and rollout steps, ChemTS utilizes an RNN model trained on simplified molecular input line entry system (SMILES) strings.¹⁴ In SBMolGen, the molecules generated by ChemTS (Step 1 in Fig. 1) were evaluated based on the molecular docking simulation (Steps 2 and 3 in Fig. 1). The evaluation value was then reflected in the search tree (Step 4 in Fig. 1). By repeating this loop, SBMolGen achieves structure-aware molecular generation. The RNN model

implemented in SBMolGen was trained using approximately 250,000 molecules from the ZINC database.¹⁵

Figure 1. Concept of this study and the molecular generator scheme. SBMolGen aims to generate a novel molecule with a better binding affinity score (docking score) than a known active compound for the target molecule. In SBMolGen, the molecules generated by a SMILES-based search tree and RNN (Step 1) are evaluated based on the molecular docking simulation (Steps 2 and 3). Subsequently, the evaluation value is reflected in the search tree (Step 4). By repeating this loop, SBMolGen achieves structure-aware molecular generation.

2.2. Target proteins and active compounds

Prior to commencing the molecular generation process, the target protein was prepared for docking. We prepared four target proteins that include two kinases (cyclin-dependent kinase 2 (CDK2) and epidermal growth factor receptor erbB1 (EGFR)) and two G protein-coupled receptors (GPCRs; adenosine A2a receptor (AA2AR) and beta-2 adrenergic receptor (ADRB2))

(Table 1). The structures of CDK2, EGFR, AA2AR, and ADRB2 were obtained from the Protein Data Bank (PDB),¹⁶ with PDB IDs of 1H00,¹⁷ 2RGP,¹⁸ 3EML,¹⁹ and 3NY8,²⁰ respectively. The structure preparation module in the Molecular Operating Environment ver. 2019.1 (MOE; Chemical Computing Group, Montreal, Canada) was used to prepare the structures of the target protein, including build loops for disordered loops and for the addition of any missing hydrogen atoms. To evaluate the generated molecules, we collected known active and decoy compounds of each target protein from the enhanced directory of useful decoys (DUD-E)²¹ (Table 1).

Table 1. Known active and decoy compounds of the target protein used in this study.

Protein name	PDB ID	Number of active compounds	Number of decoys
CDK2	1H00	474	27830
EGFR	2RGP	542	17924
AA2AR	3EML	482	31498
ADRB2	3NY8	231	14993

CDK2, cyclin-dependent kinase 2; EGFR, epidermal growth factor receptor erbB1; AA2AR, adenosine A2a receptor; ADRB2, Beta-2 adrenergic receptor.

2.3. Preparation of the docking and docking parameters

In this study, we utilized an open-source program, rDock, to conduct the docking simulations. The binding pocket (cavity) of the target protein was defined around the ligand in the crystal structure using the reference ligand method in the rDock package (radius: 6.0 Å). For the other settings, we used the default settings recommended by the rDock manual. The 3D structures of

the known compounds and the generated compounds were constructed from SMILES using rdkit ([http:// www.rdkit.org](http://www.rdkit.org)).

2.4. Reward function

In this study, the docking score was used as a reward function for molecular generation, and it was compared to the reference docking score of known active compounds. Before calculating the reward function, we filtered out any molecules with synthetic accessibility (SA) scores greater than 3.5 that did not satisfy Lipinski’s rule of five²² and an in-house filter based on the frequency of occurrence of molecular patterns in the PubChem²³ database (PubChem filter). Molecules that did not meet these conditions were assigned a score of -1 . For a generated molecule (S) that passed the filter, we calculated the reward score ($reward(S)$) using the following equation:

$$reward(S) = -\frac{s * (DS(S) - DS_{baseline})}{1 + s * |DS(S) - DS_{baseline}|}.$$

The reward function is designed to return a value between -1 and 1 . Here, $DS(S)$ is the best docking score selected by 10 docking simulations for each generated molecule. If $DS(S) < DS_{baseline}$, the reward is greater than zero. Based on the average of the values calculated from the active and decoy molecules of each target, we set $DS_{baseline} = -20$ for CDK2 and EGFR. For AA2AR and ADRB2, the value of $DS_{baseline}$ was set to -25 . In the above equation, s is the scaling parameter, which was fixed at 0.1 .

3. RESULTS AND DISCUSSION

3.1. Molecular generation

Molecular generation against four target proteins was performed using SBMolGen on an Intel Xeon E5-2698 v4 CPU and Nvidia RTX TITAN GPU, and the exploration time was set to 120 h. Five C parameters were used for the purpose of the MCTS, i.e., 0.2, 0.4, 0.6, 0.8, and 1.0. An average of approximately 25000 molecules was generated for each target protein during the experimental period (Figure S1). Figures 2A–2D show the distribution docking score of the generated molecules with time. The results of the generation experiments show that the molecules generated at later times have a better docking score, and the results for the two kinase proteins show this trend well. However, there are some exceptions, such as the AA2AR protein at parameter C settings of 0.6, 0.8, and 1.0 and the ADRB2 protein at a parameter C setting of 0.8, where almost no major changes were observed in the docking scores. This possibly because larger values of parameter C generate more diverse molecules, but the convergence of the docking scores becomes worse. To more clearly observe the effects of different values of parameter C on the generated molecules, we calculated the average docking score of all molecules generated for all target proteins at the same parameter C setting. As shown in Figure 2E, the docking score quickly decreased as C also decreased.

Figure 2. Distribution of the docking scores of the generated molecules over time for each target protein at different values of parameter C. A more negative docking score corresponds to a higher binding affinity. A–D represent the results for CDK2, EGFR, AA2AR, and ADRB2, respectively; E represents the effect of parameter C on the docking score of the generated

molecules (the docking score is the average of the docking scores of all molecules generated for all target proteins at the same parameter C setting).

3.2 Structural analysis of the generated molecules

As shown in Figure 3, we checked some examples of molecules generated at different times for the target protein CDK2 when parameter C was set to 1.0. We noticed that the docking score of the generated molecules for CDK2 was more negative, where a sulfo-like fragment and an amino group interacted with ASP145. Figures S2–S4 show the examples of molecules generated for the other targets. In addition, we checked the binding poses and the protein-ligand interaction diagrams for the generated top-scoring molecules in their docking complexes. The generated molecules were found to bind to the binding pockets of their respective target proteins. In addition, the best docking score obtained for a generated molecule was found to be superior to that of the known compound (Figure 4).

Figure 3. Examples of generated molecules for CDK2. The molecules were selected from the generated molecules with the highest top scores every 10 h during the experiment, and the protein-ligand interactions between molecules and the associated residues in the target protein

pocket were plotted using their docked complex structures (see Figures S2–S4 for examples of molecules generated for other targets). The protein-ligand interactions diagrams were prepared using MOE.

Figure 4. Binding model of the generated molecule and its crystal structure. The molecular structures shown in gray and in color are the known active compounds and generated molecules, respectively. A, B, and C correspond to CDK2; D, E, and F to EGFR; G, H, and I to AA2AR; and J, K, and L to ADRB2. C, F, I, and L show the protein-ligand interactions of the known compounds and the generated molecules. The number shown at the bottom of each panel is the docking score (unit is kcal/mol). M is the description of the interaction type. The generated molecules shown here are those generated with the highest docking scores to each target protein at a parameter C setting of 1.0.

3.3. Analysis of the discrepancy from known molecules. To evaluate the novelty of the generated molecules, we visualized their ISOMAP distributions in the chemical space explored by SBMolGen (Figure 5). A mapping from fingerprints to a two-dimensional space was constructed by applying the ISOMAP algorithm²⁴ to the training data set obtained from the ZINC database.¹⁶ Subsequently, the generated molecules were mapped to a two-dimensional space. Figure 5 shows that the chemical space of the generated molecules was larger than that of the ZINC dataset. For example, at C=0.2, in CDK2, SBMolGen initially produces molecules close to the training data, but gradually produces molecules that are outside the training data (upward), and once it returns to the center of the figure, the search proceeds further down to the right. This indicates that SBMolGen searches for molecules that are not included in the training data while changing direction in the middle of the search. In addition, the search that shows a decreasing score in Figure 2 indicates that SBMolGen searches for data that are not included in the training data. However, some isolated cases showed no significant differences because the generated molecules did not converge well. Observing the ISOMAP distributions of all cases, we note that when the value of parameter C is small, the generated molecules initially spread from the

distribution region similar to the ZINC data (upper right) and then gradually expand to the lower right as time progresses. When parameter C becomes larger, the generated molecules are distributed in two distinctly different regions and do not gradually transition from one region to the other.

Figure 5. ISOMAP visualization of the generated molecules and training data from the ZINC database.¹⁶ Gray dots represent molecules in the ZINC database, and heatmap dots represent the

distribution of the generated molecules over time. From top to bottom: C=0.2, 0.4, 0.6, 0.8, and 1.0.

To confirm the novelty of the generated molecules, we also compared the molecular similarities and molecular scaffold similarities between the generated molecules and the known active compounds of each target protein. The rdkit module was used to extract scaffolds and calculate the Morgan fingerprint and similarity score (Tanimoto score) from the molecular structures represented by SMILES strings. The similarity score ranges from 0 to 1, where a score close to 1 indicates a higher similarity. Figures 6, S5, and S6 show the distribution of the similarity scores. Here, it can be observed that the similarity scores between the majority of generated molecules and known active molecules were ~ 0.2 , while the similarity score for the molecular scaffold was ~ 0.3 , which implies that our generated molecules possess a high novelty.

Figure 6. Distribution of the nearest neighbor similarity values for the generated compounds to the known active compound. The upper panel shows the similarity of the molecule in general, while the lower panel shows the similarity of the molecular scaffold (a higher similarity score within the range of 0–1 indicates a higher similarity of the scaffold).

4. CONCLUSION

In this study, we developed a novel *de novo* molecular design method, SBMolGen, which uses the docking score as a reward function to generate high-affinity score compounds for a specified target. Our method was used to evaluate the binding affinities of generated molecules to a given target protein by molecular docking during the molecular generation process. In addition, the output binding poses of the generated molecules were used to evaluate these molecules from a structural point of view. Overall, the results of an evaluation using four target proteins showed that the generated molecules had a better binding affinity score (docking score) than the known active compounds, and they possessed a broader chemical space distribution. Such a capability is expected to be useful in future drug design.

In the future, for a more practical system, it is desirable to reduce the computation time. Parallelized methods of tree search such as virtual loss^{11,25} are considered promising. In addition, only the docking score with the filters based on the SA score and Lipinski's rule of five, and the PubChem filter were used in this study, although optimization of multiple properties is necessary for drug discovery.²⁶ We have a plan to realize such multi-objective optimization using multi-objective MCTS algorithms.^{27,28}

ASSOCIATED CONTENT

Supporting Information

Supporting Information is available free of charge on the ACS Publications website: number of generated molecules, example of generated molecules for target EGFR at C=1.0, example of generated molecules for target AA2AR at C=1.0, example of generated molecules for target ADRB2 at C=1.0; distribution of the similarity score of the generated molecules, and distribution of the scaffold similarity score of the generated molecules.

AUTHOR INFORMATION

Corresponding Author

*Correspondence: terayama@yokohama-cu.ac.jp, okuno.yasushi.4c@kyoto-u.ac.jp

Author Contributions

B.M., K.T., and Y.O. designed the study. B.M. and K.T. constructed the Python script of SBMolGen. B.M. performed the molecular generation experiments and analyses. Y.S. prepared the protein structure S.M., Y.I., H.I., and M.A. performed the discussion and manuscript editing. All authors have approved the final version of the manuscript.

Funding Sources

This work was supported by MEXT as a “Priority Issue on Post K computer (Building Innovative Drug Discovery Infrastructure Through Functional Control of Biomolecular

Systems),” MEXT as a “Program for Promoting Researches on the Supercomputer Fugaku (Application of Molecular Dynamics Simulation to Precision Medicine Using Big Data Integration System for Drug Discovery),” and the AMED under Grant Number JP20nk0101111.

Notes

The authors declare no competing financial interests.

Data Availability

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

Code Availability. The code used to generate results shown in this study is available from the corresponding author upon request.

ACKNOWLEDGMENT

We would like to thank Professor K. Tsuda (The University of Tokyo, RIKEN) for providing the ChemTS program and to Dr. M. Ohta (RIKEN) and the members of the LINC-PJ16 for their useful discussions and helpful comments regarding the filter for the molecule.

References

- (1) Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1923–1938.

- (2) Batool, M.; Ahmad, B.; Choi, S. A Structure-Based Drug Discovery Paradigm. *Int. J. Mol. Sci.* **2019**, *20* (11), 2783.
- (3) Polishchuk, P. G.; Madzhidov, T. I.; Varnek, A. Estimation of the Size of Drug-like Chemical Space Based on GDB-17 Data. *J. Comput. Aided. Mol. Des.* **2013**, *27* (8), 675–679.
- (4) Xu, Y.; Lin, K.; Wang, S.; Wang, L.; Cai, C.; Song, C.; Lai, L.; Pei, J. Deep Learning for Molecular Generation. *Future Med. Chem.* **2019**, *11* (6), 567–597.
- (5) Elton, D. C.; Boukouvalas, Z.; Fuge, M. D.; Chung, P. W. Deep Learning for Molecular Design—a Review of the State of the Art. *Mol. Syst. Des. Eng.* **2019**, *4* (4), 828–849.
- (6) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4* (2), 268–276.
- (7) Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Cent. Sci.* **2018**, *4* (1), 120–131.
- (8) Sattarov, B.; Baskin, I. I.; Horvath, D.; Marcou, G.; Bjerrum, E. J.; Varnek, A. De Novo Molecular Design by Combining Deep Autoencoder Recurrent Neural Networks with Generative Topographic Mapping. *J. Chem. Inf. Model.* **2019**, *59* (3), 1182–1196.

- (9) Grebner, C.; Matter, H.; Plowright, A. T.; Hessler, G. Automated de Novo Design in Medicinal Chemistry: Which Types of Chemistry Does a Generative Neural Network Learn? *J. Med. Chem.* **2020**, *63* (16), 8809–8823.
- (10) Xu, Z.; Wauchope, O.; Frank, A. T. Navigating Chemical Space by Interfacing Generative Artificial Intelligence and Molecular Docking. *bioRxiv* **2020**, 2020.06.09.143289.
- (11) 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 Trans. Comput. Intell. AI Games* **2012**, *4* (1), 1–43.
- (12) Yang, X.; Zhang, J.; Yoshizoe, K.; Terayama, K.; Tsuda, K. ChemTS: An Efficient Python Library for de Novo Molecular Generation. *Sci. Technol. Adv. Mater.* **2017**, *18* (1), 972–976.
- (13) Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; Barril, X.; Hubbard, R. E.; Morley, S. D. RDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLOS Comput. Biol.* **2014**, *10* (4), e1003571.
- (14) Yoshikawa, N.; Terayama, K.; Sumita, M.; Homma, T.; Oono, K.; Tsuda, K. Population-Based De Novo Molecule Generation, Using Grammatical Evolution. *Chem. Lett.* **2018**, *47* (11), 1431–1434.
- (15) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768.

- (16) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- (17) Beattie, J. F.; Breault, G. A.; Ellston, R. P. A.; Green, S.; Jewsbury, P. J.; Midgley, C. J.; Naven, R. T.; Minshull, C. A.; Pauptit, R. A.; Tucker, J. A.; Pease, J. E. Cyclin-Dependent Kinase 4 Inhibitors as a Treatment for Cancer. Part 1: Identification and Optimisation of Substituted 4,6-Bis Anilino Pyrimidines. *Bioorg. Med. Chem. Lett.* **2003**, *13* (18), 2955–2960.
- (18) Xu, G.; Abad, M. C.; Connolly, P. J.; Neeper, M. P.; Struble, G. T.; Springer, B. A.; Emanuel, S. L.; Pandey, N.; Gruninger, R. H.; Adams, M.; Moreno-Mazza, S.; Fuentes-Pesquera, A. R.; Middleton, S. A. 4-Amino-6-Arylamino-Pyrimidine-5-Carbaldehyde Hydrazones as Potent ErbB-2/EGFR Dual Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18* (16), 4615–4619.
- (19) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science* **2008**, *322* (5905), 1211–1217.
- (20) Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved Binding Mode of Human B2 Adrenergic Receptor Inverse Agonists and Antagonist Revealed by X-Ray Crystallography. *J. Am. Chem. Soc.* **2010**, *132* (33), 11443–11445.

- (21) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55* (14), 6582–6594.
- (22) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1–3), 3–26.
- (23) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem in 2021: New Data Content and Improved Web Interfaces. *Nucleic Acids Res.* **2021**, *49* (D1), D1388–D1395.
- (24) Tenenbaum, J. B.; De Silva, V.; Langford, J. C. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **2000**, *290* (5500), 2319–2323.
- (25) Zhang, J.; Terayama, K.; Sumita, M.; Yoshizoe, K.; Ito, K.; Kikuchi, J.; Tsuda, K. NMR-TS: De Novo Molecule Identification from NMR Spectra. *Sci. Technol. Adv. Mater.* **2020**, *21* (1), 552–561.
- (26) Winter, R.; Montanari, F.; Steffen, A.; Briem, H.; Noé, F.; Clevert, D. A. Efficient Multi-Objective Molecular Optimization in a Continuous Latent Space. *Chem. Sci.* **2019**, *10* (34), 8016–8024.
- (27) Wang, W.; Sebag, M.; Hoi, S. C. H.; Buntine, W. Multi-Objective Monte-Carlo Tree Search. In *Proceedings of the Asian Conference on Machine Learning*; PMLR, 2012; Vol. 25, pp 507–522.

- (28) Perez, D.; Mostaghim, S.; Samothrakis, S.; Lucas, S. M. Multiobjective Monte Carlo Tree Search for Real-Time Games. *IEEE Trans. Comput. Intell. AI Games* **2015**, 7 (4), 347–360.

Table of Contents