Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Generating potential protein-protein interaction inhibitor molecules based on physicochemical properties

Masahito Ohue[1,*], Yuki Kojima[1] and Takatsugu Kosugi[1]

[1] Department of Computer Science, School of Computing, Tokyo Institute of Technology; ohue@c.titech.ac.jp

* Correspondence: ohue@c.titech.ac.jp

Abstract: Protein-protein interactions (PPIs) are associated with various diseases; hence, they are important targets in drug discovery. However, the physicochemical empirical properties of PPI-targeted drugs are distinct from those of conventional small molecule oral pharmaceuticals, which adhere to the "rule of five (RO5)." Therefore, developing PPI-targeted drugs using conventional methods, such as molecular generation models, is difficult. In this study, we propose a molecule generation model based on deep reinforcement learning, which is specialized for generating PPI inhibitor candidates. We successfully generated potential PPI inhibitor compounds by modifying the scoring functions of the existing small molecule generation model and constructed a virtual library of generated PPI inhibitor compounds.

Keywords: protein-protein interaction inhibitor, rule of five, rule of four, QEPPI, molecular generation, virtual chemical library

1. Introduction

Significant advances in science and technology have been made since the 1950s; however, the efficiency of the drug discovery process has notably declined. Specifically, the number of drug approvals per billion US dollars of research and development spending has halved approximately every nine years [1]. One of the main reasons for this decline in drug discovery is that the current drug-target space is nearly saturated. Therefore, since the early 2000s, researchers have been actively exploring novel therapeutic targets, such as protein-protein interactions (PPIs) [2–6]. PPIs play a crucial role in various cellular processes and are associated with many diseases, such as cancer [6] and Alzheimer's disease [7]. However, targeting PPIs is complex, and only a few compounds have been approved or have progressed to the clinical trial stage as PPI inhibitors [8].

One of the difficulties in targeting PPIs is that they have different properties compared to conventional drug discovery targets. The binding interface of PPIs is wider than the average binding region of a typical protein target. As a result, PPI inhibitor molecules tend to be larger and more diverse in conformation [9].

Lipinski et al. [10, 11] proposed a rule of thumb, known as the "Rule of five (RO5)," to evaluate the likelihood of a compound with specific physicochemical properties becoming an orally active drug. According to this rule, the compounds that do not meet two or more of the following four criteria are considered to be poorly absorbed and unlikely to eventually become pharmaceutical products.

- Molecular weight ≤ 500
- Indicator of lipophilicity, LogP ≤ 5
- Number of hydrogen bond donors ≤ 10
- Number of hydrogen bond acceptors ≤ 5

On the other hand, Morelli et al. [12] studied the properties of PPI inhibitors obtained from the 2P2I database [13]. This database contains protein complex structures and interaction regions stored in the PDB file format, along with complex structures where PPI inhibitors are bound. They showed

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

that most PPI inhibitors from the database violated RO5 significantly. Furthermore, they showed that the following rule of thumb, called "Rule of four (RO4)", was more applicable to the PPI inhibitors.

- Molecular weight > 400
- Indicator of lipophilicity, LogP > 4
- Number of cyclic structures > 4
- Number of hydrogen bond acceptors > 4

High throughput screening (HTS) is commonly used to find active compounds in the early phases of drug discovery. HTS is a rapid assay process that evaluates numerous compounds quickly and can effectively identify active compounds from vast libraries for a specific target. However, when HTS is performed for PPI as a target using a compound library consisting of general small molecule compounds, the acquisition rate of hit compounds is significantly lower [14]. Therefore, studies on the development of PPI-specific libraries to obtain high hit rates have been conducted to date [15,16]. However, such existing compound libraries mainly contain derivatives of the principal core structures used in known PPI inhibitors, limiting the diversity of compounds in the libraries.

To address this restricted range of compounds, we focused on a large-scale search method in this study. Moreover, we used a deep-reinforcement-learning-based molecular generation model to produce potential novel PPI inhibitors. Molecular generation aims at discovering novel compounds with desirable properties and activities from a vast compound space. To produce potential PPI inhibitors, the scoring function of the molecular generation model was modified according to their properties. In addition, a virtual library containing the generated PPI-target compounds was constructed.

2. Materials and Methods

2.1. Molecular generation model

In this study, we used a molecular generation model called REINVENT, developed by Blaschke et al. [17]. This model is based on character string (SMILES notation) and uses recurrent neural networks as the architecture. Moreover, it generates molecules with desired properties in combination with reinforcement learning. A pre-trained model based on ChEMBL [18], which is a database of chemical compounds, was obtained from [19] and used for reinforcement learning in the same way as described in a previous study [17].

2.2. Scoring function

The scoring function of REINVENT $S(x)$ of one generated compound x is calculated in the range of 0 to 1. Three scoring functions used in this study were quantitative estimates of drug-likeness (QED) [20], RO4 [12], and quantitative estimate index for early-stage screening of compounds targeting PPIs (QEPPI) [21]. QED is an index that quantifies the drug-likeness of small molecules and is equivalent to making RO5 continuous. It is defined in the range of 0 to 1 and was used as the score $S(x)$. RO4 was introduced as the REINVENT score, and for each fulfilled RO4 four conditions, 0.25 is added to the total score $S(x)$; thus, the score ranges from 0 to 1. The QEPPI is a quantitative index of PPI inhibitor suitability, ranging from 0 to 1. It has also been used for PPI inhibitor evaluation in molecular generation studies [22]. The QEPPI was used directly as the overall score $S(x)$ in this study.

2.3 Computational experiments

The three scoring functions, QED-based, RO4-based, and QEPPI-based, described above, were used in this work, and reinforcement learning was performed for 3,000 steps each. Parameters other than the scoring function $S(x)$ were set to default values [19]. The numbers of compounds generated by applying the QED, RO4, and QEPPI scoring functions in 3,000 steps of REINVENT training were 357,456; 368,140; and 359,722, respectively. To obtain higher scoring molecules among the generated compounds, the compounds generated in the last 100 steps (i.e., between steps 2,901 and 3,000) were extracted, resulting in 11,714; 12,547; and 12,097 compounds for QED, RO4, and QEPPI, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

3. Results

3.1 Inducing exploration through reinforcement learning

The QED, RO4 equivalent, and QEPPI scores were calculated for each of the three generated compound sets, and their distributions for each score are shown in Figures 1–3. As anticipated and in line with expectations, the investigation confirmed that compounds generated with QED as the scoring function tend to have higher QED scores (Figure 1), compounds generated with RO4 as the scoring function tend to have higher RO4 scores (Figure 2), and compounds generated with QEPPI as the scoring function tend to have higher QEPPI scores (Figure 3). As can be seen, the exploration was induced to the region where the scoring function was higher, indicating that reinforcement learning was properly performed.

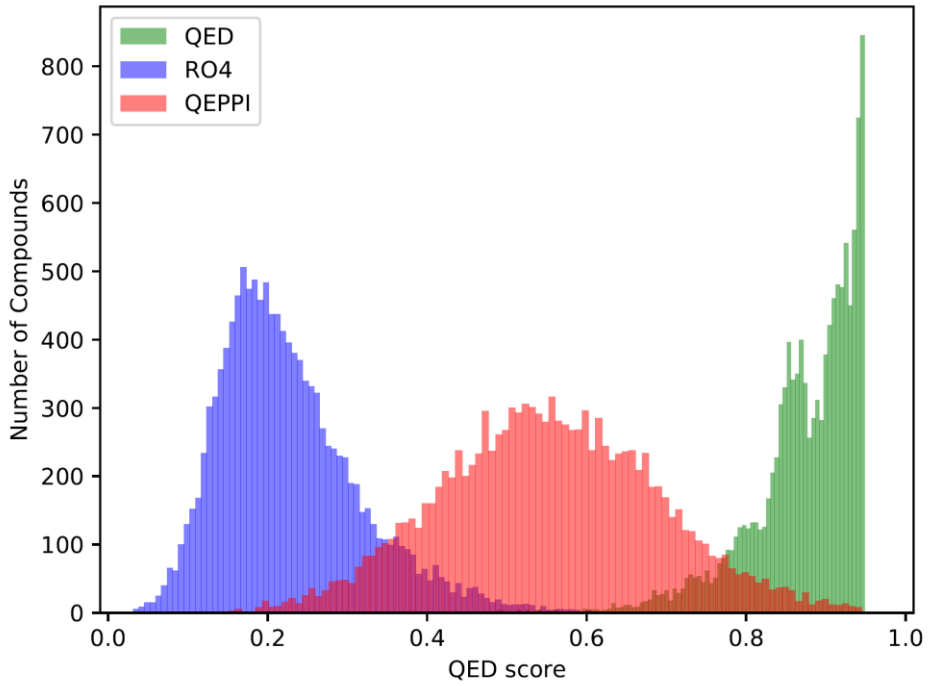


Figure 1. Distribution of QED scores of the compounds generated between the REINVENT steps 2,901 and 3,000.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

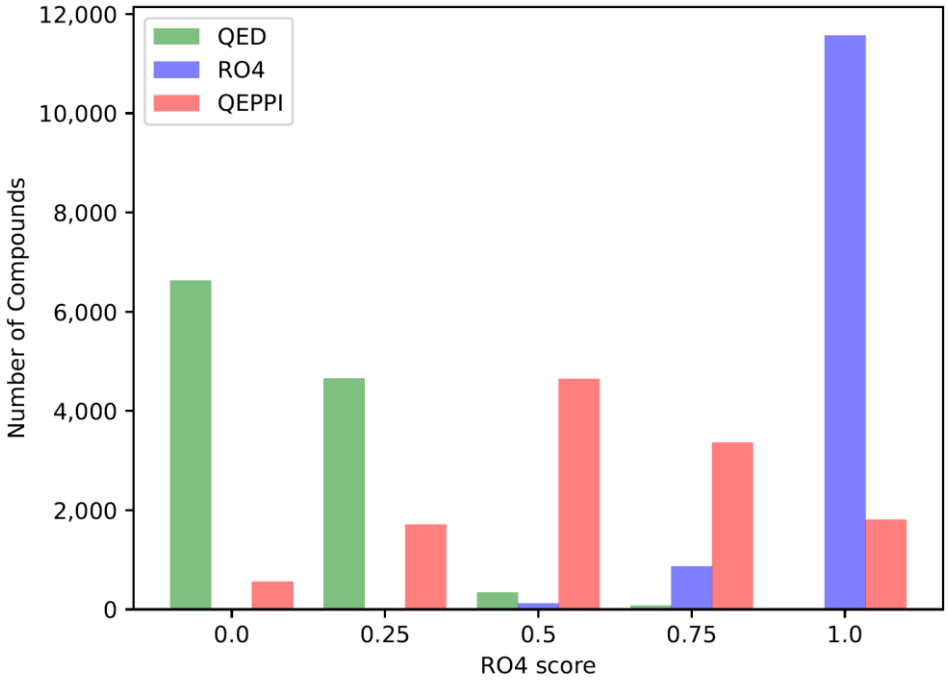


Figure 2. Distribution of RO4 equivalent scores of the compounds generated between the REINVENT steps 2,901 and 3,000.

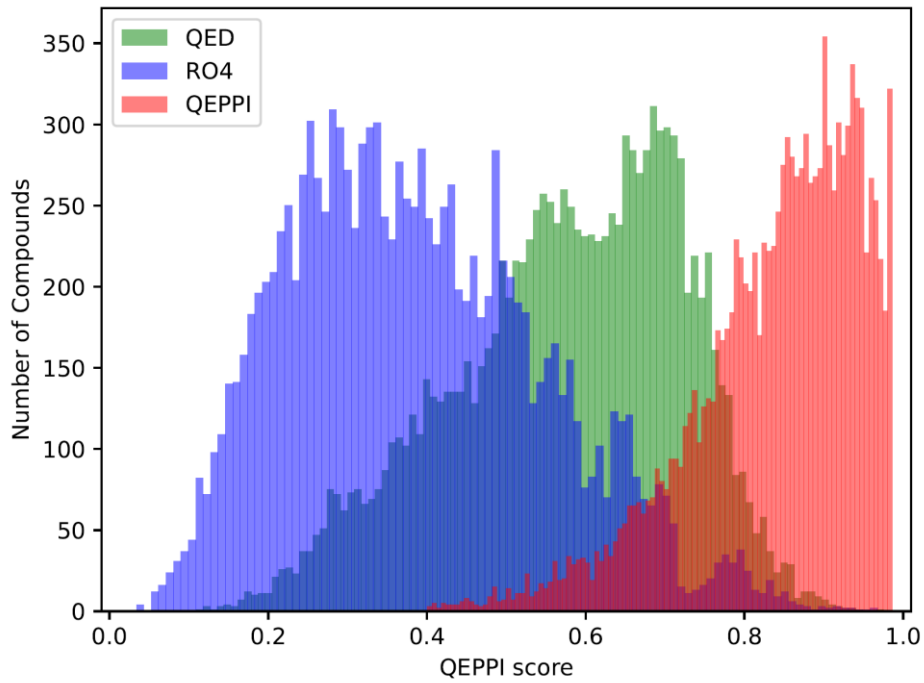
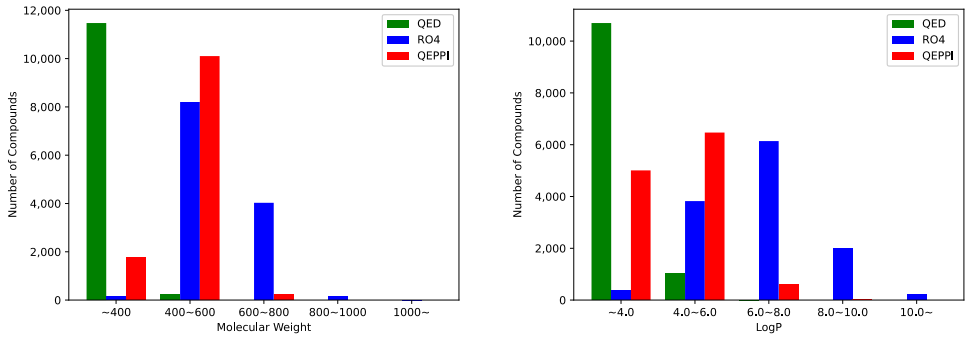


Figure 3. Distribution of QEPPI scores of the compounds generated between the REINVENT steps 2,901 and 3,000.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

3.2 Distribution of compounds generated by REINVENT

Molecular weight and LogP, a measure of lipophilicity, are both related to RO4 and can be used to evaluate PPI inhibitor-likeness. Figures 4(a) and 4(b) show the distribution of molecular weight and LogP of the generated compounds for each scoring function. In the case of QED, both molecular weight and LogP were skewed toward the smaller values, and the generated molecules could not be suitable for PPI inhibitors. On the other hand, when RO4 was used, most of the molecules generated satisfied RO4 (molecular weight > 400, LogP > 4). Although the QEPPI-generated molecules did not satisfy RO4, 85.3% and 58.6% of all the generated compounds fulfilled RO4 conditions for molecular weight and LogP, respectively. In addition, Figure 5 plots a two-dimensional scatter plot of the overlaid molecular weight and LogP distributions for the RO4 and QEPPI cases. Molecules with excessively high molecular weight and LogP, such as in RO4, were not generated by the QEPPI scoring function, and the compounds were concentrated in the appropriate chemical space where high QEPPI scores can be obtained.



(a) Molecular Weight (b) LogP

Figure 4. Distribution of the generated compounds. (a) Distribution of molecular weight, (b) distribution of lipophilicity (LogP).

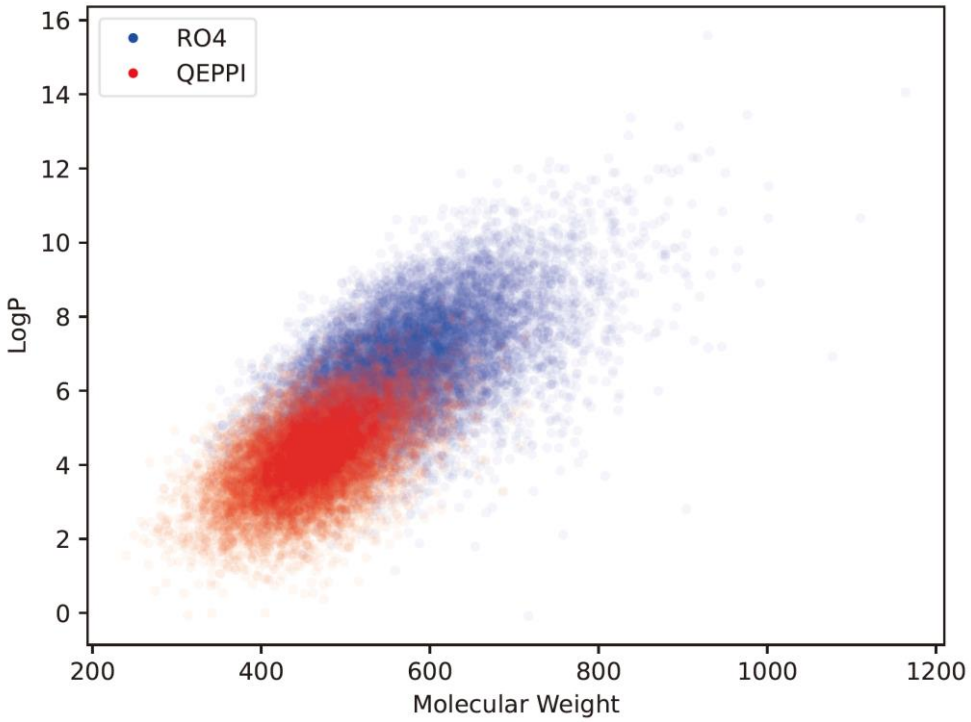


Figure 5. Molecular weight–LogP scatter plots of the RO4- and QEPPI-generated compounds.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

3.3 Indicators for oral bioavailability

Veber's rule is an indicator of oral bioavailability, and compounds that satisfy the following two properties tend to have good membrane permeability when administered orally [23].

- Number of rotatable bonds, Rbond ≤ 10
- Topological polar surface area TPSA ≤ 140

The numbers of QED-, RO4-, and QEPPI-generated molecules satisfying Veber's rule were 11,712; 9,617; and 12,029, with percentages of 99.9% (11,712/11,714), 76.6% (9,617/12,547), and 99.4% (12.029/12,097), respectively. Thus, molecules generated using the QEPPI scoring function can also be expected to have good oral bioavailability at ratios almost equal to those generated using the QED.

3.4 Constructing virtual libraries of PPI-target compounds

Based on the results, the compounds generated based on the QEPPI (proposed method 2) are more likely to be able to target PPIs, as they combine general drug-like properties while satisfying PPI inhibitor-specific properties such as RO4. Therefore, 12,097 compounds generated based on QEPPI were used as the basis for the virtual library for developing PPI inhibitors.

In addition, the pan assay interference compounds (PAINS) filter [24] was used to identify substructural features of generated compounds that appear in promiscuous compounds or frequent hitters in many biochemical high throughput screening campaigns and removed their compounds. PAINS filter removed 845 (7.0%) compounds. Then, the remaining 11,252 compounds constituted the virtual library, published on https://github.com/ohuelab/iPPI-REINVENT. Figure 6 shows some examples of compounds included in this library. Compounds **a**, **b**, and **c** in Figure 6 had the average, minimum, and maximum molecular weights, respectively, of the compounds in the virtual library. Compound **d** had the highest number of similar known PPI inhibitors. Compound **e** is an example of a compound containing a slightly longer alkyl chain. In fact, these compounds will not be suitable as drugs as they are, especially **c**, which has a strange structure with an aromatic ring containing a sulfonamide. However, we hope that it will be obtained PPI inhibitor hits from this virtual library and contribute to drug discovery.

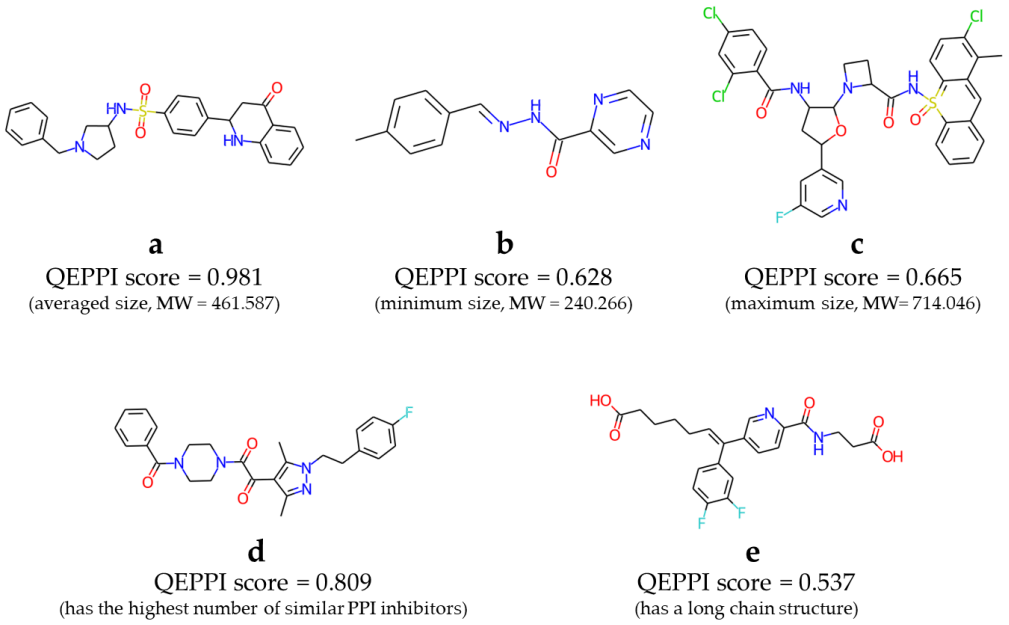


Figure 6. Examples of compounds that have been generated and included in the virtual library.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

4. Discussion

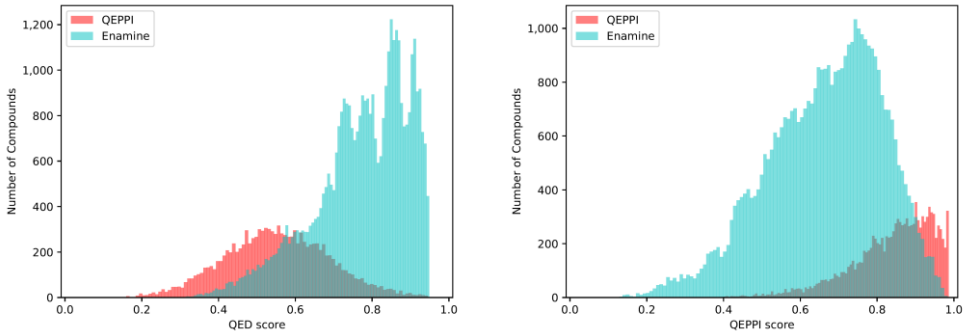
4.1 Chemical space of generated compounds

One of the objectives of this study is to generate novel PPI inhibitor compound candidates from an unexplored compound space by modifying a scoring function according to the characteristics of PPI inhibitors in the molecule generation model. In this study, three types of scoring functions, QED, RO4, and QEPPI, were used to explore the compound space and generate molecules. The distributions of molecular weight and LogP for each compound shown in Figures 4 and 5 are skewed toward the smaller values for QED and are excessively large for RO4. On the other hand, the QEPPI-generated molecules exhibited smaller and more coherent distributions while satisfying RO4. Therefore, the QEPPI-based molecular generation achieved the above objective by exploring the compound space that remained unexplored in the cases of QED and RO4. Note that the pre-trained model used in this study utilized ChEMBL and was not PPI inhibitor specific. Therefore, pre-training with a dataset derived from a known PPI inhibitor may improve the efficiency of the exploration and generation.

4.2 Comparison with existing PPI libraries

To determine whether the QEPPI-based compound library, which is considered more suitable for recent PPI inhibitor design, could be a useful virtual library, we compared the corresponding 12,097 compounds with those in an existing PPI compound library—the Enamine PPI library [25]. This commercial PPI library contains 40,640 compounds with core structures that are expected to bind to specific substructures extracted from more than 20 different protein complex structures.

Figure 7(a) and Figure 7(b) show the distributions of the QED and QEPPI scores. In addition, Figure 8 shows two-dimensional scatter plots for molecular weight and LogP. These results show that the compounds in the Enamine library have small molecular weights and LogP values and high QED scores, while those generated by QEPPI have higher molecular weights and LogP values. Moreover, the QEPPI scores of the compounds generated in this study were also higher. Therefore, we can conclude that commercial PPI libraries are oriented toward more oral drug-like tendencies and that our proposed generated compounds contain more PPI inhibitor-like compounds.



(a) QED (b) QEPPI

Figure 7. Distribution of scores (number of steps: 100) for the compounds generated by the proposed method 2 and obtained from the Enamine PPI library: (a) QED and (b) QEPPI scores.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

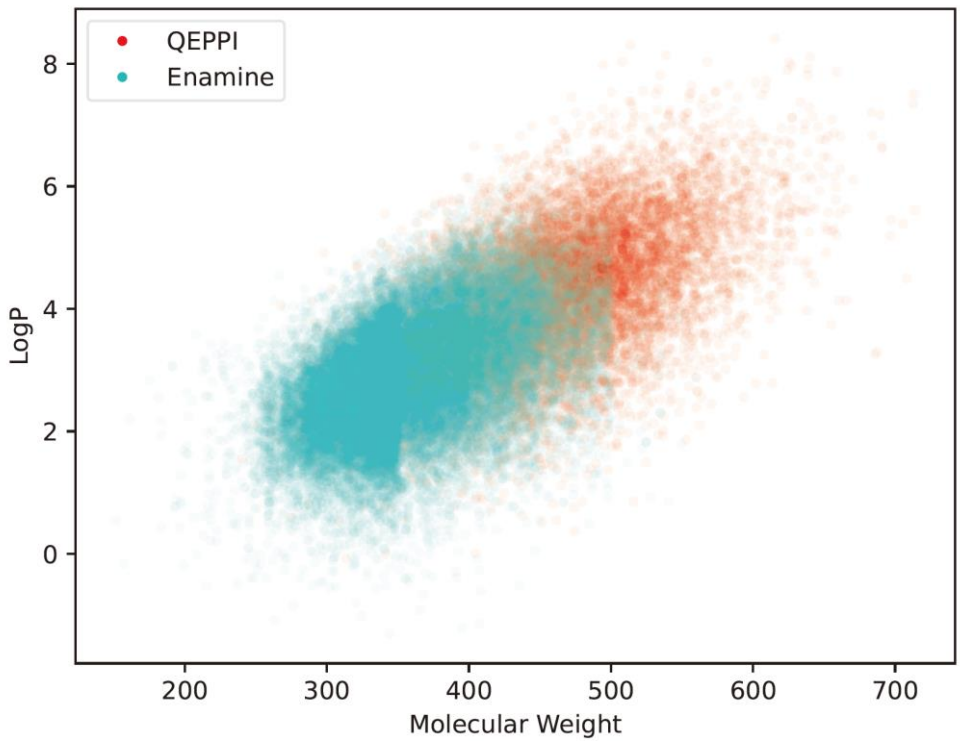


Figure 8. Molecular weight–LogP scatter plots of the compounds generated by the proposed method 2 and obtained from the Enamine PPI library.

5. Conclusions

In this study, we used REINVENT—a molecular generation tool—to generate compounds using three types of scoring functions: QED, RO4, and QEPPI. In all cases, we could generate compounds with the desired properties targeted by the scoring functions. However, this study aimed to generate candidate compounds for novel PPI inhibitors, and most of the compounds generated based on QED did not meet RO4 conditions. In addition, using the QEPPI scoring function, a higher ratio of molecules was generated with superior oral bioavailability compared to the RO4-generated molecules. From this collection of QEPPI-based compounds with PPI inhibitor-likeness, 11,252 compounds were selected, excluding those with improper structures with PAINS filter. These compounds are for use as a virtual library in assay experiments. Compared with the Enamine PPI library, these compounds can be considered a complement to the existing PPI libraries.

Only the QEPPI-generated compounds were considered in this study. Since REINVENT supports complex scoring functions, we believe that the search for compounds in chemical space can be broadened by applying and modifying various scoring functions. In future work, we intend to explore ways to enhance the scoring functions and expand our search to include chemical spaces that cannot be adequately covered by the existing PPI libraries.

The synthesizability of virtual compounds is also crucial for actual biochemical assays. PAINS filter was applied to the generated compounds to remove inappropriate compounds; however, this method alone cannot consider synthetic feasibility. In fact, in the Enamine PPI library, those compounds that fit all of the multiple medicinal chemistry filters, including PAINS, were selected [25]; for example, the synthetic accessibility score [26] guesses the difficulty of synthesizing a compound. There is also room for consideration of a more reliable retrosynthetic analysis [27], although it requires long computation times. In future work, we would like to incorporate such a method that allows us to computationally evaluate the synthetic feasibility of compounds. It would be necessary to consider and provide the contribution of molecular substructures with interpretable artificial intelligence and other technologies [28]. We aim to construct an even more helpful chemical library that will allow efficient screening in the future.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

Author Contributions: Conceptualization, M.O.; methodology, M.O and Y.K.; software, Y.K.; validation, Y.K. and T.K.; formal analysis, M.O. and Y.K.; investigation, M.O.; resources, Y.K.; data curation, M.O. and Y.K.; writing—original draft preparation, M.O. and Y.K.; writing—review and editing, M.O.; visualization, M.O. and Y.K.; supervision, M.O.; project administration, M.O.; funding acquisition, M.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by JST FOREST (JPMJFR216J), JST ACT-X (JPMJAX20A3), JSPS KAKENHI (20H04280 and 23H04887).

Data Availability Statement: Data on the compounds generated and codes created in this study are provided at https://github.com/ohuelab/iPPI-REINVENT.

Acknowledgments: We would like to thank Dr. Kazuki Yamamoto and Dr. Kentaro Rikimaru for useful discussion.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Scannell, J.W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov*. **2012**, *11*, 191–200. DOI:10.1038/nrd3681.
2. Toogood, P.L. Inhibition of protein-protein association by small molecules: Approaches and progress. *J Med Chem*. **2002**, *45*, 1543–1558. DOI:10.1021/jm010468s.
3. Arkin, M.R.; Wells, J.A. Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nat Rev Drug Discov*. **2004**, *3*, 301–317. DOI:10.1038/nrd1343.
4. Dev, K.K. Making protein interactions druggable: Targeting PDZ domains. *Nat Rev Drug Discov*. **2004**, *3*, 1047–1056. DOI:10.1038/nrd1578.
5. Jin, L.; Wang, W.; Fang, G. Targeting protein-protein interaction by small molecules. *Annu Rev Pharmacol Toxicol*. **2014**, *54*, 435–456. DOI:10.1146/annurev-pharmtox-011613-140028.
6. Ivanov, A.A.; Khuri, F.R.; Fu, H. Targeting protein–protein interactions as an anticancer strategy. *Trends Pharmacol Sci*. **2013**, *34*, 393–400. DOI:10.1016/j.tips.2013.04.007.
7. Mao, Y.; Fisher, D.W.; Yang, S.; Keszycki, R.M.; Dong, H. Protein-protein interactions underlying the behavioral and psychological symptoms of dementia (BPSD) and Alzheimer's disease. *PLOS ONE*. **2020**, *15*, e0226021. DOI:10.1371/journal.pone.0226021.
8. Shin, W.H.; Kumazawa, K.; Imai, K.; Hirokawa, T.; Kihara, D. Current challenges and opportunities in designing protein-protein interaction targeted drugs. *Adv Appl Bioinform Chem*. **2020**, *13*, 11–25. DOI:10.2147/AABC.S235542.
9. Shin, W.H.; Christoffer, C.W.; Kihara, D. In silico structure-based approaches to discover protein-protein interaction-targeting drugs. *Methods*. **2017**, *131*, 22–32. DOI:10.1016/j.ymeth.2017.08.006.
10. 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*. **1997**, *23*, 3–26.
11. Lipinski, C.A. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov Today Technol*. **2004**, *1*, 337–341. DOI:10.1016/j.ddtec.2004.11.007.
12. Morelli, X.; Bourgeas, R.; Roche, P. Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I). *Curr Opin Chem Biol*. **2011**, *15*, 475–481. DOI:10.1016/j.cbpa.2011.05.024.
13. Basse, M.J.; Betzi, S.; Morelli, X.; Roche, P. 2P2Idb v2: Update of a structural database dedicated to orthosteric modulation of protein-protein interactions. *Database (Oxford)*. **2016**, *2016*, No. baw007. DOI:10.1093/database/baw007.
14. Jnoff, E.; Albrecht, C.; Barker, J.J.; Barker, O.; Beaumont, E.; Bromidge, S.; Brookfield, F.; Brooks, M.; Bubert, C.; Ceska, T.; et al. Binding mode and structure-activity relationships around direct inhibitors of the Nrf2-Keap1 complex. *ChemMedChem*. **2014**, *9*, 699–705. DOI:10.1002/cmdc.201300525.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0704.v1

15. Bosc, N.; Muller, C.; Hoffer, L.; Lagorce, D.; Bourg, S.; Derviaux, C.; Gourdel, M.E.; Rain, J.C.; Miller, T.W.; Villoutreix, B.O.; et al. Fr-PPIChem: An academic compound library dedicated to protein-protein interactions. *ACS Chem Biol*. **2020**, *15*, 1566–1574. DOI:10.1021/acschembio.0c00179.
16. Shimizu, Y.; Yonezawa, T.; Sakamoto, J.; Furuya, T.; Osawa, M.; Ikeda, K. Identification of novel inhibitors of Keap1/Nrf2 by a promising method combining protein–protein interaction-oriented library and machine learning. *Sci Rep*. **2021**, *11*, 7420. DOI:10.1038/s41598-021-86616-1.
17. Blaschke, T.; Aŕus-Pous, J.; Chen, H.; Margreitter, C.; Tyrchan, C.; Engkvist, O.; Papadopoulos, K.; Patronov, A.; REINVENT 2.0: An AI tool for de novo drug design. *J Chem Inf Model*. **2020**, *60*, 5918–5922. DOI:10.1021/acs.jcim.0c00915.
18. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrían-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños M.P.; Overington, J.P.; Papadatos, G.; Smit, I.; Leach, A.R. The ChEMBL database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. DOI: 10.1093/nar/gkw1074.
19. MolecularAI. ReinventCommunity. Available online: https://github.com/MolecularAI/ReinventCommunity.
20. Bickerton, G.R.; Paolini, G.V.; Besnard, J.; Muresan, S.; Hopkins, A.L. Quantifying the chemical beauty of drugs. *Nat Chem*. **2012**, *4*, 90–98. DOI:10.1038/nchem.1243.
21. Kosugi, T.; Ohue, M. Quantitative estimate index for early-stage screening of compounds targeting protein-protein interactions. *Int J Mol Sci*. **2021**, *22*, 10925. DOI:10.3390/ijms222010925.
22. Wang, J.; Chu, Y.; Mao, J.; Jeon, H.N.; Jin, H.; Zeb, A.; Jang, Y.; Cho, K.H.; Song, T.; . De novo molecular design with deep molecular generative models for PPI inhibitors. *Brief Bioinform*. **2022**, *23*, bbac285. DOI:10.1093/bib/bbac285.
23. Veber, D.F.; Johnson, S.R.; Cheng, H.Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**, *45*, 2615–2623. DOI:10.1021/jm020017n.
24. Baell, J.B.; Holloway, G.A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem*. **2010**, *53*, 2719–2740. DOI:10.1021/jm901137j.
25. Enamine; PPI Library. Available online: https://enamine.net/compound-libraries/targeted-libraries/ppi-library.
26. Ertl, P.; Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Cheminform*. **2009**, *1*, 8. DOI:10.1186/1758-2946-1-8.
27. Genheden, S.; Thakkar, A.; Chadimová, V.; Reymond, J.L.; Engkvist, O.; Bjerrum, E. AiZynthFinder: A fast, robust and flexible open-source software for retrosynthetic planning. *J Cheminform*. **2020**, *12*, 70. DOI:10.1186/s13321-020-00472-1.
28. Kengkanna, A.; Ohue, M. Enhancing Model Learning and Interpretation Using Multiple Molecular Graph Representations for Compound Property and Activity Prediction. *arXiv* **2023**, 2304.06253. DOI:10.48550/arXiv.2304.06253.