

Empirical Evidence for the Fragment level Understanding on Drug Molecular Structure of LLMs

Xiuyuan Hu¹, Guoqing Liu², Yang Zhao¹, Hao Zhang^{1*}

¹Department of Electronic Engineering, Tsinghua University, Beijing, China, ²Microsoft Research AI4Science
huxy22@mails.tsinghua.edu.cn, guoqingliu@microsoft.com, {zhao-yang,haozhang}@tsinghua.edu.cn

Abstract

AI for drug discovery has been a research hotspot in recent years, and SMILES-based language models has been increasingly applied in drug molecular design. However, no work has explored whether and how language models understand the chemical spatial structure from 1D sequences. In this work, we pre-train a transformer model on chemical language and fine-tune it toward drug design objectives, and investigate the correspondence between high-frequency SMILES substrings and molecular fragments. The results indicate that language models can understand chemical structures from the perspective of molecular fragments, and the structural knowledge learned through fine-tuning is reflected in the high-frequency SMILES substrings generated by the model.

1 Introduction

With the explosion of computational technologies such as deep learning and the accumulation of data across various disciplines including biology, chemistry, medicine, and pharmacology, the potential of artificial intelligence (AI) in drug discovery has increasingly attracted attention in recent years (Vamathevan et al. 2019; Paul et al. 2021). Compared to traditional wet-laboratory-based drug research and development processes, computer-aided drug discovery (CADD) has evident advantages in speed and cost, exemplified by techniques such as high-throughput screening (Sabe et al. 2021). Notably, AI tools may drive a greater pharmaceutical revolution, as they not only further improve computational accuracy and efficiency but also demonstrate unprecedented creativity (Zhang et al. 2023).

Large language models (LLMs) based on transformer architecture have swept the field of natural language processing (NLP), especially achieving stunning breakthroughs in understanding and generation abilities for conversational intelligence (Ouyang et al. 2022; OpenAI 2023; Zhao et al. 2023). However, the performance of the currently best general language model, GPT-4, is still limited in the domain of drug discovery; for example, its reliability in accomplishing quantitative tasks and processing molecular sequences is doubtful (AI4Science and Quantum 2023). As it stands,

training specialized language models for chemical language to deal with problems in drug discovery seems more meaningful than using general language models.

At present, generative language models have been widely applied in drug molecular design, and these models can generate a set of desirable molecules according to given objectives. Particularly, existing works have employed transformer-based language models to process chemical molecular strings; examples include MolGPT (Bagal et al. 2022), Chemformer (Irwin et al. 2022) and MolRL-MGPT (Hu et al. 2023). For this reason, tokenizers specifically designed for chemical languages have been introduced, and models are trained on various molecular data. The chemical language most commonly used by current AI researchers is the simplified molecular input line entry system (SMILES) (Weininger 1988), which can represent any molecular structure by a sequence of characters.

However, although SMILES-based drug design approaches are competitive (Gao et al. 2022), understanding and generating 1D molecular strings appear less intuitive and less interpretable than methods based on 2D molecular graphs. For instance, two atoms connected by a chemical bond in a molecule may not be adjacent in a SMILES string, posing a challenge for language models to correctly comprehend chemical structures. Therefore, to design better and more reliable LLMs for drug discovery, it is necessary to reveal whether and how language models understand the chemical structure of molecules.

To address this problem, we follow the pre-training and fine-tuning paradigm of LLMs: pre-training a transformer language model on chemical molecular data and then fine-tuning this model using reinforcement learning for drug molecular design. We carry out experiments on three design tasks, not only validating the effectiveness of the algorithm but also studying changes in the model’s understanding of chemical language during the fine-tuning process. Specifically, we analyze the learning of molecular spatial structures by the language model from the perspective of molecular fragments. We demonstrate an increase in the quantity and quality of high-frequency SMILES substrings generated by the language model during the fine-tuning process and point out that these substrings correspond well with substructures in the 2D molecular diagrams. These results indicate that the language model for molecular generation

*Corresponding author: Hao Zhang.

understands the spatial structure of molecules, rather than merely fitting the SMILES sequences. Our code is available at: <https://github.com/HXYfighter/LLMsUnderstandMol>.

2 Related Works

In recent years, pharmacology has attracted increasing attention, especially as global public health crises such as the COVID-19 pandemic have demanded greater efficiency in drug development. Consequently, researchers worldwide are committed to applying advanced AI technologies to drug discovery, aiming to tackle practical problems such as drug-target binding prediction, molecular property prediction, retrosynthesis, and drug molecular design (Dara et al. 2022). Although algorithms have seen some performance improvements, they still fall short of meeting the demands of real-world drug discovery, one key challenge being their lack of interpretability.

Drug molecular design is a generative task with the objective of generating a diverse set of drug candidates that fulfill given properties. Given that the chemical space is unstructured and extremely vast (Polishchuk, Madzhidov, and Varnek 2013), drug molecular design is quite challenging. Machine learning techniques have been extensively utilized to address this issue, including genetic algorithms (Jensen 2019; Fu et al. 2022), reinforcement learning (You et al. 2018; Zhou et al. 2019; Gottipati et al. 2020; Ahn et al. 2020; Fu et al. 2021), generative adversarial networks (Guimaraes et al. 2017), flow models (Zang and Wang 2020; Bengio et al. 2021), etc. Common representations for small chemical molecules encompass 1D molecular strings, 2D molecular graphs, and 3D geometries, among which 1D molecular strings are the most suitable for processing and generation using language models (Du et al. 2022).

2.1 Language Models for Drug Design

Significant advancements in the NLP field in recent years have driven rapid updates of drug design methods based on 1D string representation. Currently, the most commonly used 1D string molecular representation is SMILES (Weininger 1988), characterized by its non-necessarily valid strings and the possibility of a single molecular structure to correspond to multiple different SMILES strings. Besides, other 1D sequence molecular representations, including SELFIES (Krenn et al. 2020) and IUPAC (Favre and Powell 2013), are also utilized in drug discovery algorithms.

Reinvent (Olivecrona et al. 2017) was the first to propose the use of reinforcement learning to fine-tune a recurrent neural network (RNN) to accomplish SMILES-based drug design, and it continues to lead in molecular design benchmarks to date (Gao et al. 2022). Subsequently, numerous SMILES-based molecular generation methods were proposed, such as ReLeaSE (Popova, Isayev, and Tropsha 2018) and ChemGE (Yoshikawa et al. 2018), among others. In addition, some algorithms utilize transformer models for molecule generation, including MCMG (Wang et al. 2021), MolGPT (Bagal et al. 2022), Chemformer (Irwin et al. 2022), GMTransformer (Wei et al. 2023), FSM-DDTR (Monteiro et al. 2023), and MolRL-MGPT (Hu et al. 2023).

While much research is dedicated to uncovering the working mechanisms of LLMs, no study has focused on explaining chemical language models. We believe it is necessary to meticulously explore whether and how existing language models model and understand complex relationships between molecular structures and properties. This exploration holds significant implications for further tapping into the potential of language models in drug discovery.

2.2 Fragment-based Drug Design

Fragment-based drug design (FBDD) is a contemporary strategy employed in medicinal chemistry for the discovery and optimization of novel drug compounds, which uses smaller chemical compounds or 'fragments' as starting points in the creation of new drugs. This method offers several potential advantages over traditional ones, including (1) better coverage of the chemical space with fewer molecules, (2) lower combinatorial complexity to build up compounds, and (3) higher hit rates due to the rational design process (Hajduk and Greer 2007; Kirsch et al. 2019).

As a result, the concept of FBDD has also been embraced by AI researchers. For instance, HierG2G (Jin, Barzilay, and Jaakkola 2020a), RationaleRL (Jin, Barzilay, and Jaakkola 2020b), FREED (Yang et al. 2021) RS-VAE (Kong et al. 2022) and MiCaM (Geng et al. 2023) propose different approaches to mine molecular substructures/motifs to generate molecules.

Up until now, no work has attempted to implement fragment-based drug design using language models, nor has any work analyzed drug design language models from the perspective of fragments. Therefore, we aim to fill this gap.

3 Pre-training on Chemical Language

To equip language models with the fundamental capability of processing chemical language, we pre-train them on a dataset of molecules in SMILES representation, aiming to enable the model to comprehend the chemical syntax of SMILES and generate valid SMILES strings. Following MolGPT (Bagal et al. 2022), we construct an autoregressive generative model based on GPT-2 architecture (Radford et al. 2019), which comprises only 6.4M parameters, including eight transformer blocks. For the chemical language of SMILES, we establish a vocabulary consisting of over 100 atomic-level tokens that incorporate atoms, functional groups, numbers, symbols, etc., and build a corresponding tokenizer.

On the ChEMBL dataset (Mendez et al. 2019), which contains over two million drug molecules, we train the aforementioned model unsupervised after filtering out some overly long SMILES strings or those containing rare ions. Specifically, we pre-train the GPT model on the dataset for 10 epochs with a batch size of 4096, and the learning rate is 0.001 with a cosine decay schedule. During the training process, we monitor the change in cross-entropy losses and the valid ratio of SMILES strings generated by the model, as shown in Figure 1. The loss appears to converge, and the valid ratio is stable at higher than 97%, demonstrating that the model is indeed gradually mastering the ability to handle chemical language.

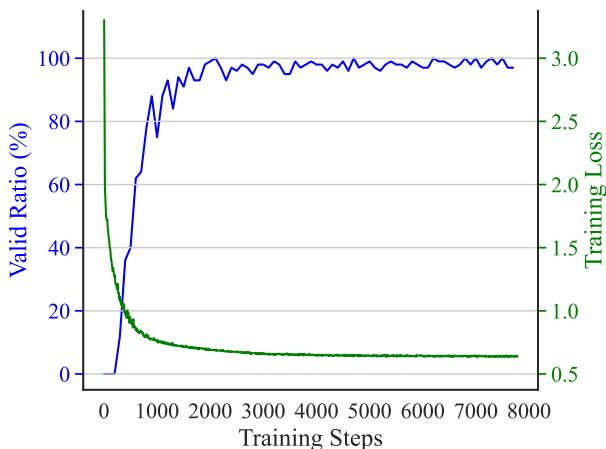


Figure 1: The changing curves of valid ratio and training loss during the pre-training process on chemical language.

It is worth noting that, compared with popular large models in the NLP field, our model is not "large". This is because the number of tokens in the chemical language is far fewer than in natural language, and the length of one "sentence", i.e., a SMILES string, is relatively limited (usually dozens of tokens). Therefore, a smaller model is enough to master chemical syntax without needing a large model with billions of parameters.

4 Fine-tuning for Drug Molecular Design

4.1 RL Fine-tuning

In the task of drug molecular design, once we have obtained a language model capable of generating valid SMILES strings, we hope to fine-tune it to generate molecules that exhibit specific properties, i.e., drug candidates. We use scoring functions (also known as oracles) to evaluate the properties of molecules, such as using docking software to predict whether a molecule can bind well with a given protein target. Generally speaking, there are virtually no compounds satisfying conditions (i.e., *de novo* drug design) in existing datasets, so supervised learning for fine-tuning is not feasible, while reinforcement learning (RL) is currently the most recognized solution for this problem. In an RL fine-tuning process for drug molecular design, the language model is the agent to be fine-tuned; its auto-regressive generation of SMILES tokens represents RL actions, and the scores of molecular properties are RL rewards.

The transformer model pre-trained on the SMILES dataset is adopted to initialize our RL agent. Following Reinvent (Olivecrona et al. 2017), we use REINFORCE algorithm (Williams 1992) to fine-tune the agent at each step with the following loss function:

$$L(x) = [\log P(x, \Theta_{\text{pre}}) - \log P(x, \Theta_{\text{agent}}) + \sigma \cdot s(x)]^2 \quad (1)$$

where $x = "x^0x^1\dots x^n"$ is a sequence of tokens generated by the agent, Θ_{pre} and Θ_{agent} respectively represents the weights of the pre-trained chemical language model and

Algorithm 1: RL fine-tuning algorithm

Input:

(1) Pre-trained weights Θ_{pre} , (2) Scoring function $s(\cdot)$

Hyper-parameters:

(1) Number of fine-tuning steps t , (2) Batch size m , (3) Learning rate η , (4) Coefficient σ

Output:

A set of SMILES strings of high-scoring molecules

Algorithm flow:

- 1: Create an empty memory of high-scoring molecules M
 - 2: Initialize an agent Θ_{agent} with Θ_{pre}
 - 3: **for** $i = 1, 2, \dots, t$ **do**
 - 4: The agent samples a batch of strings x_1, \dots, x_m
 - 5: Using $s(\cdot)$ to score the strings
 - 6: Update M with pairs $(x_k, s(x_k))$, $k = 1, \dots, m$
 - 7: Calculate losses $L(x_k)$, $k = 1, \dots, m$
 - 8: Update Θ_{agent} using losses with the learning rate η
 - 9: **end for**
 - 10: **return** M
-

the weights of the agent, $s(\cdot)$ refers to the scoring function given by the drug design task, and σ is a coefficient hyper-parameter for controlling the term of scores. The likelihood P of generating the sequence x equals the product of the probabilities of generating each token in turn:

$$P(x, \Theta) = \prod_{i=1}^n p(x^i, \Theta | x^0, \dots, x^{i-1}) \quad (2)$$

4.2 Experiments

For experiments, we select three drug rediscovery tasks from the GuacaMol benchmark (Brown et al. 2019) as fine-tuning objectives. The target drug compounds are Celecoxib, Troglitazone, and Thiothixene, respectively (as shown in Figure 3). The scoring function calculates the Tanimoto similarity (Peter et al. 1998) between the molecule corresponding to a generated SMILES string and the target structure, with values ranging from 0 to 1, where a score of 1 signifies successful discovery of the target structure; invalid strings produced receive a score of -1.

For each design objective, we perform 1000 steps of RL fine-tuning, where the batch size sampled in each step is 256, the learning rate is set to 0.0001, and σ is set to 1000.

Figure 2 demonstrates the score curves of the SMILES strings generated by the agent during the RL fine-tuning process. This includes the average score of all strings sampled at each step (**mean score**) and the highest score among the previously discovered molecules (**top-1 score**), which is the official drug design evaluation metric in GuacaMol benchmark. Results show that in the three drug rediscovery tasks, the top-1 scores achieve a full score of 1.0 for both Celecoxib and Troglitazone, indicating successful completion of the task objectives. Although Thiothixene is not successfully rediscovered, the top-1 score reaches 0.72, demonstrating that RL fine-tuning has an effective impact.

Moreover, the mean scores of the sequences produced by the GPT agent also increase during the RL fine-tuning pro-

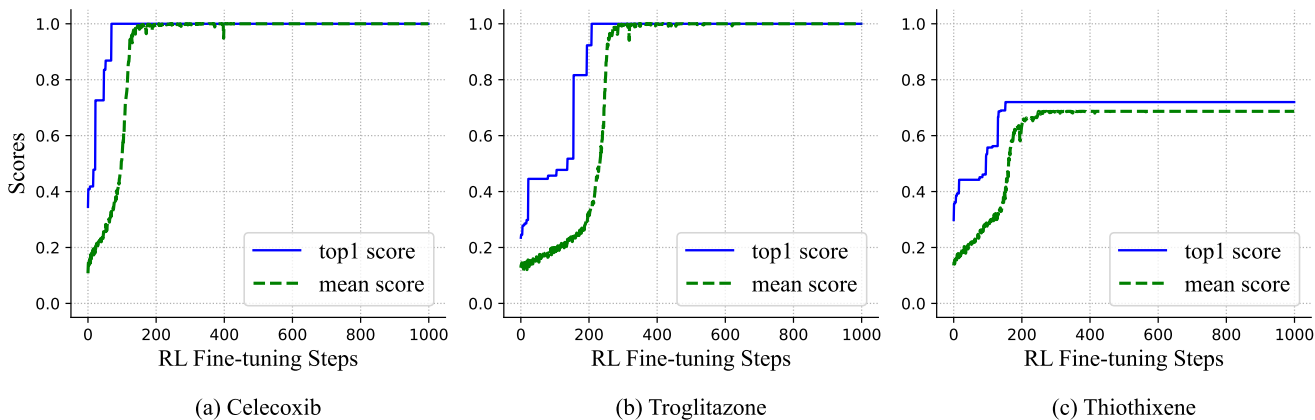


Figure 2: The score curves during the RL fine-tuning processes on three drug design tasks.

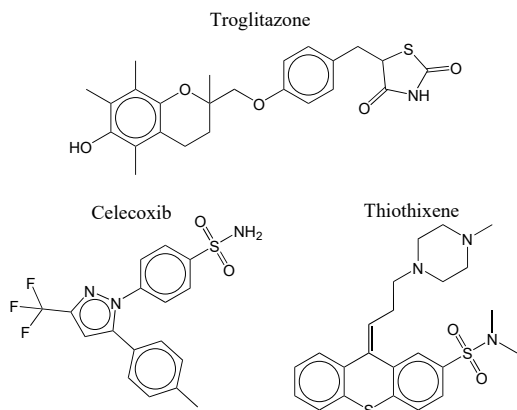


Figure 3: Three target drug structures for rediscovery tasks.

cess and steadily converged to a level close to the top-1 scores. This suggests that the chemical knowledge acquired by the language model indeed transfers from a relatively broad distribution to a small range close to the drug design objectives. Furthermore, this RL-based transferring strategy proved to be applicable to various objectives.

5 Language Models’ Understanding of Molecular Fragments

In this section, we aim to explore the language models’ understanding of chemical structures. Specifically, we seek to verify whether language models, through task-oriented fine-tuning, truly learn spatial structural fragments of the target molecules instead of merely fitting the SMILES sequences. To this end, we employ the SMILES pair encoding algorithm to extract high-frequency substrings from the set of SMILES strings sampled by the language model, which reflect the patterns grasped by the language model toward the task objectives. In the three tasks of drug molecular design, we study the quantity and quality changes of high-frequency substrings during the fine-tuning process; here, quality refers to their efficiency in encoding the SMILES strings of the tar-

get molecules.

5.1 SMILES Pair Encoding

To analyze ”molecular fragments” in chemical language, we use the SMILES Pair Encoding (SPE) algorithm (Li and Fourches 2021) to identify high-frequency substrings in a set of chemical strings. Its principle is derived from the Byte Pair Encoding (BPE) algorithm (Gage 1994), which is commonly used in NLP and is effective in data compression and text tokenization.

Specifically, for a given set of SMILES strings, we first apply SMILES randomization (enumeration) (Bjerrum 2017; Arús-Pous et al. 2019) as a data augmentation, thereby ensuring that adjacency in the chemical space is also reflected in the SMILES set as far as possible. Subsequently, the SPE algorithm identifies all high-frequency substrings in this set through an iterative process. In each iteration step, the occurrence frequencies of all token pairs are counted, and the highest-frequency token pair is merged and encoded into a new token. The iterative process continues until the occurrence frequencies of all token pairs are lower than a pre-defined minimum frequency (MF) hyper-parameter.

It is worth noting that the SMILES substrings recognized by the SPE algorithm do not necessarily perfectly correspond to substructures of molecules, such as ”O)”, ”c1”, etc. This is because SMILES, as a sequence representation, incorporates numbers, parentheses, etc., to encode chemical spatial structures such as cycles. These tokens do not correspond to atoms, chemical bonds, or other chemical entities. However, like punctuation in natural language, they also play significant syntactic roles in chemical language, so they should also be included in our analysis of the chemical language model.

5.2 Model’s Learning of SMILES Substrings

According to the experimental results in Section 4.2, RL fine-tuning effectively transfers the language model’s chemical knowledge toward the neighborhoods of the drug design objectives. Building on this qualitative conclusion, we

Table 1: The canonical and randomized SMILES strings of the three target drug structures for experimental analysis. The colored substrings correspond to the colored substructures in Figure 6.

| Target Drugs | SMILES strings |
|--------------|--|
| Celecoxib | (canonical) <chem>Cc1ccc(-c2cc(C(F)F)nn2-c2ccc(S(N)(=O)=O)cc2)cc1</chem> |
| | rand1 <chem>c1(-c2ccc(C)cc2)n(-c2ccc(S(N)(=O)=O)cc2)nc(C(F)F)c1</chem> |
| | rand2 <chem>c1c(S(N)(=O)=O)ccc(-n2nc(C(F)F)cc2-c2ccc(C)cc2)c1</chem> |
| Troglitazone | (canonical) <chem>Cc1c(C)c2c(c(C)c1O)CCC(C)(COc1ccc(CC3SC(=O)NC3=O)cc1)O2</chem> |
| | rand1 <chem>CC1(COc2ccc(CC3C(=O)NC(=O)S3)cc2)Oc2c(C)c(C)c(O)c(C)c2CC1</chem> |
| | rand2 <chem>c12c(c(C)c(O)c(C)c1C)CCC(C)(COc1ccc(CC3C(=O)NC(=O)S3)cc1)O2</chem> |
| Thiothixene | (canonical) <chem>CN1CCN(CC/C=C/C2/C3CCCC3Sc3ccc(S(=O)(=O)N(C)C)cc32)CC1</chem> |
| | rand1 <chem>c1cc2c(cc1)Sc1c(cc(S(=O)(=O)N(C)C)cc1)/C2=C\CCN1CCN(C)CC1</chem> |
| | rand2 <chem>c1cc2c(cc1)/C(=C/CCN1CCN(C)CC1)c1cc(S(=O)(N(C)C)=O)ccc1S2</chem> |

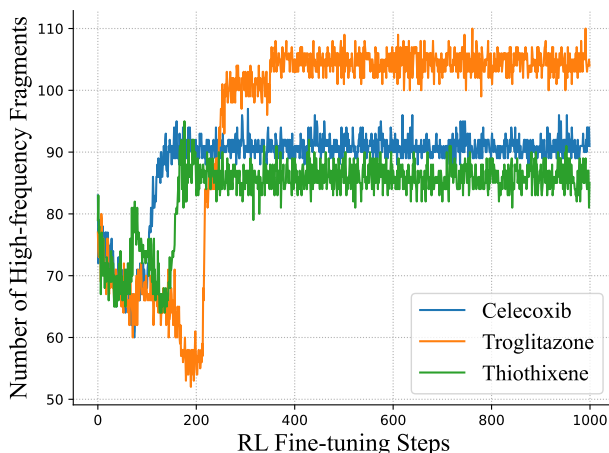


Figure 4: The changing curves of numbers of high-frequency fragments extracted by SPE algorithm during the RL fine-tuning processes on three drug design tasks.

aim to further elucidate what kind of knowledge the language model has learned at the SMILES string level that enables it to generate target molecules more effectively. With the tool of the SPE algorithm, we can quantitatively analyze changes in the language model during RL fine-tuning from the perspective of substrings, which aligns with fragment-based strategy in drug design. Our three experimental tasks of drug rediscovery are well-suited for analysis of molecular structure learning, as our learning objectives themselves are structures rather than properties, making it possible to precisely analyze the relationship between the model’s knowledge and the structures of the target compounds.

Firstly, we report the quantities of high-frequency substrings extracted by the SPE algorithm from the batch of 256 strings sampled at each RL step during the fine-tuning process of the three tasks; here, the hyper-parameter MF is set to 200. As shown in Figure 4, the number of high-frequency substrings generated by the pre-trained model (i.e., the agent at the starting point of the fine-tuning process) is approximately between 70 and 80. However, it stabilizes at a higher level with small fluctuations after the fine-tuning converges. This observation is consistent with intuitive cognition: when

a language model moves to a small range close to the drug design objective, the chemical knowledge it grasps also focuses on a smaller range. Therefore, the generated sequences should have more identical fragments, i.e., high-frequency substrings.

Next, we select three different SMILES strings for each drug structure, one of which is canonical, while the other two are generated through SMILES randomization, as listed in Table 1. To demonstrate the quality of task-oriented chemical knowledge learned by the language model, we monitor changes in the efficiency of generating different SMILES representations of the target molecule during the fine-tuning process. This refers to how many high-frequency substrings (as defined above) are needed to construct the target’s SMILES string. The smaller this number, the more consistent the chemical spatial structure knowledge mastered by the language model is with the drug design objectives.

Figure 5 demonstrates the changing curves of numbers of high-frequency fragments to compose into SMILES strings of drug structures during the RL fine-tuning processes. The following conclusions are drawn from Figure 5:

1. For each task, the number of high-frequency substrings required to constitute different SMILES strings of the target drug molecule converge to a level lower than that of the pre-training model during fine-tuning. This suggests that the language model has acquired knowledge for efficiently generating SMILES strings of the target molecule.
2. For Celecoxib and Troglitazone, the tasks successfully completed, the numbers of high-frequency substrings needed to constitute the SMILES strings of the target drug molecules are less than that for Thiothixene, a task that does not achieve a full score. Furthermore, fewer high-frequency substrings are required to form canonical SMILES compared to randomized SMILES, which could be related to the pre-training model’s better grasp of canonical SMILES.
3. For molecular structures irrelevant to the task objective (here, we use target molecules of other tasks), the number of high-frequency substrings required to form their SMILES strings converge to a level higher than the pre-training model during fine-tuning. This indicates that the

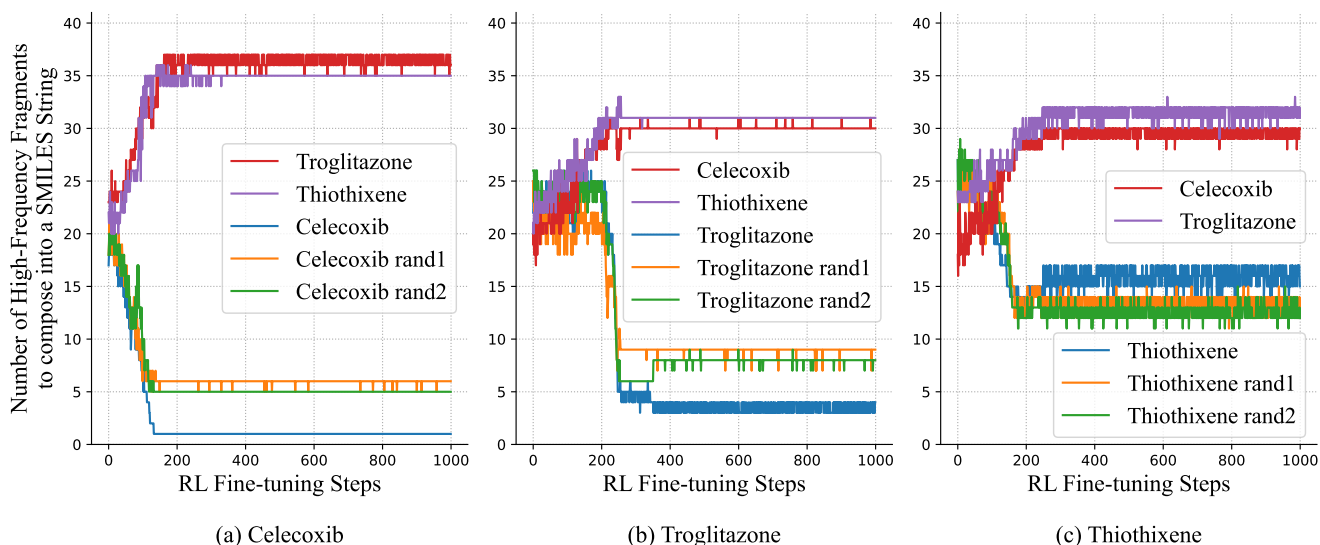


Figure 5: The changing curves of numbers of high-frequency fragments to compose into the given SMILES strings during the RL fine-tuning processes on three drug design tasks.

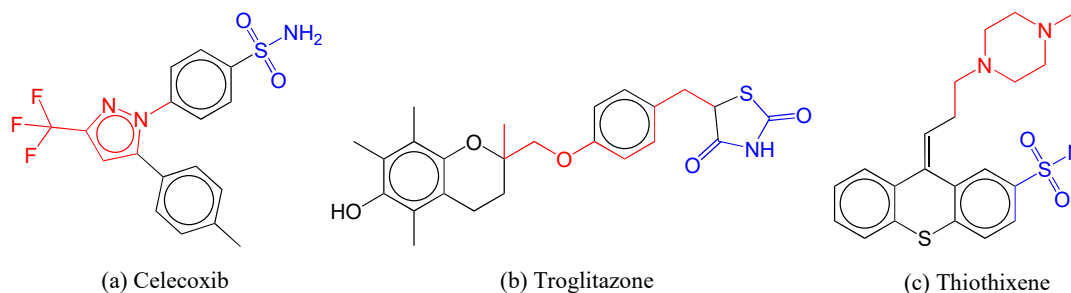


Figure 6: Visualization of substructures in target drug molecules corresponding to high-frequency fragments. The colored substructures correspond to the colored substrings in Table 1.

model “forgets” some chemical knowledge unrelated to the task objectives.

5.3 Model’s Understanding of Molecular Fragments

We have pointed out that during the RL fine-tuning process, the language model learns chemical knowledge related to the task objective. This knowledge is reflected in the high-frequency SMILES substrings within the sequences sampled by the model. These substrings both outnumber and are more consistent with the target molecule’s SMILES compared to the pre-training model. We aim to further investigate the correspondence between these task-related high-frequency SMILES substrings and the 2D fragments of the target molecules.

In Figure 6 and Table 1, we have color-coded the fragments in the 2D graphs of the target molecules and the corresponding substrings in different SMILES strings. These substrings are all high-frequency ones learned by the GPT agent. From this analysis, we can draw the following conclusions:

1. The high-frequency substrings learned by the language model indeed correspond well with the substructures in the 2D molecular graphs, forming wholes in both sequences and the space.
2. Substrings containing several rare tokens like non-carbon atoms and double bonds are learned by the language model as a whole (even in different sequence forms), and correspond to substructures in the 2D molecular graphs. Examples include $S(N)(=O)=O$ in Celecoxib, $SC(=O)NC3=O$ and $C(=O)NC(=O)S$ in Troglitazone, and $S(=O)(=O)N(C)C)cc$ and $S(=O)(N(C)C)=O)cc$ in Thiothixene. These substructures often happen to be key functional groups in drug molecules, playing a decisive role in molecular properties. This indicates that the language model has the potential to learn critical structural features of drug molecules.
3. In some cases, the 2D molecular fragments do not correspond to contiguous token sequences in SMILES strings. However, this does not prevent the language model from

learning corresponding continuous sequences in other SMILES, as exemplified by c2cc(C(F)(F)F)nn2 in Celecoxib (rand1 SMILES) and CCN1CCN(C)CC in Thiothixene (canonical SMILES). This suggests that the language model can overcome the limitations of 1D representations to understand the spatial structure of chemical molecules.

In summary, we believe the language model’s understanding of the chemical spatial structure is superior to that at a 1D string level, because it indeed learns structural knowledge related to fragments of target molecules in different tasks, including their key functional groups.

6 Conclusion

In this study, we pre-train and fine-tune a transformer model on SMILES chemical language for drug molecular design. We pay particular attention to the language model’s understanding of chemical structures during the fine-tuning process, analyze changes in the quantity and quality of high-frequency SMILES substrings generated by the model, and clarify that the language model indeed understands chemical structures from the correspondence between molecular fragments and these substrings.

Our work lays the foundation for better understanding and designing chemical language models and drug molecular design algorithms. Future research directions include but are not limited to: (1) how the language model understands the relationship between molecular structure and properties; (2) fragment-based language models for drug design; (3) the influence of data on chemical language models.

References

- Ahn, S.; Kim, J.; Lee, H.; and Shin, J. 2020. Guiding deep molecular optimization with genetic exploration. *Advances in neural information processing systems*, 33: 12008–12021.
- AI4Science, M. R.; and Quantum, M. A. 2023. The Impact of Large Language Models on Scientific Discovery: a Preliminary Study using GPT-4. *arXiv preprint arXiv:2311.07361*.
- Arús-Pous, J.; Johansson, S. V.; Prykhodko, O.; Bjerrum, E. J.; Tyrchan, C.; Reymond, J.-L.; Chen, H.; and Engkvist, O. 2019. Randomized SMILES strings improve the quality of molecular generative models. *Journal of cheminformatics*, 11(1): 1–13.
- Bagal, V.; Aggarwal, R.; Vinod, P. K.; and Priyakumar, U. D. 2022. MolGPT: Molecular Generation Using a Transformer-Decoder Model. *Journal of Chemical Information and Modeling*, 62(9): 2064–2076.
- Bengio, E.; Jain, M.; Korablyov, M.; Precup, D.; and Bengio, Y. 2021. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems*, 34: 27381–27394.
- Bjerrum, E. J. 2017. SMILES enumeration as data augmentation for neural network modeling of molecules. *arXiv preprint arXiv:1703.07076*.
- Brown, N.; Fiscato, M.; Segler, M. H.; and Vaucher, A. C. 2019. GuacaMol: benchmarking models for de novo molecular design. *Journal of chemical information and modeling*, 59(3): 1096–1108.
- Dara, S.; Dhamecherla, S.; Jadav, S. S.; Babu, C. M.; and Ahsan, M. J. 2022. Machine learning in drug discovery: a review. *Artificial Intelligence Review*, 55(3): 1947–1999.
- Du, Y.; Fu, T.; Sun, J.; and Liu, S. 2022. Molgensurvey: A systematic survey in machine learning models for molecule design. *arXiv preprint arXiv:2203.14500*.
- Favre, H. A.; and Powell, W. H. 2013. *Nomenclature of organic chemistry: IUPAC recommendations and preferred names 2013*. Royal Society of Chemistry.
- Fu, T.; Gao, W.; Coley, C.; and Sun, J. 2022. Reinforced genetic algorithm for structure-based drug design. *Advances in Neural Information Processing Systems*, 35: 12325–12338.
- Fu, T.; Xiao, C.; Glass, L. M.; and Sun, J. 2021. MOLER: incorporate molecule-level reward to enhance deep generative model for molecule optimization. *IEEE transactions on knowledge and data engineering*, 34(11): 5459–5471.
- Gage, P. 1994. A New Algorithm for Data Compression. *C Users Journal*, 12(2): 23–38.
- Gao, W.; Fu, T.; Sun, J.; and Coley, C. 2022. Sample efficiency matters: a benchmark for practical molecular optimization. *Advances in Neural Information Processing Systems*, 35: 21342–21357.
- Geng, Z.; Xie, S.; Xia, Y.; Wu, L.; Qin, T.; Wang, J.; Zhang, Y.; Wu, F.; and Liu, T.-Y. 2023. De Novo Molecular Generation via Connection-aware Motif Mining. In *International Conference on Learning Representations*.
- Gottipati, S. K.; Sattarov, B.; Niu, S.; Pathak, Y.; Wei, H.; Liu, S.; Blackburn, S.; Thomas, K.; Coley, C.; Tang, J.; et al. 2020. Learning to navigate the synthetically accessible chemical space using reinforcement learning. In *International conference on machine learning*, 3668–3679. PMLR.
- Guimaraes, G. L.; Sanchez-Lengeling, B.; Outeiral, C.; Farias, P. L. C.; and Aspuru-Guzik, A. 2017. Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models. *arXiv preprint arXiv:1705.10843*.
- Hajduk, P. J.; and Greer, J. 2007. A decade of fragment-based drug design: strategic advances and lessons learned. *Nature reviews Drug discovery*, 6(3): 211–219.
- Hu, X.; Liu, G.; Zhao, Y.; and Zhang, H. 2023. De novo Drug Design using Reinforcement Learning with Multiple GPT Agents. In *Thirty-seventh Conference on Neural Information Processing Systems*.
- Irwin, R.; Dimitriadis, S.; He, J.; and Bjerrum, E. J. 2022. Chemformer: a pre-trained transformer for computational chemistry. *Machine Learning: Science and Technology*, 3(1): 015022.
- Jensen, J. H. 2019. A graph-based genetic algorithm and generative model/Monte Carlo tree search for the exploration of chemical space. *Chemical science*, 10(12): 3567–3572.

- Jin, W.; Barzilay, R.; and Jaakkola, T. 2020a. Hierarchical generation of molecular graphs using structural motifs. In *International conference on machine learning*, 4839–4848. PMLR.
- Jin, W.; Barzilay, R.; and Jaakkola, T. 2020b. Multi-Objective Molecule Generation using Interpretable Substructures. In *International Conference on Machine Learning*, 4849–4859. PMLR.
- Kirsch, P.; Hartman, A. M.; Hirsch, A. K.; and Empting, M. 2019. Concepts and core principles of fragment-based drug design. *Molecules*, 24(23): 4309.
- Kong, X.; Huang, W.; Tan, Z.; and Liu, Y. 2022. Molecule generation by principal subgraph mining and assembling. *Advances in Neural Information Processing Systems*, 35: 2550–2563.
- Krenn, M.; Häse, F.; Nigam, A.; Friederich, P.; and Aspuru-Guzik, A. 2020. Self-referencing embedded strings (SELF-IES): A 100% robust molecular string representation. *Machine Learning: Science and Technology*, 1(4): 045024.
- Li, X.; and Fourches, D. 2021. SMILES pair encoding: a data-driven substructure tokenization algorithm for deep learning. *Journal of chemical information and modeling*, 61(4): 1560–1569.
- Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; et al. 2019. ChEMBL: towards direct deposition of bioassay data. *Nucleic acids research*, 47(D1): D930–D940.
- Monteiro, N. R.; Pereira, T. O.; Machado, A. C. D.; Oliveira, J. L.; Abbasi, M.; and Arrais, J. P. 2023. FSM-DDTR: End-to-end feedback strategy for multi-objective De Novo drug design using transformers. *Computers in Biology and Medicine*, 164: 107285.
- Olivecrona, M.; Blaschke, T.; Engkvist, O.; and Chen, H. 2017. Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics*, 9.
- OpenAI. 2023. GPT-4 Technical Report. *ArXiv*, abs/2303.08774.
- Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35: 27730–27744.
- Paul, D.; Sanap, G.; Shenoy, S.; Kalyane, D.; Kalia, K.; and Tekade, R. K. 2021. Artificial intelligence in drug discovery and development. *Drug discovery today*, 26(1): 80.
- Peter; Willett; John; M.; Barnard; Geoffrey; M.; and Downs. 1998. Chemical Similarity Searching. *Journal of Chemical Information and Modeling*, 38(6): 983–996.
- Polishchuk, P. G.; Madzhidov, T. I.; and Varnek, A. 2013. Estimation of the size of drug-like chemical space based on GDB-17 data. *Journal of computer-aided molecular design*, 27(8): 675–679.
- Popova, M.; Isayev, O.; and Tropsha, A. 2018. Deep reinforcement learning for de novo drug design. *Science advances*, 4(7): eaap7885.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8): 9.
- Sabe, V. T.; Ntombela, T.; Jhamba, L. A.; Maguire, G. E.; Govender, T.; Naicker, T.; and Kruger, H. G. 2021. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *European Journal of Medicinal Chemistry*, 224: 113705.
- Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; et al. 2019. Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, 18(6): 463–477.
- Wang, J.; Hsieh, C.-Y.; Wang, M.; Wang, X.; Wu, Z.; Jiang, D.; Liao, B.; Zhang, X.; Yang, B.; He, Q.; et al. 2021. Multi-constraint molecular generation based on conditional transformer, knowledge distillation and reinforcement learning. *Nature Machine Intelligence*, 3(10): 914–922.
- Wei, L.; Fu, N.; Song, Y.; Wang, Q.; and Hu, J. 2023. Probabilistic generative transformer language models for generative design of molecules. *Journal of Cheminformatics*, 15(1): 88.
- Weininger, D. 1988. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of chemical information and computer sciences*, 28(1): 31–36.
- Williams, R. J. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Reinforcement learning*, 5–32.
- Yang, S.; Hwang, D.; Lee, S.; Ryu, S.; and Hwang, S. J. 2021. Hit and lead discovery with explorative rl and fragment-based molecule generation. *Advances in Neural Information Processing Systems*, 34: 7924–7936.
- Yoshikawa, N.; Terayama, K.; Sumita, M.; Homma, T.; Oono, K.; and Tsuda, K. 2018. Population-based de novo molecule generation, using grammatical evolution. *Chemistry Letters*, 47(11): 1431–1434.
- You, J.; Liu, B.; Ying, Z.; Pande, V.; and Leskovec, J. 2018. Graph convolutional policy network for goal-directed molecular graph generation. *Advances in neural information processing systems*, 31.
- Zang, C.; and Wang, F. 2020. Moflow: an invertible flow model for generating molecular graphs. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 617–626.
- Zhang, X.; Wang, L.; Helwig, J.; Luo, Y.; Fu, C.; Xie, Y.; Liu, M.; Lin, Y.; Xu, Z.; Yan, K.; et al. 2023. Artificial Intelligence for Science in Quantum, Atomistic, and Continuum Systems. *arXiv preprint arXiv:2307.08423*.
- Zhao, W. X.; Zhou, K.; Li, J.; Tang, T.; Wang, X.; Hou, Y.; Min, Y.; Zhang, B.; Zhang, J.; Dong, Z.; et al. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223*.
- Zhou, Z.; Kearnes, S. M.; Li, L.; Zare, R. N.; and Riley, P. F. 2019. Optimization of Molecules via Deep Reinforcement Learning. *Scientific Reports*, 9.