EPiC Series in Technology

Volume 3, 2026, Pages 218–220

Proceedings of AI for Accelerated Research Symposium

Large Language Models in Drug Discovery: Insights from Reasoning and Planning

Shengchao Liu

Department of Computer Science and Engineering, Chinese University of Hong Kong
scliu@cse.cuhk.edu.hk

Abstract

Recent efforts have explored the use of large language models (LLMs) in drug discovery. As pioneers in this research line, we share our perspective on its current state and where it may lead. Our work, ChatDrug, demonstrates how LLM–human interaction, supported by a domain agent, can improve reliability: when LLMs generate incorrect or invalid outputs, the agent retrieves reference information to guide correction. While ChatDrug enhances answer accuracy and insight generation, a key limitation of current LLMs, their lack of direct perception of the physical world, remains. We believe that overcoming this boundary will require multi-modal tools integrating LLMs with domain-specific capabilities, an important direction for future research.

1 Introduction and Method

Recent advances in large language models (LLMs) have driven significant progress in computer vision and natural language processing, and similar trends are emerging in AI for drug discovery. However, key bottlenecks in this domain remain unresolved. Our research aims to explore these challenges through a series of works leveraging the diverse capabilities of LLMs for drug discovery. Central to our inquiry is a fundamental question: *where is the boundary of LLMs in drug discovery?* Addressing this requires revisiting the core attributes of both LLMs and the nature of drug discovery tasks.

Attributes of language and LLMs Language is often noted for its ambiguity, yet it also possesses remarkable flexibility and an open vocabulary. It can be used to describe virtually any object or property, and new words can be invented (e.g., "selfie", "doomscrolling"), borrowed from other languages, or coined as proper names. This reflects the generativity of language—the ability to create and comprehend an unlimited number of novel words, sentences, and expressions. Generativity, a core property of language, enables humans to combine a finite set of elements into infinitely many new utterances. This capacity underpins the reasoning and planning abilities of LLMs, allowing them to represent and manipulate recursive structures, hypothetical scenarios, and complex multi-step ideas.

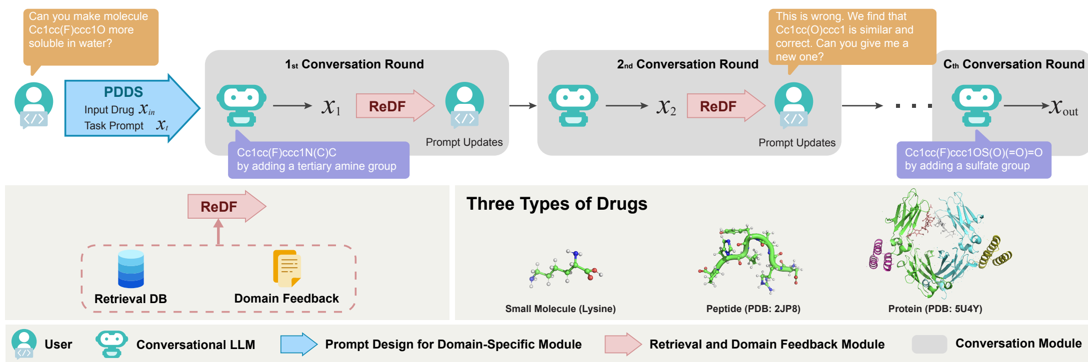

Figure 1: The pipeline for ChatDrug with 3 modules. PDDS generates drug editing prompts. ReDF updates the prompts using retrieved information and domain feedback. Finally, ChatDrug adopts the conversational module for interactive refinement.

Attributes of drug discovery tasks In drug discovery, many important problems are inherently challenging and lack existing solutions. Unlike typical tasks in other AI domains, these problems often have no established ground-truth answers. In AI terminology, valuable drug discovery tasks can be seen as out-of-distribution challenges—aligning with the very notion of "discovery", where the goal is to explore the unknown rather than replicate known solutions.

Connecting LLMs and drug discovery Building on the attributes of LLMs and drug discovery, the next question is: *Why use LLMs for drug discovery?* From the earlier discussion, this can be reframed as: *How can we maximize the utility of LLMs' reasoning and planning abilities to address rare and valuable drug discovery tasks?* This acute scientific insight prompts a closer examination of problem definitions, leading us to the view that **LLMs can play a key role in augmenting drug optimization tasks**. In drug optimization, the goal is to take an input molecule and a target property, and iteratively modify the molecule to achieve the desired properties—a process that inherently requires reasoning and planning.

Motivated by this, we have developed a series of research projects aimed at harnessing the unique strengths of LLMs to advance drug discovery. Among them, ChatDrug [2] is a key contribution. Its core idea is to enable human users to interact with LLMs to obtain desired answers, while introducing a domain agent to intervene when the LLM produces incorrect or invalid outputs. In such cases, the agent searches for and extracts relevant reference information to spark insights that guide correction. The overall pipeline is illustrated in Figure 1.

2 Result and Conclusion

Results and accomplishments In ChatDrug [2], we empirically demonstrate that it achieves the best performance across all 39 drug editing tasks, covering small molecules, peptides, and proteins. Through 10 detailed case studies, we further show that ChatDrug can accurately identify key substructures for manipulation, generating diverse and valid suggestions for drug optimization. Notably, ChatDrug also provides insightful, domain-specific explanations, enhancing interpretability and supporting informed decision-making.

Limitation and next steps LLMs currently lack direct perception of the physical world, which limits their understanding of tasks grounded in 3D Euclidean space. This limitation can lead to critical failures in applications where spatial and structural reasoning is essential. To address this, future work requires a multi-modal framework equipped with a key inter-modality alignment module to supplement the missing information. Certain efforts have already explored encoding geometric structures [1], and a promising direction is to integrate such geometry-aware Transformers into existing LLM-based frameworks.

References

[1] Haorui Li, Weitao Du, Yuqiang Li, Hongyu Guo, and Shengchao Liu. Inertialar: Autoregressive 3d molecule generation with inertial frames. *arXiv preprint arXiv:2510.27497*, 2025.

[2] Shengchao Liu*, Jiongxiao Wang*, Yijin Yang, Chengpeng Wang, Ling Liu, Hongyu Guo*, and Chaowei Xiao*. Conversational drug editing using retrieval and domain feedback. In *ICML SynS and ML Workshop 2023 (Oral); ICLR*, 2024.