SPARK: Harnessing Human-Centered Workflows with Biomedical Foundation Models for Drug Discovery

Bum Chul Kwon , Simona Rabinovici-Cohen , Beldine Moturi , Ruth Mwaura , Kezia Wahome , Oliver Njeru , Miguel Shinyenyi , Catherine Wanjiru , Sekou Remy , William Ogallo , Itai Guez , Partha Suryanarayanan , Joseph Morrone , Shreyans Sethi , Seung-Gu Kang , Tien Huynh , Kenney Ng , Diwakar Mahajan , Hongyang Li , Matan Ninio , Shervin Ayati , Efrat Hexter , Wendy Cornell

IBM Research

bumchul.kwon@us.ibm.com, simona@il.ibm.com, Beldin.Moturi@ibm.com, ruth.mwaura@ibm.com, kezia.wahome@ibm.com, oliver.njeru@ibm.com, miguel.shinyenyi@ibm.com, catherine.wanjiru@ibm.com, sekou@ke.ibm.com, william.ogallo@ibm.com, itai.guez@ibm.com, psuryan@us.ibm.com, jamorron@us.ibm.com, shreyans.sethi@ibm.com, sgkang@us.ibm.com, gdesktop@us.ibm.com, kenney.ng@us.ibm.com, dmahaja@us.ibm.com, hongyang.li@ibm.com, matann@il.ibm.com, sayati@us.ibm.com, efrathex@il.ibm.com, cornell@us.ibm.com

Abstract

Biomedical foundation models, trained on diverse sources of small molecule data, hold great potential for accelerating drug discovery. However, their complex nature often presents a barrier for researchers seeking scientific insights and drug candidate generation. SPARK addresses this challenge by providing a user-friendly, web-based interface that empowers researchers to leverage these powerful models in their scientific workflows. Through SPARK, users can specify target proteins and desired molecule properties, adjust pre-trained models for tailored inferences, generate lists of potential drug candidates, analyze and compare molecules through interactive visualizations, and filter candidates based on key metrics (e.g., toxicity). By seamlessly integrating human knowledge and biomedical AI models' capabilities through an interactive web-based system, SPARK can improve the efficiency of collaboration between human experts and AI, thereby accelerating drug candidate discovery and ultimately leading to breakthroughs in finding cures for various diseases.

1 Introduction

Drug discovery provides the promise of battling diseases that have afflicted humanity, paving the way for new interventions, treatments, and therapies. Recent advancements in biomedical foundation models, fueled by artificial intelligence and vast volumes of chemical and biological data, have opened a new avenue for identifying and evaluating potential drug candidates. For instance, such biomedical foundation models, trained on diverse sources of small molecule data, can be used to infer a list of potential drug candidates given biological targets and predict their physicochemical and biological properties.

However, one of the major challenges in using biomedical foundation models for drug discovery is to seamlessly integrate the capabilities of biomedical researchers and the biomedical AI models. AI models are often regarded as black boxes to human experts due to the intricate neural network architectures, so it is difficult for researchers to comprehend their capabilities and use the models to gain scientific insights [Dey *et al.*, 2022]. Furthermore, it often takes a long-term collaboration among multiple experts to be able to generate initial candidates, evaluate the outcomes, and then iteratively refine them. To achieve this goal, we need to design a human-centered workflow that empowers users to steer model inferences based on their hypotheses and to translate outcomes into meaningful biomedical insights.

In this work, we introduce a web-based system, called SPARK, designed to support user tasks with biomedical foundation models for drug discovery. Our system integrates pre-trained and fine-tuned models with diverse capabilities, various open-source analysis toolkits, and interactive visualizations into a cohesive workflow. It allows researchers to specify target proteins, adjust model parameters, and set inclusion criteria for the generated molecules by combining interactive visualization tools on the frontend with biomedical models on the backend. Our system aims to facilitate an AI-powered, human-centered approach to drug discovery, accelerating research efforts within the AI and biomedical communities. This can ultimately foster the development of advanced AI models for drug discovery and the identification of novel drug candidates, particularly for neglected diseases awaiting effective treatments.

2 Design of SPARK

In this section, we introduce user workflows and the design of individual components of the system. An illustration of the system is displayed in Figure 1. A demonstration video is available: https://vimeo.com/937777860

2.1 User Workflows and System Overview

Before conducting preclinical and clinical development in vivo or in vitro, identifying new drug candidates computa-

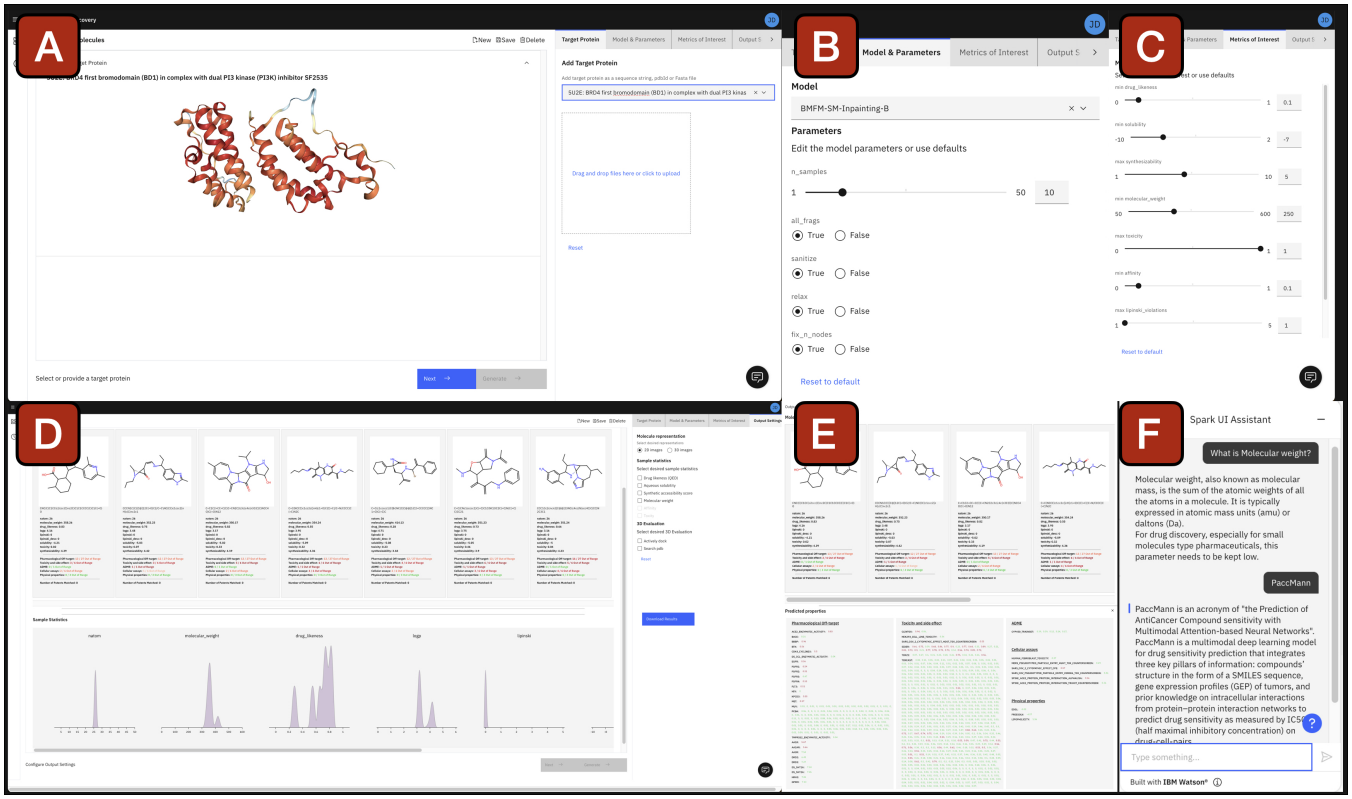

Figure 1: Overview of the system SPARK. (A) Users select target proteins, (B) adjust parameters of models, (C) specify desired ranges for metrics of interest, and (D) analyze generated molecules that fit the criteria. (E) Users can expand the results to view the details of predicted properties and (F) they can ask questions about molecules, chemistry, and small molecule foundation models on a chatbot.

tionally usually consists of two stages: 1) Generating hypotheses: scientists first develop a mechanistic understanding of the disease and identify potential protein targets that could be modulated to treat it; 2) Finding effective molecules: Next, scientists search for molecules that effectively interact with the chosen targets and ultimately cure the disease. Biomedical foundation models can assist in this process by generating molecular candidates for given protein targets. However, the generated candidates may not always be satisfactory. In such cases, users can refine the model parameters to generate new batches of candidates for further evaluation. Specifically, these molecules must also meet specific requirements like low toxicity and high solubility to be viable as pharmaceutical drugs. Fine-tuned models can predict these properties, enabling visualization and detailed analysis to assess suitability. Throughout this process, users may have questions about biomedical concepts, models, and the system itself. SPARK facilitates exploration and decision-making in the computational drug discovery process by integrating biomedical foundation models for generation and prediction with user-friendly interfaces and interactive visualizations.

2.2 Target Protein Selection

To start the drug discovery process, users first decide on one protein from the database. The database includes protein samples that are populated from the Protein Data Bank (PDB). Users may also directly upload their target protein sequences in the FASTA file format. Once users click on one protein, the main view shows the 3-dimensional structure of the protein. The 3D representation of the target protein, implemented via NGL Viewer [Rose and Hildebrand, 2015], allows users to interactively navigate the protein structure using panning and zooming as Figure 1 (A) shows. By using the view, users can inspect the structural characteristics and speculate docking results for small molecule candidates.

2.3 Models & Parameters Refinement

Once the protein target is determined, users move to the next step to choose biomedical foundation models and their parameter conditions. The system integrates several small molecule foundation models, such as PaccMann within GT4SD [Cadow *et al.*, 2020; Manica *et al.*, 2023], a MolFormer-inspired model [Ross *et al.*, 2022; Belgodere *et al.*, 2022] trained on drug-like molecules with diffusion architectures [Schneuing *et al.*, 2022]. To enable API requests from users, we used a microservices architecture, called Caikit open source[1], on the backend to manage the biomedical foundation models. The weights of the previously trained models are stored as checkpoints (files) in the web server. When users select a model on the website, they are prompted to adjust the values of model parameters for inferences as Figure 1 (B) shows. The kinds of parameters vary among models, such as the number of generated samples and temperature. Users can update the parameters by adjusting

[1]https://github.com/caikit/caikit

the number of generated samples from 10 to 50, for instance.

2.4 Metrics of Interest Specification

In the next step, users can specify value ranges for metrics of interest so that the generated molecules meet user-specified inclusion criteria. The metrics of interest include properties of molecules that are important to be used as pharmaceutical drugs, such as affinity, drug-likeness score (QED), lipinski score, synthetic accessibility score, molecular weight, LogP, aqueous solubility, and toxicity, as illustrated in Figure 1 (C). For instance, it is important to estimate the ease of synthesis through synthetic accessibility score so that scientists can virtually screen and prioritize molecular compounds for rapid experiments and drug discovery. The constraints specified with slider input on the user interface are then used to filter the generated molecules that do not meet the conditions.

2.5 Visual Analysis of Generated Molecules and Their Predicted Properties

With all user inputs provided, the biomedical foundation model generates a list of small molecule candidates. Subsequently, the fine-tuned models predict the properties of each generated molecule. The combined results are shown as two visualizations: 1) molecule card; and 2) density chart. The molecule card consists of a 2D atomic representation of the candidate and a summary of its properties (Figure 1 (D)). By using this view in parallel with the target protein, users can investigate structural and atomic characteristics and explore potential binding regions and orientations. The summary view can be expanded by clicking to show individual property values. By using the predicted properties of molecules, users can comparatively assess the efficacy, safety, and synthesizability of the candidates. In this process, the density chart, built with D3.js, can provide a comprehensive overview of all generated molecules with respect to the metrics of interest (Figure 1 (D)). Users can export the information about the generated molecules into various file formats (e.g., json, excel, pdf) for further inspection with third-party tools.

2.6 Chatbot Assistant

The AI-powered drug discovery takes a long, iterative process. The users of the system, scientists and AI developers, collaborate to achieve their shared goals of discovering drug candidates, but they may not necessarily have a deep understanding of each other's domains. To help users ask questions about biology, chemistry, and biomedical models, we implemented a chatbot assistant. The chatbot built with IBM watsonx Assistant[2] provides users with natural language responses to users' queries about technical terms and helpful resources to understand more about them as Figure 1 (E) shows.

3 Use Case: Chagas Disease

This section demonstrates a usage scenario where SPARK can be used to generate and explore small molecule candidates for drug discovery. The insights we describe here are speculative resulting from exploratory data analysis, so they need further scientific validation.

Chagas disease is an infectious disease caused by a parasite found in certain insects. It is a rare disease primarily found in Central and South America. On the first page, users can search for target proteins related to Chagas disease. Previous research has identified that some kinases containing bromodomains as a potential drug target [Pezza *et al.*, 2022]. Among the list, users select a protein target, 5U2E[3], from the list. Once selected, the 3D representation of the protein target is shown to the users. Users zoom in and pan to explore the structure in detail.

Moving on to the next page, users choose models and parameters. Users choose a custom-built biomedical foundation model that was previously implemented on the backend. Here, users set the number of molecule candidates as 10 to generate diverse candidates.

On the subsequent page, users set desired properties for generated molecules through input sliders. Specifically, they update the drug-likeness and synthetic accessibility score to be above minimum threshold levels (0.1 and 5, respectively) so that the generated molecule candidates resemble existing drugs and are predictably easy to synthesize. Setting the constraints on the generated molecules can help bring suitable drug candidates to the market in a rapid manner.

Finally, the system generates the molecules using the foundation model. As Figure 1 (D) shows, the 10 generated molecules display diverse structural patterns. By scanning the predicted properties, one molecule stands out as it appears to include desired properties, especially with measures of cellular assays and physical properties less than the corresponding threshold values. Users click on the card to expand the results. Upon investigation, users realize that the molecule is predicted to have toxicity through multiple measures as Figure 1 (E) shows. Users then go back to the previous step and choose a different model, $PaccMann^{RL}$ [Cadow *et al.*, 2020; Manica *et al.*, 2023], and request to generate 10 molecules with it. The resulting view shows another set of molecules, which shows another set of interesting molecules. Users repeat this process iteratively to get a set of interesting molecules to discuss with other scientists, so they export the results to PDF and share it with them.

4 Conclusion

In this work, we introduce SPARK, a web-based system that empowers researchers to harness the power of biomedical foundation models. By seamlessly integrating AI technologies with intuitive interfaces, SPARK enables users to guide model discovery efforts, visualize and interact with results, and extract valuable scientific insights. This human-centered interactive AI system can serve as a great example of how collaborative intelligence between AI models and human experts can help accelerate real-world advancements in biomedicine, particularly in the exciting domain of small molecule drug candidate generation. Through interactive exploration and iterative refinement, SPARK opens doors for biomedical scientists to leverage cutting-edge AI models alongside their scientific knowledge, ultimately leading to more efficient and potentially groundbreaking discoveries.

[2]https://www.ibm.com/products/watsonx-assistant

[3]https://www.rcsb.org/structure/5u2e

Acknowledgments

We would like to thank the Cleveland Clinic-IBM Discovery Accelerator partnership for their support and express our appreciation to the anonymous reviewers for their thoughtful and detailed feedback on this paper.

References

[Belgodere *et al.*, 2022] Brian Belgodere, Vijil Chenthamarakshan, Payel Das, Pierre Dognin, Toby Kurien, Igor Melnyk, Youssef Mroueh, Inkit Padhi, Mattia Rigotti, Jarret Ross, et al. Cloud-based real-time molecular screening platform with molformer. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 641–644. Springer, 2022.

[Cadow *et al.*, 2020] Joris Cadow, Jannis Born, Matteo Manica, Ali Oskooei, and María Rodríguez Martínez. Paccmann: a web service for interpretable anticancer compound sensitivity prediction. *Nucleic acids research*, 48(W1):W502–W508, 2020.

[Dey *et al.*, 2022] Sanjoy Dey, Prithwish Chakraborty, Bum Chul Kwon, Amit Dhurandhar, Mohamed Ghalwash, Fernando J. Suarez Saiz, Kenney Ng, Daby Sow, Kush R. Varshney, and Pablo Meyer. Human-centered explainability for life sciences, healthcare, and medical informatics. *Patterns*, 3(5), May 2022.

[Manica *et al.*, 2023] Matteo Manica, Jannis Born, Joris Cadow, Dimitrios Christofidellis, Ashish Dave, Dean Clarke, Yves Gaetan Nana Teukam, Giorgio Giannone, Samuel C Hoffman, Matthew Buchan, et al. Accelerating material design with the generative toolkit for scientific discovery. *npj Computational Materials*, 9(1):69, 2023.

[Pezza *et al.*, 2022] Alejandro Pezza, Luis E. Tavernelli, Victoria L. Alonso, Virginia Perdomo, Raquel Gabarro, Rab Prinjha, Elvio Rodríguez Araya, Inmaculada Rioja, Roberto Docampo, Felix Calderón, Julio Martin, and Esteban Serra. Essential Bromodomain TcBDF2 as a Drug Target against Chagas Disease. *ACS Infectious Diseases*, 8(5):1062–1074, May 2022.

[Rose and Hildebrand, 2015] Alexander S. Rose and Peter W. Hildebrand. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Research*, 43(W1):W576–W579, 2015.

[Ross *et al.*, 2022] Jerret Ross, Brian Belgodere, Vijil Chenthamarakshan, Inkit Padhi, Youssef Mroueh, and Payel Das. Large-scale chemical language representations capture molecular structure and properties. *Nature Machine Intelligence*, 4(12):1256–1264, 2022.

[Schneuing *et al.*, 2022] Arne Schneuing, Yuanqi Du, Charles Harris, Arian Jamasb, Ilia Igashov, Weitao Du, Tom Blundell, Pietro Lió, Carla Gomes, Max Welling, et al. Structure-based drug design with equivariant diffusion models. *arXiv preprint arXiv:2210.13695*, 2022.