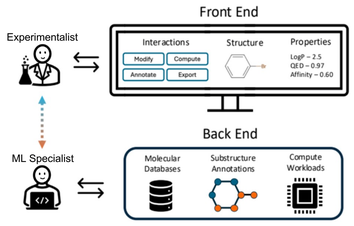
Super-Glue: Interactive Web Tool to Facilitate Chemistry <> ML Collaboration

EXPEDITION MEDICINES

Overview

Background and Motivation



Machine learning (ML) is transforming Medicinal Chemistry enabling rapid pharmaceutical development. However, it lacks the intuition honed by expert chemists through decades of knowledge acquisition. Super-Glue aims to "glue" together expert knowledge with ML, enhancing model performance, fostering cross-domain trust, and integrating AI into existing drug discovery pipelines.

Overall Goal: To create an interactive web tool to facilitate the collaboration of expert medicinal chemistry knowledge with ML in a scalable and data-driven manner.

Objectives: Aggregated User Stories (from over 20)

- **Experimentalist** exploring drug discovery with limited coding experience, seeking a user-friendly interface to **visualize** molecules, **highlight** key functional groups, and leverage AI for **property prediction** based on existing data.
- **ML specialist** proficient in predictive modeling but new to cheminformatics, seeking a curated, **annotated molecular dataset** in a **structured format** that integrates seamlessly with **Python workflows**.

Technologies Used

Frontend

- Ketcher[1] – a web-based chemical structure editor that is open-source and free to use
- React
- Vite
- Ant Design
- 3Dmol.js[2] – enables 3-D molecular modeling

Backend

- Flask – web framework that is simple and easy to learn and offers high flexibility
- RDKit[3] – open-source toolkit for cheminformatics
- Unittest
- ADMET_ai[4] – Python package that provides fast and accurate ADMET predictions

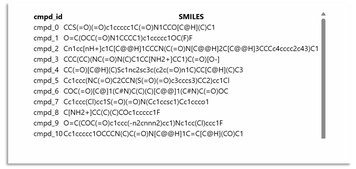
Design Flow

Home Page

- Login, **upload** a .csv file, and view previously uploaded files.

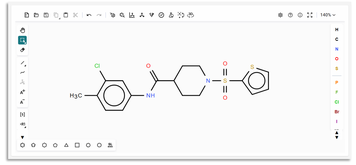
Summary Page

- After uploading, **preview** the dataset in a tabular view.
- Toggle to view **all saved highlights** across all molecules.



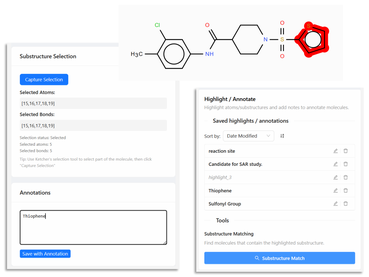
Molecular Visualization

- Molecules are rendered with **Ketcher**, enabling real-time interaction with structures directly in the browser.



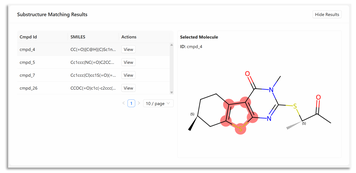
Annotation + Highlighting

- **Highlight atoms or substructures, add annotations** for reactions or key features.
- Saved highlights are listed in a table for easy review.



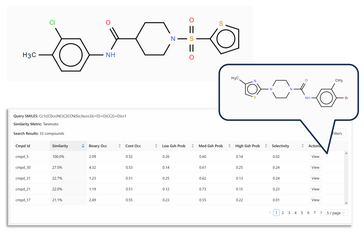
Substructure Matching

- Search for molecules with **matching substructures** using saved highlights.
- Results include a table with **highlighted regions** for quick pattern recognition.



Similarity Search

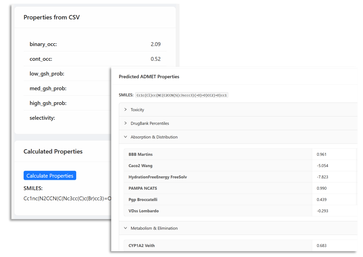
- Identify **structurally similar compounds** using various similarity **metrics.**
- Adjust **thresholds** and view ranked results with similarity scores.



Design Flow

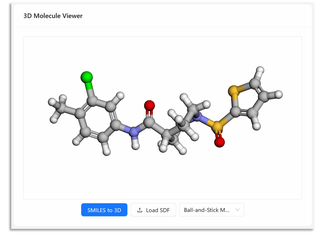
Property Computations

- Calculate simple properties like **logP** with **RDKit**.
- Run drug-likeness predictions using **ADMET_ai**.



3-D Modeling

- Visualize molecules in an **interactive 3D viewer** – rotate, zoom, explore.
- Upload **SDF Files** for accurate geometries, or auto-generate from SMILES.



Conclusion

Main Takeaways

- Super-Glue allows users to upload their molecular data, annotate their findings, find similar molecules, and predict complex properties crucial drug development,
- Super-Glue is a platform that combines AI with human intuition, by fostering trust, enabling collaboration, and optimizing drug discovery

Future Work

- Containerization for cloud deployment
- We plan to expand similarity search to larger databases like PubChem for greater accuracy and utility
- Finetuning of ADMET using client specific data for generative AI

References

1. Ketcher: Web-based chemical structure editor. https://github.com/epam/ketcher/releases/tag/v3.2.0, doi: https://doi.org/10.1186/1758-2946-3-S1-P3.
2. Nicholas Rego and David Koes, "3Dmol.js: molecular visualization with WebGL", Bioinformatics, 31(8): 1322-1324, 2015, doi: https://doi.org/10.1093/bioinformatics/btu829.
3. RDKit: Open-source cheminformatics. https://www.rdkit.org, doi: https://doi.org/10.5281/zenodo.591637.
4. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. https://github.com/swansonk14/admet_ai, doi: https://doi.org/10.1093/bioinformatics/btae416

https://github.com/SuperGlue2025/SuperGlue2025

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

CHEMICAL ENGINEERING
UNIVERSITY of WASHINGTON

ADVISORS: Dr. Mahdi Horbani, Dr. Orion Dollar, Dr. David Beck

SPONSOR: Expedition Medicines