Open-Source Developments for Data Driven Kinase-Centric Drug Design

Andrea Volkamer, Data Driven Drug Design, Saarland University

Human protein kinases play a significant role in numerous diseases, making them a crucial protein family for targeted therapy. To date there are over 6,000 human kinase structures in the PDB and around 70 small molecule kinase inhibitors available. However, challenges such as drug promiscuity, resistance, and unexplored kinase territory persist.

In this presentation, we will provide an overview of methods that leverage openly available kinase data to generate new insights and facilitate community engagement, with a focus on machine learning (ML) tasks. Utilizing the TeachOpenCADD platform [1,2], we will demonstrate how diverse computer-aided drug design (CADD) tasks can be orchestrated for individual kinases. Additionally, we will introduce freely available tools to support kinase research, including: (i) KinFragLib for fragment-based kinase inhibitor design [3], (ii) KiSSim – a KLIFS-based kinase structural similarity fingerprint [4], and (iii) a pipeline for assessing kinase similarity from various data perspectives [5].

Lastly, we will present ongoing projects in structure-informed ML for kinase inhibitor design and affinity prediction across kinases [6,7], employing deep learning techniques.

[1] S. Dominique, *et al.*, TeachOpenCADD 2022: Open Source and FAIR Python Pipelines to Assist in Structural Bioinformatics and Cheminformatics Research. *Nucleic Acids Research,* **2022**. https://doi.org/10.1093/nar/gkac267

[2] M. Backenköhler, *et al.,* TeachOpenCADD goes Deep Learning: Open-source Teaching Platform Exploring Molecular DL Applications. *ChemRxiv,* **2023**. https://doi.org/10.26434/chemrxiv-2023-kz1pb

[3] S. Dominique, *et al.*, KinFragLib: Exploring the Kinase Inhibitor Space Using Subpocket-Focused Fragmentation and Recombination. *Journal of Chemical Information and Modeling*, **2020**. https://doi.org/10.1021/acs.jcim.0c00839

[4] S. Dominique, *et al.*, KiSSim: Predicting Off-Targets from Structural Similarities in the Kinome. *Journal of Chemical Information and Modeling*, **2022**. https://doi.org/10.1021/acs.jcim.2c00050

[5] T.B. Kimber, *et al.*, Kinase Similarity Assessment Pipeline for Off-Target Prediction [v1.0]. *Living Journal of Computational Molecular Science*, **2022**. https://doi.org/10.33011/livecoms.3.1.1599

[6] D. Schaller, C.D. Christ, J.D. Chodera, A. Volkamer. Benchmarking Cross-Docking Strategies for Structure-Informed Machine Learning in Kinase Drug Discovery. bioRxiv., 2023. https://doi.org/10.1101/2023.09.11.557138

[7] M. Backenköhler, J. Groß, V. Wolf, A. Volkamer, Guided docking as a data generation approach facilitates structure-based machine learning on kinases. ChemRxiv, **2023**. https://doi.org/10.26434/chemrxiv-2023-prk53