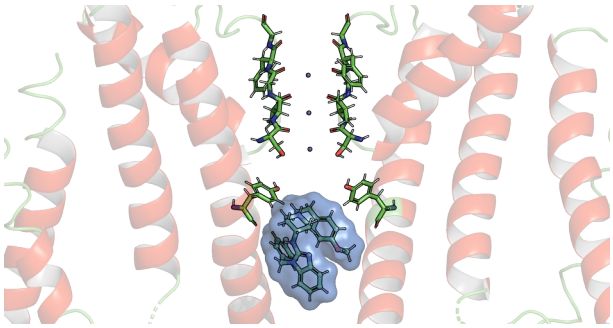
A combined deep learning and structure based cheminformatic approach to understand ligand blockage activity on the hERG channel

Nathaniel Smith, David Bushiri Pwesombo, Han Sun

FMP-Berlin



The hERG (human *Ether-à-go-go*-Related Gene) encoded potassium channel Kv11.1 is highly expressed in heart tissue and is well-known for its contribution to the repolarisation of the cardiac action potential. Dysfunctional or drug inhibition of the hERG channel may cause prolongation of the QT interval (long QT syndrome), which leads to arrythmia, such as *torsades de pointes.* A number of common drugs have been shown to bind hERG channel, causing unwanted side effects. Therefore, tremendous effort has been put into assessing hERG-related cardiotoxicity in the preclinical stages of drug discovery.

The recent advancements in machine learning (ML) techniques have enabled the use of powerful cheminformatic approaches to predict the activities of small molecule ligands. However, generally a large amount of labeled dataset is required for training. The hERG channel is particularly appealing for ML-based approaches, as a high number of blockers have been previously identified from the screening of large collections of chemical libraries. Furthermore, the recent high-resolution cryo-EM structures of hERG channel [1,2] has enabled several structure-based studies on hERG inhibition.[3] In the current study, we aim to combine deep learning and structure-based cheminformatics approaches to understand ligand blockage activity on the hERG channel. We chose 298,016 molecules labeled for hERG inhibition from the hERGcentral database as the training dataset, where we calculated the molecular descriptors (PaDEL descriptors).[4] Different ML approaches, including deep neural networks (DNN) and graph neural networks (GNN), were trained on the aforementioned dataset to predict their activities. For training of GNN models, molecules were represented as graphs. We compared the accuracy obtained from different models. The positive outcomes of ML will be rationalised in the near future using molecular docking and molecular dynamics simulation. Furthermore, we envision extending this approach to identify small molecule modulators for other potassium channels where limited experimental data exist.

[1] Wang, W. & MacKinnon, R. Cryo-EM structure of the open human ether-à-go-go-related K+ channel hERG. Cell 169,(**2017**), 422–430

[2] Asai, T., Adachi, N., Moriya, T., Oki, H., Maru, T., Kawasaki, M., et al., Cryo-EM structure of K(+)- bound hERG Channel Complexed With The Blocker astemizole. Structure, (**2021**), 29 203–12 e4.

[3] Creanza T. M, Delre P, Ancona N, Lentini G, Saviano M, Mangiatordi G. F. Structure-Based Prediction of HERG-Related Cardiotoxicity: A Benchmark Study. J. Chem. Inf. Model. (**2021**), 61, 4758−4770

[4] Du F, Yu H, Zou B, Babcock J, Long S, Li M. hERGCentral: a large database to store, retrieve, and analyze compound-human Ether-à-go-go related gene channel interactions to facilitate cardiotoxicity assessment in drug development. Assay Drug Dev Technol. (**2011**), 9(6):580-8. PMID: 22149888, PMCID: PMC3232635.