

# Towards a structural characterization of Pan Assay Interference Compounds

*Maksim Koptelov Albrecht Zimmermann Bruno Crémilleux*

GREYC (CNRS UMR 6072), Université de Caen Normandie, Caen, France

Virtual screening is a critical step of a drug design process whose success can be compromised by the selection of pan-assay interference compounds (PAINS) as leads. The goal of our work is the early detection of as many real PAINS as possible in order to exclude them from the set of drug candidates before further development, thus minimizing money and time spent on drug development.

We propose a new method for predicting candidates as being PAINS (or not) in an input set of molecules with reliable confidence. It combines resampling, discriminative subgraph mining, numerical features related to chemical structure, and ideas from QSAR. The method relies on the identification of chemical fragments whose occurrences correlate with a PAIN profile.

The method has been implemented, tested, and verified. It shows good accuracy and recall depending on the classifier employed. The results are promising, the method proposed has some limitations which can be improved during the future work, however.

In addition to the model, a tool with a graphic user interface has been designed: it allows a chemist to classify a chemical compound, and to explore the related model (see Figure 1).

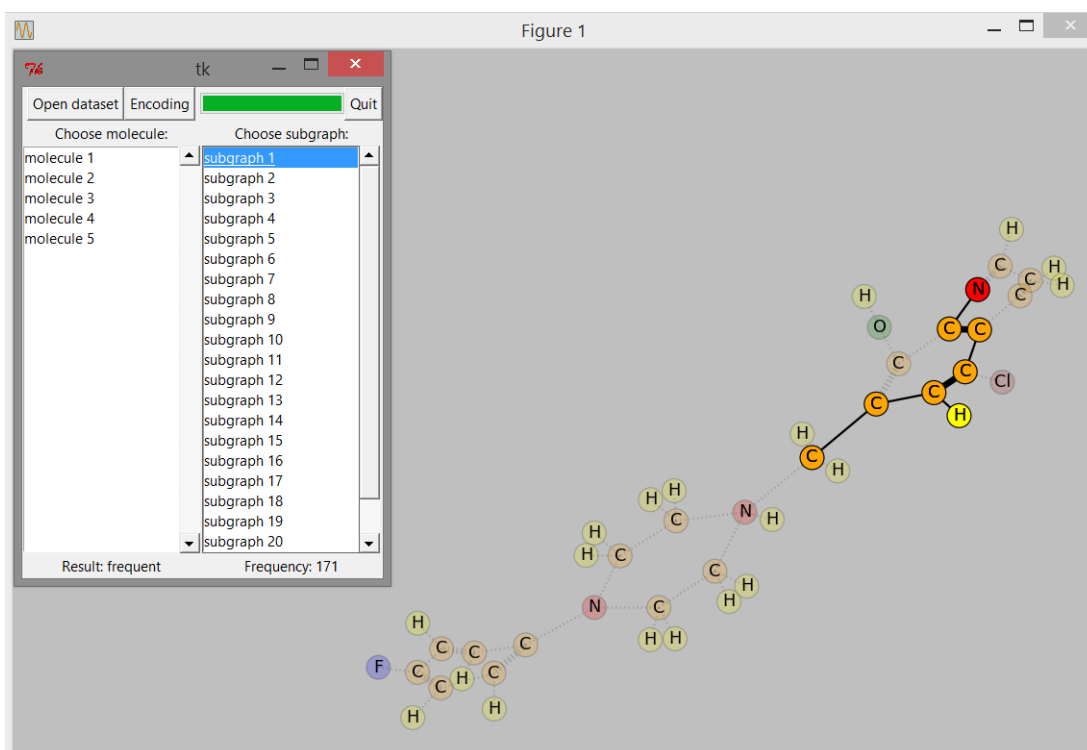


Figure 1: A view of the user interface