In Silico Prediction of Pharmaceutical Cocrystal Formation Using Explainable Machine Learning

Filip Vretenar and Željka Ujević Andrijić

fvretenar@fkit.unizg.hr

University of Zagreb Faculty of Chemical Engineering and Technology, Marulićev trg 19, Zagreb, Croatia

Pharmaceutical cocrystals are multicomponent crystalline systems composed of an active pharmaceutical ingredient (API) and a coformer, assembled by noncovalent forces. The design of such cocrystals has become an important strategy for improving the physicochemical properties of APIs, such as solubility, chemical stability, and bioavailability [1]. In general, screening of new cocrystals involves experimental identification with a large number of coformers, thus being time-consuming, laborious and expensive. The application of computational (*in silico*) methods for the prediction of cocrystal formation, such as machine learning (ML), can significantly reduce the time and cost required for cocrystal screening [2]. Classification ML models can be used to predict whether a given API–coformer combination will form a cocrystal, providing useful preliminary information for the selection of potential coformers. In this study, a classification ML model was developed using experimental data reported in the literature. Molecular descriptors were generated from the SMILES representations of APIs and coformers using the RDKit toolkit. These descriptors quantitatively describe molecular properties and are used as input features for the ML model. Although ML models can be useful for *in silico* cocrystal screening, they are often considered black-box models with limited interpretability. The lack of interpretability can reduce trust in model predictions and limit the extraction of meaningful scientific insights [3]. To address this limitation, explainable artificial intelligence (XAI) methods are used to provide insight into model predictions and help identify the key factors influencing cocrystal formation.

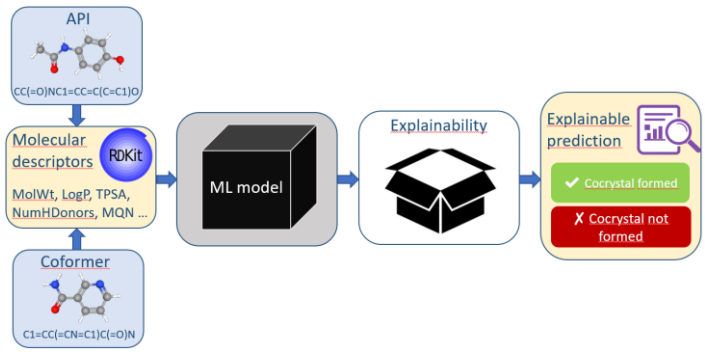

Figure 1. Workflow of explainable machine learning classification for cocrystal prediction

References:

[1] R. Birolo, F. Bravetti, E. Alladio, E. Priola, G. Bianchini, R. Novelli, A. Aramini, R. Gobetto, M. R. Chierotti, *Cryst. Growth Des.* **23** (2023) 7898–7911.

[2] Y. Song, Y. Ding, J. Su, J. Li, Y. Ji, *Angew. Chem. Int. Ed.* **64** (2025) e202502410.

[3] F. Oviedo, J. L. Ferres, T. Buonassisi, K. T. Butler, *Acc. Mater. Res.* **3** (2022) 597–607.