

Reviewer's report of the habilitation thesis by

M.Sc. Pavlo Polishchuk, Ph.D.

New methodologies of Machine-Learning modeling of complex chemical systems: mixtures, reactions and ligand-protein complexes

Reviewer: Doc. RNDr. Petr Jurečka, Ph.D.

Although I am not an expert in the field of bioinformatics, I accepted the invitation to review this habilitation thesis knowing that the candidate has already been awarded the HDR (Habilitation à diriger des recherches) title by the respected University of Strasbourg, France. This allowed me to read the work with a sense of confidence and curiosity, focusing on the areas that were new and enriching to me. Indeed, I found the thesis to be a compelling, well-written, and substantial contribution that is both rich in content and beneficial to the field.

The habilitation thesis is written as a comprehensive, self-contained text with references to works published by the author. The published articles have appeared in high-quality scientific journals in the fields of cheminformatics or computational chemistry. The list of publications is respectably long, and Pavlo Polishchuk frequently appears as the first or corresponding author, particularly in works related to the main topics of this habilitation. In my opinion, this demonstrates a strong scientific record of the candidate.

I will not spend much time here describing the main achievements of the author. The thesis presents a coherent research program that revolves around computer-aided drug design and addresses several of its aspects, such as extension of QSAR/SAR methodology to **mixtures and chemical reactions**, then a **universal, structure-level interpretation** of QSAR models with an accompanying benchmark, **3D pharmacophore representation**, retrieval and modeling, **multi-instance (multi-conformer) learning** for ligand-protein modeling and **de novo molecular design** via chemically reasonable mutations (CReM), supported by applied case studies. The unifying theme is to identify weak points in the current computational drug design protocols and replace them with physically inspired, open, reproducible solutions. This is a systematic, focused and productive body of work that, in my view, warrants recognition at the habilitation level.

After reviewing the thesis submitted by Pavlo Polishchuk, entitled **New methodologies of Machine-Learning modeling of complex chemical systems: mixtures, reactions and ligand-protein complexes**, I have to conclude that it meets the requirements for a habilitation work, and I recommend it as a basis for awarding the title *docent (associate professor)* in Physical Chemistry.

Questions for the habilitation thesis defense:

1) The Universal Interpretation Approach (UIA) suggested for better QSAR interpretability appears truly universal, useful, and a strong basis for further development. However, one question comes to my mind. How does UIA treat stereochemistry of the drug molecules? Also, could the

masking procedure itself alter the stereochemistry of a compound and thereby affect the result? If masking does alter stereochemistry, is there a risk that we step outside the domain the QSAR model was trained for?

2) Multi-instance (multi-conformer) learning is undoubtedly a very important element of drug design workflows. In your experience, how reliable are conformer-generating tools like RDKit? The most stable conformer *in vacuum* can differ from that in solvent. How is this accounted for? Do you think there is a room for improvement in conformer generation?

In Olomouc, on October 20, 2025

doc. RNDr. Petr Jurečka, Ph.D.