How Soluble Are Metal Complexes?

Nadine BENZIADA[1], Luka Praprotnik Čurič[1], Fabio Golini[1], Artem Lennoi[1], Thanh-An Pham[1], Igor V. Tetko[2], Dragos Horvath[3], João Aires de Sousa[4]

[1] *Master de Chemoinformatics, University of Strasbourg*

[2] *Tetko Lab, Institute of Structural Biology, Helmholtz Munich, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.*

[3] *Laboratory of Chemoinformatics, UMR 7140 CNRS, University of Strasbourg, Institut Le Bel, 4 rue Blaise Pascal, 67081 Strasbourg, France.*

[4] *LAQV-REQUIMTE, Department of Chemistry, NOVA School of Science and Technology, Universidade NOVA de Lisboa, Caparica, Portugal.*

Solubility is a key property controlling the performance of metal complexes in pharmaceutical, catalytic, and energy-storage applications. However, unlike organic molecules, metal complexes lack large, standardized, and machine-learning-ready solubility datasets [1].

This project addresses this gap by building a thermodynamically consistent data infrastructure for predicting metal-complex properties. A transfer-learning strategy was proposed to investigate whether knowledge learned from organic solubility data can support the prediction of metal-complex solubility-related properties, while evaluating the respective roles of ligand environment and metal identity.

A complementary three-pillar dataset was collected from different sources: a curated organic solubility dataset from AqSolDB [2] was used as the source domain for future model pre-training. In parallel, NIST SRD46 [3] was ported into a queryable relational database containing metal-ligand thermodynamic records. Because direct solubility data for metal complexes are scarce, metal-ligand stability constants were used as thermodynamic proxies. A literature-derived corpus of stability constants was assembled, covering multiple metal centers and ligand families with ML-ready datasets.

Overall, this work provides a data infrastructure and a reproducible foundation for future machine-learning models of metal-complex solubility and stability constants.

Bibliography :

[1] N. Mousa; H.P. Varbanov; V. Kaipanchery; E. Gabano; M. Ravera; A.A. Toropov; L. Charochkina; F. Menezes; G. Godin; I.V. Tetko. *J. Inorg. Biochem.* **269** (2025) 112890.
[2] P. Llompart; C. Minoletti; S. Baybekov; D. Horvath; G. Marcou; A. Varnek. *Sci. Data* **11** (2024) 303.
[3] D.R. Burgess. *NIST SRD 46: Critically Selected Stability Constants of Metal Complexes*, NIST (2004).