Digital Driving Exploration of

Metal-Organic Supramolecular Architectures

Yishuo Chen[1], Filip T. Szczypiński[1,2], Jean-François Ayme*[1], Andrew I. Cooper*[1]

[1]*Leverhulme Research Centre for Functional Materials Design and Materials Innovation Factory, University of Liverpool, Liverpool, UK.*

[2]*Department of Chemistry, Durham University, Durham, UK*

Abstract Text :

Owing to their unique three-dimensional structures, metal–organic supramolecular architectures have found applications in molecular recognition, sensing, catalysis, biology and the stabilisation of reactive species. However, the intricate self-assembly processes that govern the formation of these architectures make their rational design challenging and largely serendipitous, their discovery still relies heavily on inefficient trial and error. Combining high-throughput experimentation (HTE) with statistical and machine learning (ML) methods offers a clear strategy for mapping and rationalising the metal–organic architecture chemical space, enabling the formalisation of design principles and accelerating the discovery of novel structures.

Building on our previous work on the autonomous synthesis of imine-based metal–organic supramolecular architectures,[1] we developed a semi-automated synthetic and analytical workflow, together with an autonomous decision-making protocol, to systematically map the reaction landscape arising from 11 diamines, 15 heteroaromatic aldehydes and 9 transition-metal salts. The high-quality dataset generated (comprising both success and failure labels) will be used to train ML models that predict the likelihood of forming an imine-based metal–organic architecture from a given set of building blocks (one amine, one aldehyde one metal salt per set), and to decipher the underlying design rules of this chemistry.

To inform the predictive models, we developed a wide range of molecular descriptors tailored specifically to supramolecular chemistry. This approach transfers a methodology that has proved highly successful in organometallic catalysis to the previously underexplored domain of supramolecular self-assembly.[2] Specifically, a high-throughput density functional theory (DFT) pipeline was established to evaluate organic building blocks and model metal complexes through conformational searches, geometry optimisations and single-point energy calculations. A custom Python workflow was then used to extract electronic, geometric and global molecular descriptors and compile them into a structured database. By drawing on these chemically meaningful features, the resulting machine learning models are expected to offer significantly enhanced predictive power and physical interpretability.

Bibliography :

[1] T. Dai, S. Vijayakrishnan, F. T. Szczypiński, J.-F. Ayme, E. Simaei, T. Fellowes, R. Clowes, L. Kotopanov, C. E. Shields, Z. Zhou, J. W. Ward, A. I. Cooper, *Nature* **2024**, 635, 890-897

[2] T. Gensch, G. dos Passos Gomes, P. Friederich, E. Peters, T. Gaudin, R. Pollice, K. Jorner, A. Nigam, M. Lindner-D'Addario, M. S. Sigman, A. Aspuru-Guzik, *J. Am. Chem. Soc.* **2022**, 144, 1205-1217