S05-OC02

ACCELERATED CATALYST DISCOVERY VIA UNSUPERVISED MACHINE LEARNING AND SPECIFIC DESCRIPTOR CALCULATIONS

I. Funes-Ardoiz[a]

[a] *Departamento de química, Centro de Investigación en Síntesis Química, Universidad de La Rioja. C/ Madre de Dios, 53, 26004, Logroño, España*

e-mail: Ignacio.funesa@unirioja.es

Keywords: *Unsupervised machine learning, DFT calculations, Speciation, Catalysis*

Speciation of chemical compounds is one of the hardest challenges in homogeneous catalysis.[1] Determining the nature of the catalyst structure, as well as the oxidation state, nuclearity and other factors, is key for reactivity, efficiency and selectivity. This challenge is not simple and usually required complex experimentation (*in situ* NMR, x-ray characterization, etc.) and complementary computational studies. In this context, Pd(I) dimer catalysis has emerged as an alternative and easy to handle catalyst to classical Pd(0)/Pd(II) catalysts for cross-coupling reactions.[2] However, identifying which ligands favors Pd(I) dimer vs Pd(0) species is far from being an easy task, as the driving force, determined by DFT calculations, strongly favors the dimer form in all the cases, but experimentally, only a very specific phosphine ligands yield it. Then, we hypothesized that many factors should be involved and due to the limited access to experimental dimers (only five were reported in the literature) and the possibility of computing a large library of ligands, we identify unsupervised machine learning as a powerful tool to achieve this goal.

Herein, we report a new workflow for Pd(I)-dimer ligand discovery based on unsupervised machine learning (k-means), that consists of two consecutive analyses: the first one is based on a large phosphine database (LKB-P) that was used to map the general trend of ligands in Pd(I) dimer formation. The second one was a refinement of the group of interest via specific descriptor calculations and a second *k*-means analysis. By these means, 21 ligands were predicted and the whole process was verified by the synthesis of new 8 representative new Pd(I) dimers, including one ligand that was never synthesized before.[3]

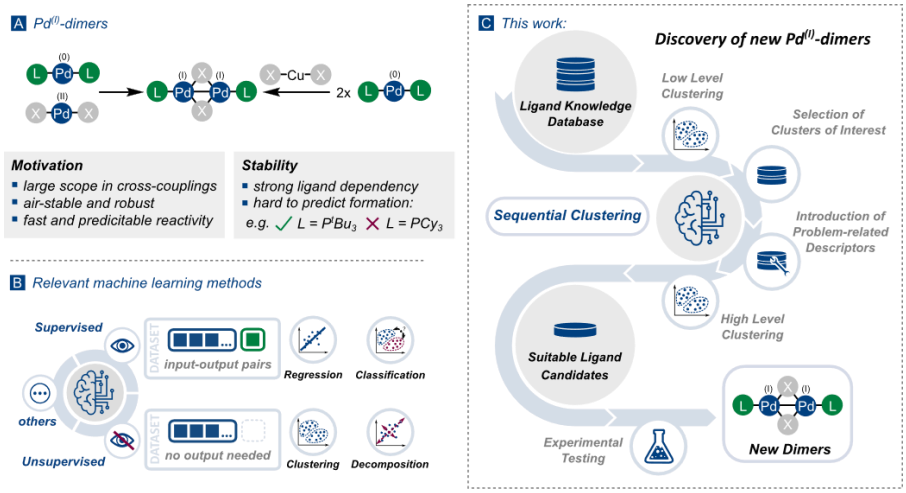

Figure 1. Unsupervised machine learning workflow for Pd(I) dimer discovery.

References

[1] I. Funes-Ardoiz, F. Schoenebeck, *Chem*, **6**, 1904-1913.

[2] I. Kalvet, K. Deckers, I. Funes-Ardoiz, G. Magnin, T. Sperger, M. Kremer, F. Schoenebeck, *Angew. Chem. Int. Ed.* **2020**, *59*, 7721-7725.

[3] J. A. Hueffel, T. Sperger, I. Funes-Ardoiz, J. S. Ward, K. Rissanen, F. Schoenebeck, *Science*, **2021**, *374*, 1134-1140