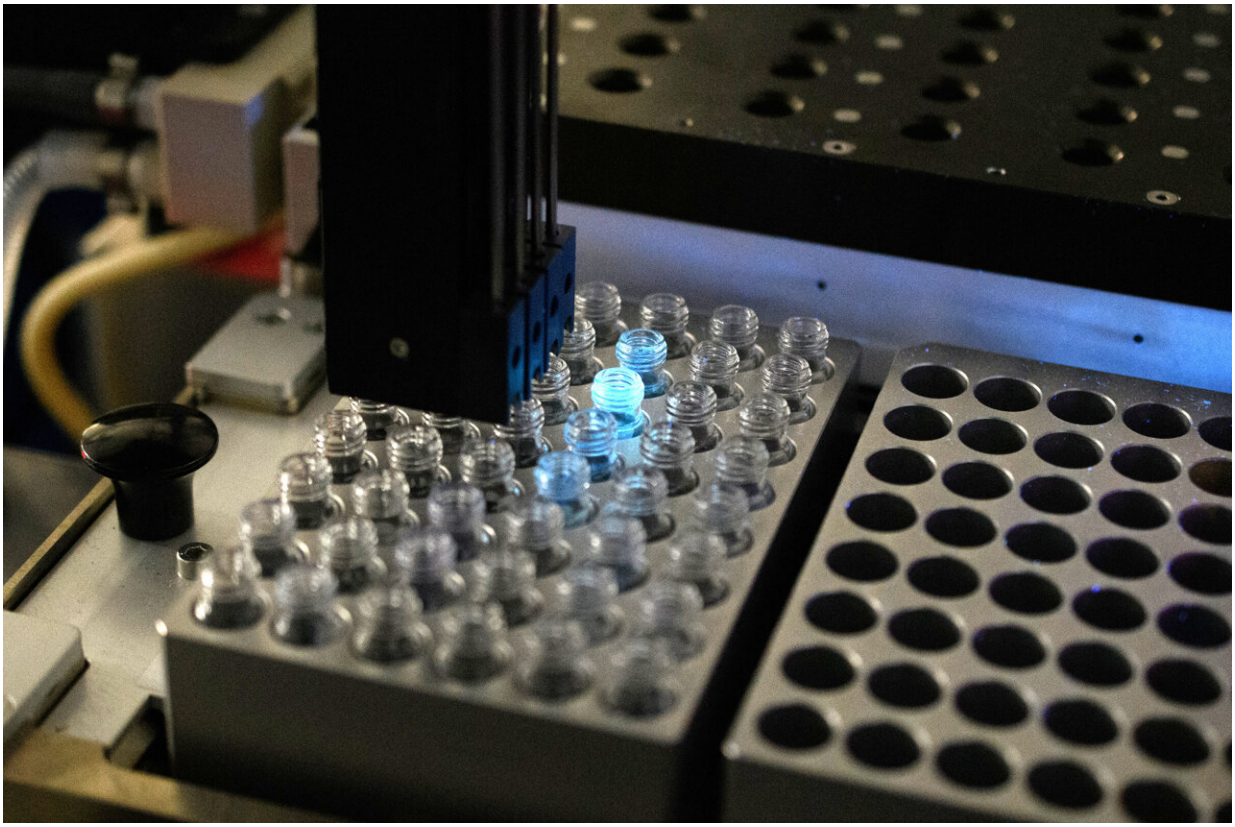
Robots and AI team up to discover highly selective catalysts

February 3 2023



Close up of the semi-automated synthesis robot used to generate training data. Credit: ICReDD

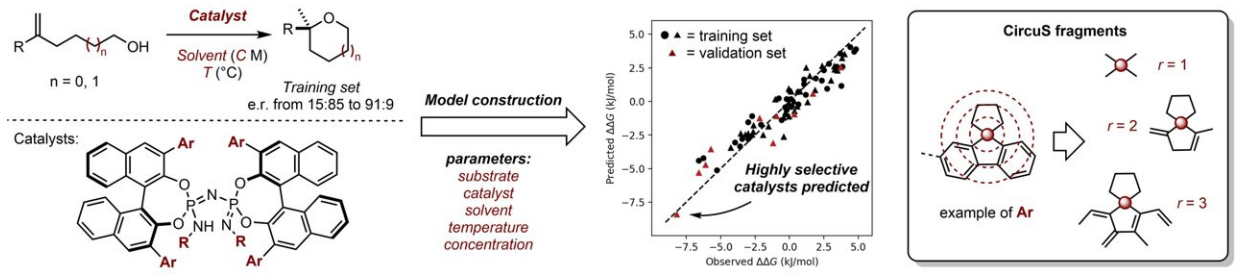
Researchers used a chemical synthesis robot and computationally cost-effective AI model to successfully predict and validate highly selective

catalysts. Their results are published in the journal *Angewandte Chemie International Edition*.

Artificial intelligence (AI) has made headlines recently with the advent of ChatGPT's language processing capabilities. Creating a similarly powerful tool for chemical reaction design remains a significant challenge, especially for complex catalytic reactions. To help address this challenge, researchers at the Institute for Chemical Reaction Design and Discovery and the Max Planck Institut für Kohlenforschung have demonstrated a machine learning method that utilizes advanced yet efficient 2D chemical descriptors to accurately predict highly selective asymmetric catalysts—without the need for quantum chemical computations.

"There have been several advanced technologies which can 'predict' catalyst structures, but those methods often required large investments of calculation resources and time; yet their accuracy was still limited," said joint first author Nobuya Tsuji. "In this project, we have developed a predictive model which you can run even with an everyday laptop PC."

For a computer to learn chemical information, molecules are usually represented as a collection of descriptors, which often consist of small parts, or fragments, of those molecules. These are easier for AI to process and can be arranged and rearranged to construct different molecules, much like Lego pieces can be arranged and connected in different ways to construct different structures.



Fast and robust predictive models using 2D descriptors particularly suited for asymmetric catalysis. Highly selective catalysts were predicted and validated using training data with only moderate selectivities. Credit: Nobuya Tsuji, Pavel Sidorov, et al. *Angewandte Chemie International Edition*. January 23, 2023

However, computationally cheaper 2D descriptors have struggled to accurately represent complex catalyst structures, leading to inaccurate predictions. To improve this issue, researchers developed new Circular Substructure (CircuS) 2D descriptors that explicitly represent cyclic and branched hydrocarbon structures, which are common in catalysis. Training data for the AI was obtained through experiments via a streamlined, semi-automatic process utilizing a synthesis robot. This experimental data was then converted into descriptors and used to train the AI model.

Researchers used the fully trained model to virtually test 190 catalysts not part of the training data. In this set, the AI model was able to predict highly selective catalysts after only having been trained on the data of catalysts with moderate selectivity, showing an ability to extrapolate beyond the training data. The catalyst predicted to have the highest selectivity was then tested experimentally, exhibiting a selectivity nearly identical to that predicted by the AI model. Obtaining high selectivity is especially crucial for the design of new medicines, and this technique provides chemists with a powerful framework for optimizing selectivity that is efficient in both computational and labor cost.

"Often, to predict new selective catalysts chemists would use models based on quantum chemical calculations. However, such models are

computationally costly, and when the number of compounds and the size of molecules increases, their application becomes limited," commented joint first author Pavel Sidorov. "Models based on 2D structures are much cheaper and therefore can process hundreds and thousands of molecules in seconds. This allows chemists to filter out the compounds they may not be interested in much more quickly."

More information: Nobuya Tsuji et al, Predicting Highly Enantioselective Catalysts Using Tunable Fragment Descriptors, *Angewandte Chemie International Edition* (2023). DOI: 10.1002/anie.202218659

Provided by Hokkaido University