

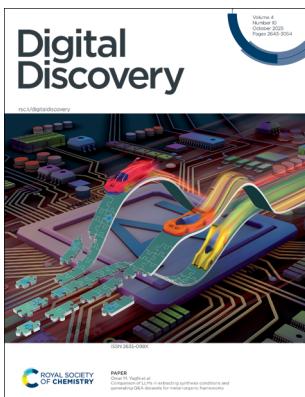
# Digital Discovery

rsc.li/digitaldiscovery

The Royal Society of Chemistry is the world's leading chemistry community. Through our high impact journals and publications we connect the world with the chemical sciences and invest the profits back into the chemistry community.

## IN THIS ISSUE

ISSN 2635-098X CODEN DDIIAI 4(10) 2643–3054 (2025)



### Cover

See Omar M. Yaghi et al., pp. 2676–2683. Image reproduced by permission of Omar M. Yaghi from *Digital Discovery*, 2025, 4, 2676. Image created by Yuang Shi and Wanrongmiao Zhang.



### Inside cover

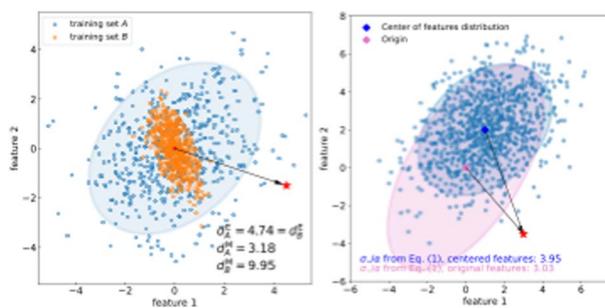
See Jerret Ross, Payel Das et al., pp. 2684–2696. Image reproduced by permission of IBM Corporation from *Digital Discovery*, 2025, 4, 2684. Image created by Samuel Hoffman.

## PERSPECTIVE

2654

### Uncertainty in the era of machine learning for atomistic modeling

Federico Grasselli,\* Sanggyu Chong, Venkat Kapil, Silvia Bonfanti and Kevin Rossi\*

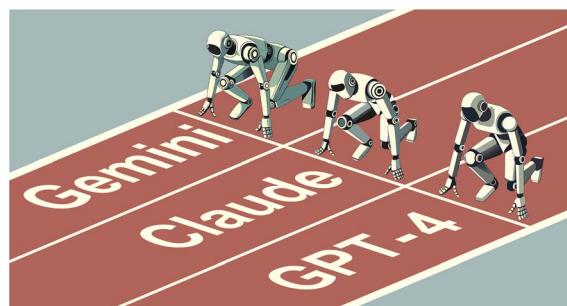


## PAPERS

2676

### Comparison of LLMs in extracting synthesis conditions and generating QA datasets for metal–organic frameworks

Yuang Shi, Nakul Rampal, Chengbin Zhao, Dongrong Joe Fu, Christian Borgs, Jennifer T. Chayes and Omar M. Yaghi\*



# RSC Advances

**At the heart of open access for  
the global chemistry community**

**Editor-in-chief**

**Russell J Cox**

Leibniz Universität Hannover, Germany

**We stand for:**



**Breadth** We publish work in all areas of chemistry and reach a global readership



**Quality** Research to advance the chemical sciences undergoes rigorous peer review for a trusted, society-run journal



**Affordability** Low APCs, discounts and waivers make publishing open access achievable and sustainable



**Community** Led by active researchers, we publish quality work from scientists at every career stage, and all countries

**Submit your work now**

[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

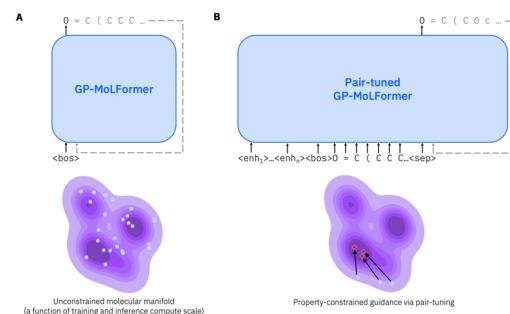
@RSC\_Adv

## PAPERS

2684

**GP-MoLFormer: a foundation model for molecular generation**

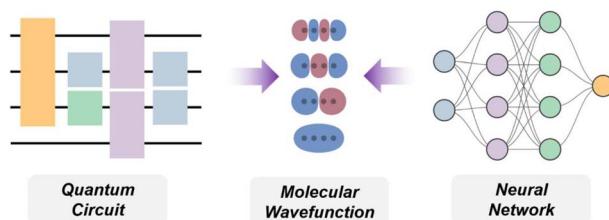
Jerret Ross,\* Brian Belgodere, Samuel C. Hoffman, Vijil Chenthamarakshan, Jiri Navratil, Youssef Mroueh and Payel Das\*



2697

**Quantum machine learning of molecular energies with hybrid quantum-neural wavefunction**

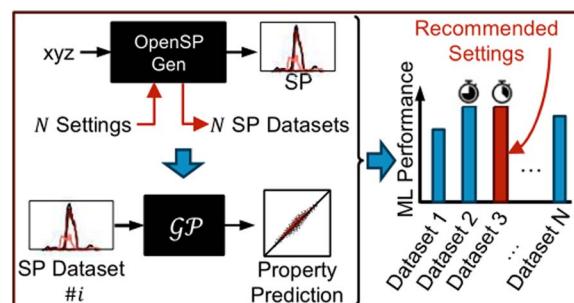
Weitang Li,\* Shi-Xin Zhang, Zirui Sheng, Cunxi Gong, Jianpeng Chen and Zhigang Shuai



2711

**Open-source generation of sigma profiles: impact of quantum chemistry and solvation treatment on machine learning performance**

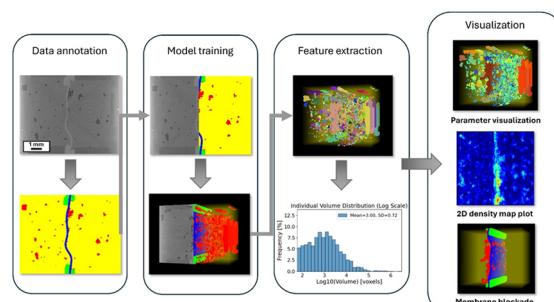
Fathyah Y. M. Salih, Dinis O. Abrantes, Edward J. Maginn and Yamil J. Colón\*



2724

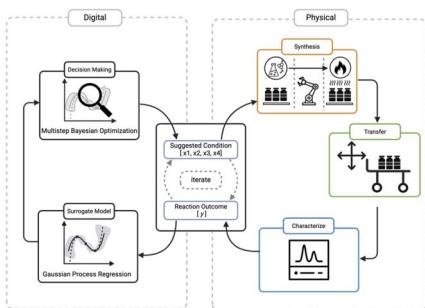
**Advancing vanadium redox flow battery analysis: a deep learning approach for high-throughput 3D visualization and bubble quantification**

André Colliard-Granero, Kangjun Duan, Roswitha Zeis, Michael H. Eikerling, Kourosh Malek and Mohammad J. Eslamibidgoli\*



## PAPERS

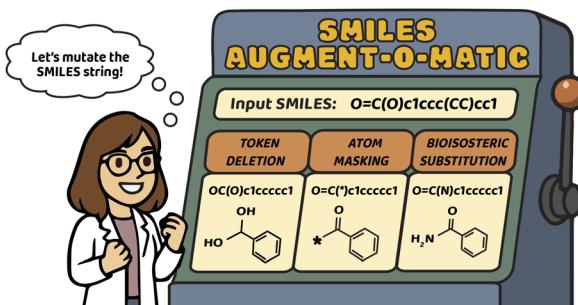
2737



## Autonomous organic synthesis for redox flow batteries *via* flexible batch Bayesian optimization

Clara Tamura, Heather Job, Henry Chang, Wei Wang, Yangang Liang\* and Shijing Sun\*

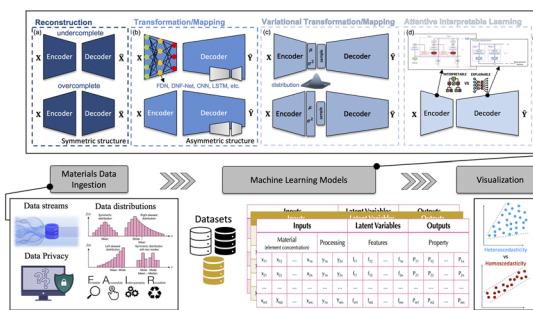
2752



## Going beyond SMILES enumeration for data augmentation in generative drug discovery

Helena Brinkmann, Antoine Argante, Hugo ter Steege and Francesca Grisoni\*

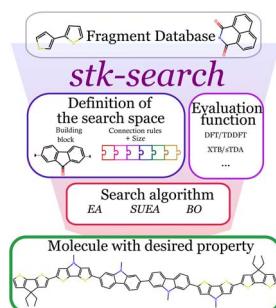
2765



## Decoding non-linearity and complexity: deep tabular learning approaches for materials science

Vahid Attari\* and Raymundo Arroyave

2781



## Comparative analysis of search approaches to discover donor molecules for organic solar cells

Mohammed Azzouzi,\* Steven Bennett, Victor Posligua, Roberto Bondesan, Martijn A. Zwijnenburg and Kim E. Jelfs\*

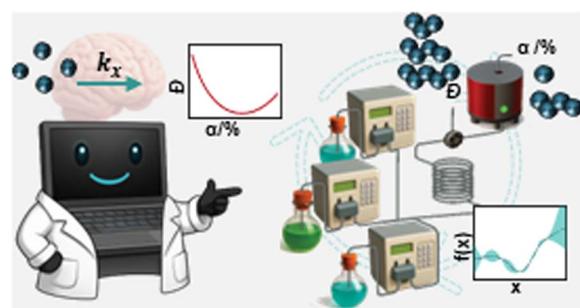


## PAPERS

2797

**Integrating mechanistic modelling with Bayesian optimisation: accelerated self-driving laboratories for RAFT polymerisation**

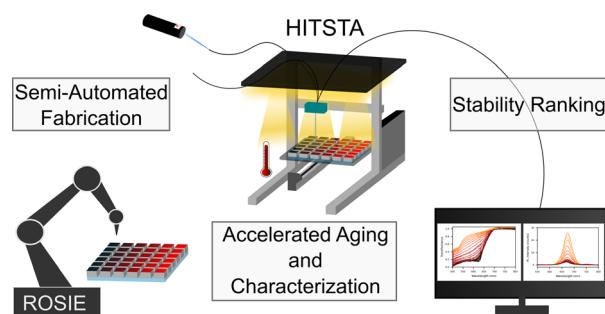
Clarissa Y. P. Wilding,\* Richard A. Bourne and Nicholas J. Warren\*



2804

**Accelerating optimization of halide perovskites: two blueprints for automation**

Hilal Aybike Can,\* Daniel Anthony Jacobs, Nicolas Fürst, Christophe Ballif and Christian Michael Wolff\*



2816

**Computer vision for polymer characterisation using lasers**

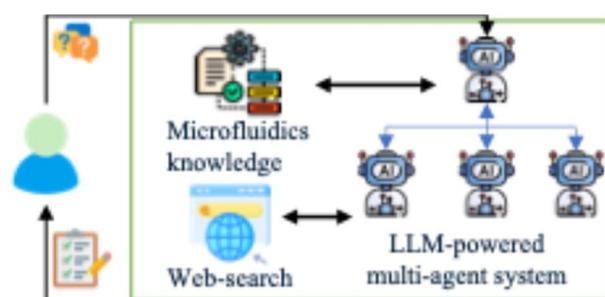
Seda Uyanik, Sam Parkinson, George Killick, Biplab Dutta, Rob Clowes, Charlotte E. Boot\* and Andrew I. Cooper\*



2827

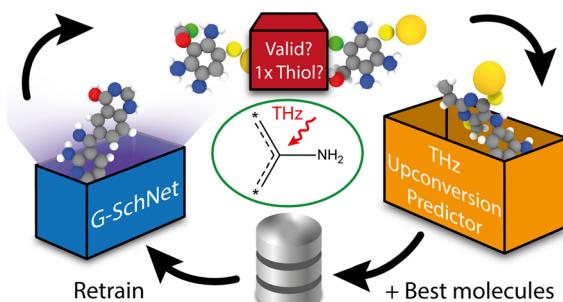
**DropMicroFluidAgents (DMFAs): autonomous droplet microfluidic research framework through large language model agents**

Dinh-Nguyen Nguyen, Raymond Kai-Yu Tong and Ngoc-Duy Dinh\*



## PAPERS

2852

**Generative design of functional organic molecules for terahertz radiation detection**

Zsuzsanna Koczor-Benda,\* Shayantan Chaudhuri, Joe Gilkes, Francesco Bartucca, Liming Li and Reinhard J. Maurer\*

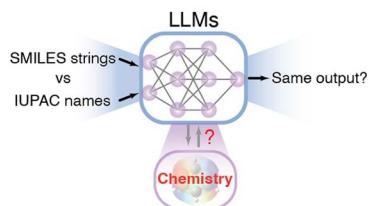
2864

**An open-source peristaltic pump with multiple independent channels for laboratory automation**

Michael Buchhorn, Gun Deniz Akkoc and Dominik Dworschak\*

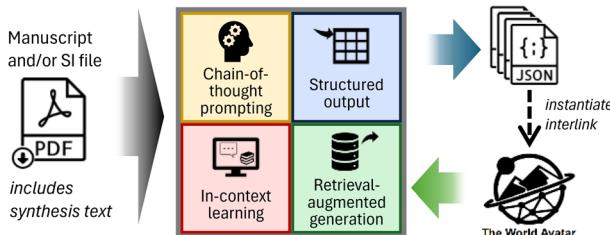
2876

Do LLMs understand **chemistry** behind symbolic representations?

**Inconsistency of LLMs in molecular representations**

Bing Yan, Angelica Chen and Kyunghyun Cho\*

2893

**Information extraction****Extraction of chemical synthesis information using the World Avatar**

Simon D. Rihm, Fabio Saluz, Aleksandar Kondinski, Jiaru Bai, Patrick W. V. Butler, Sebastian Mosbach, Jethro Akroyd and Markus Kraft\*

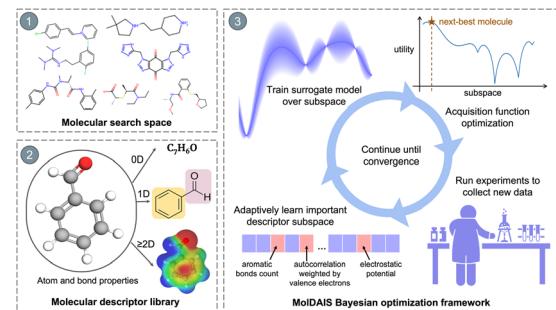


## PAPERS

2910

**Adaptive subspace Bayesian optimization over molecular descriptor libraries for data-efficient chemical design**

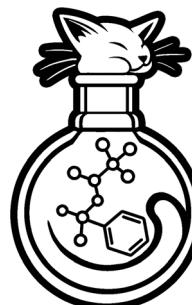
Farshud Sorourifar, Thomas Banker and Joel A. Paulson\*



2927

**Moment of inertia as a simple shape descriptor for diffusion-based shape-constrained molecular generation**

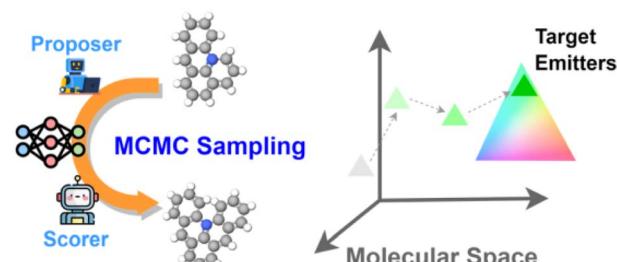
Denis Sapegin,\* Fedor Bakharev, Dmitriy Krupenya, Azamat Gafurov, Konstantin Pildish and Joseph C. Bear\*



2942

**Generative AI-powered inverse design for tailored narrowband molecular emitters**

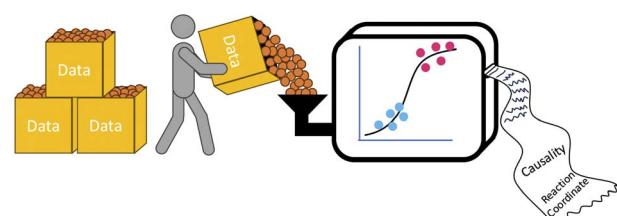
Mianzhi Pan, Tianhao Tan, Yawen Ouyang, Qian Jin, Yougang Chu, Wei-Ying Ma, Jianbing Zhang,\* Lian Duan,\* Dong Wang\* and Hao Zhou\*



2954

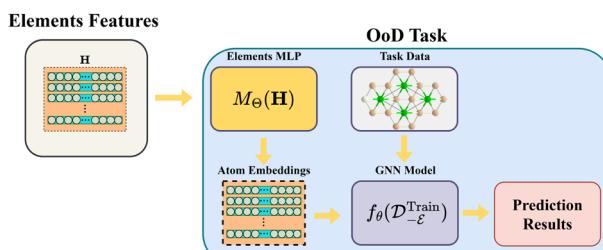
**Coupling causality and interpretable machine learning to reveal the reaction coordinate of C–N coupling with a supramolecular Cu-calix[8]arene catalyst**

R. A. Talmazan, J. Gamper, I. Castillo, T. S. Hofer\* and M. Podewitz\*



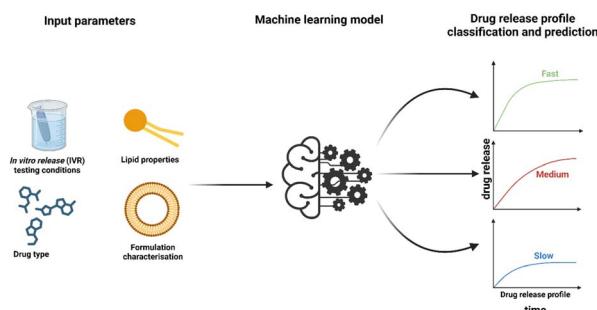
## PAPERS

2972

**Beyond training data: how elemental features enhance ML-based formation energy predictions**

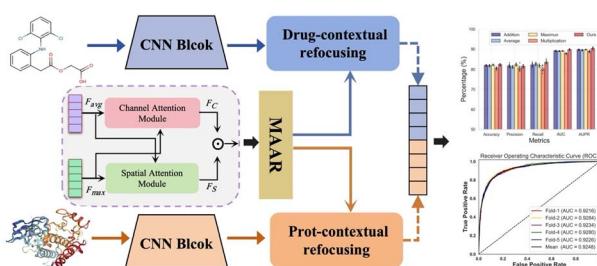
Hamed Mahdavi,\* Vasant Honavar and Dane Morgan

2983

**A machine learning workflow to accelerate the design of *in vitro* release tests from liposomes**

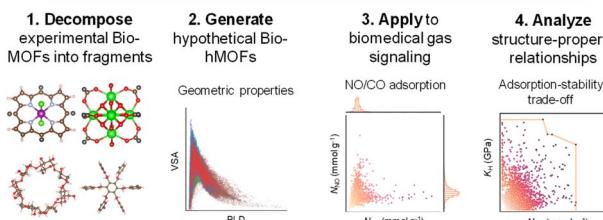
Daniel Yanes, Vasiliki Paraskevopoulou, Heather Mead, James Mann, Magnus Röding, Maryam Parhizkar, Cameron Alexander, Jamie Twycross\* and Mischa Zelzer\*

2994

**MAARDTI: a multi-perspective attention aggregation model for the prediction of drug–target interactions**

Xinke Zhan, Tiantao Liu, Changqing Yu, Yu-An Huang, Zhuhong You and Shirley W. I. Siu\*

3008

**Digital Discovery and Design of Bio-hMOFs****Digital design and discovery of biological metal-organic frameworks for gas signaling**

Yifei Yue, Athulya S. Palakkal, Saad Aldin Mohamed and Jianwen Jiang\*

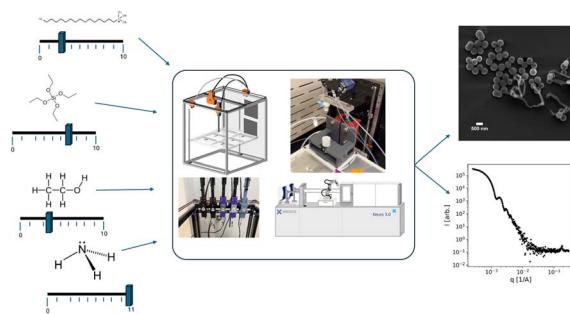


## PAPERS

3018

**Accelerated sol–gel synthesis of nanoporous silica via integrated small angle X-ray scattering with an open-source automation platform**

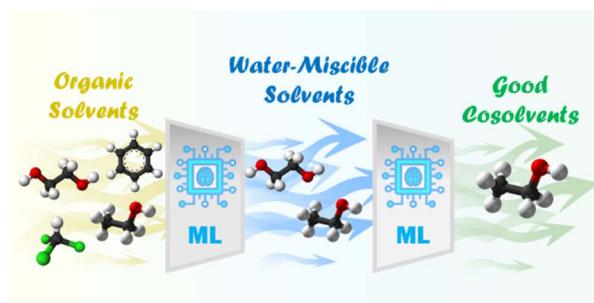
Brenden Pelkie, Chi Yuet Yung, Zachery R. Wylie and Lilo D. Pozzo\*



3031

**Predicting aqueous and organic solubilities with machine learning: a workflow for identifying organic cosolvents**

Maurycy Krzyżanowski, Sirazam Munira Aishee, Nirala Singh\* and Bryan R. Goldsmith\*



3043

**“Twisting” the data: a universal machine-learning approach to classify single-molecule curves and beyond**

C. Roldán-Piñero,\* M. Teresa González, Pablo M. Olmos, Linda A. Zotti and Edmund Leary\*

