

# 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 3(1) 1–222 (2024)



### Cover

See J. Pérez-Ríos *et al.*, pp. 34–50. Image reproduced by permission of Jesús Pérez-Ríos from *Digital Discovery*, 2024, 3, 34.



### Inside cover

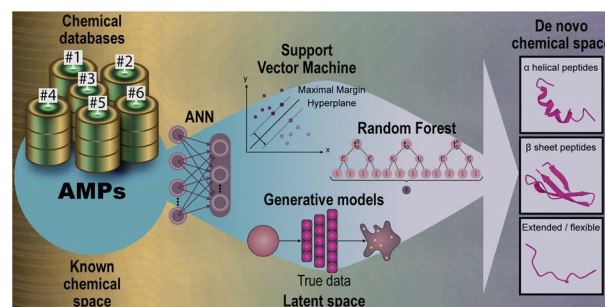
See Miroslava Nedyalkova, Marco Lattuada *et al.*, pp. 9–22. Image reproduced by permission of Miroslava Alexandrova Nedyalkova from *Digital Discovery*, 2024, 3, 9.

## PERSPECTIVES

9

### Progress and future of the computational design of antimicrobial peptides (AMPs): bio-inspired functional molecules

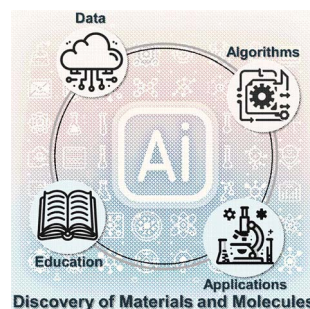
Miroslava Nedyalkova,\* Andrew S. Paluch, Diana Potes Vecini and Marco Lattuada\*



23

### Accelerated chemical science with AI

Seoin Back,\* Alán Aspuru-Guzik, Michele Ceriotti, Ganna Gryn'ova, Bartosz Grzybowski, Geun Ho Gu, Jason Hein, Kedar Hippalgaonkar, Rodrigo Hormázabal, Yousung Jung, Seonah Kim, Woo Youn Kim, Seyed Mohamad Moosavi, Juhwan Noh, Changyoung Park, Joshua Schrier, Philippe Schwaller, Koji Tsuda, Tejs Vegge, O. Anatole von Lilienfeld and Aron Walsh



# Royal Society of Chemistry approved training courses

Explore your options.  
Develop your skills.  
Discover learning  
that suits you.

**Courses in the classroom,  
the lab, or online**

Find something for every  
stage of your professional  
development. Search our  
database by:

- subject area
- location
- event type
- skill level

Members **get at least 10% off**

Visit [rsc.li/cpd-training](https://rsc.li/cpd-training)

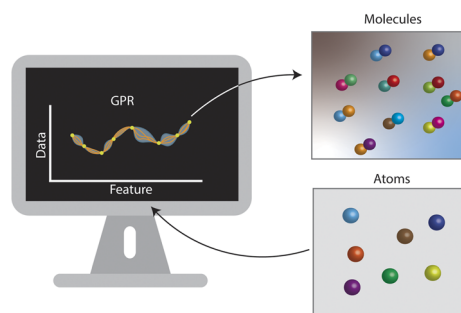
**SAVE  
10%**



34

## Spectroscopic constants from atomic properties: a machine learning approach

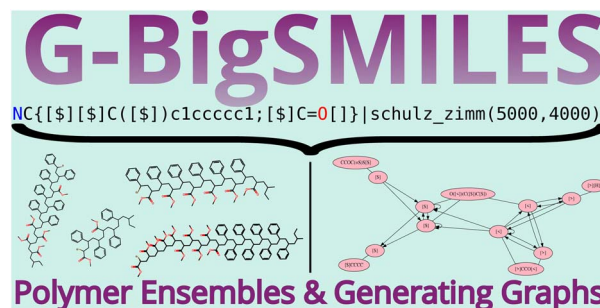
Mahmoud A. E. Ibrahim, X. Liu and J. Pérez-Ríos\*



51

## Generative BigSMILES: an extension for polymer informatics, computer simulations & ML/AI

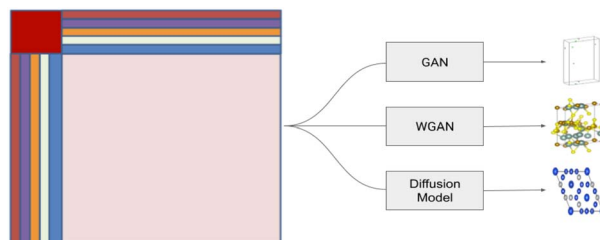
Ludwig Schneider,\* Dylan Walsh, Bradley Olsen and Juan de Pablo\*



62

## Generative adversarial networks and diffusion models in material discovery

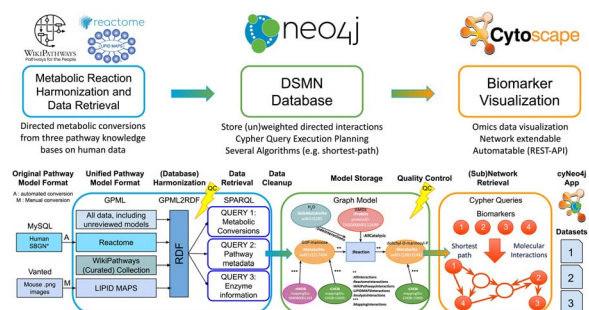
Michael Alverson,\* Sterling G. Baird, Ryan Murdock, (Enoch) Sin-Hang Ho, Jeremy Johnson and Taylor D. Sparks



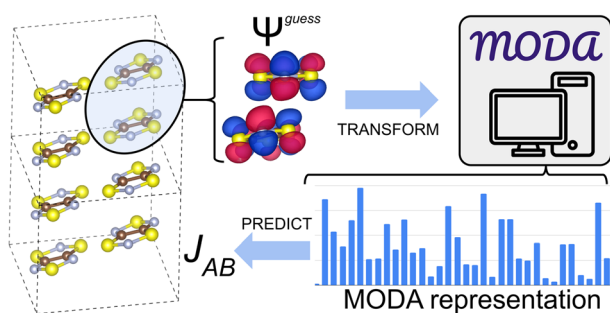
81

## Discovering life's directed metabolic (sub)paths to interpret human biochemical markers using the DSMN tool

Denise Slenter,\* Martina Kutmon, Chris T. Evelo and Egon L. Willighagen



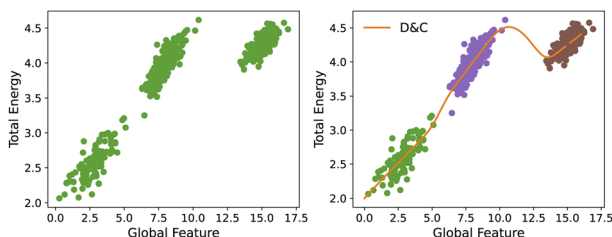
99



## Unlocking the predictive power of quantum-inspired representations for intermolecular properties in machine learning

Raul Santiago,\* Sergi Vela, Mercè Deumal and Jordi Ribas-Arino

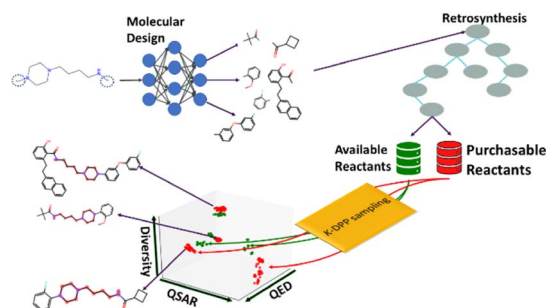
113



## Divide-and-conquer potentials enable scalable and accurate predictions of forces and energies in atomistic systems

Claudio Zeni,\* Andrea Anelli, Aldo Glielmo, Stefano de Gironcoli and Kevin Rossi\*

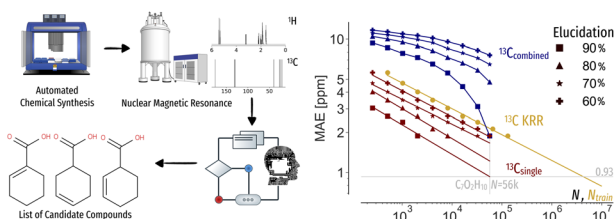
122



## De novo generated combinatorial library design

Simon Viet Johansson,\* Morteza Haghir Chehreghani, Ola Engkvist and Alexander Schliep

136



## Impact of noise on inverse design: the case of NMR spectra matching

Dominik Lemm, Guido Falk von Rudorff and O. Anatole von Lilienfeld



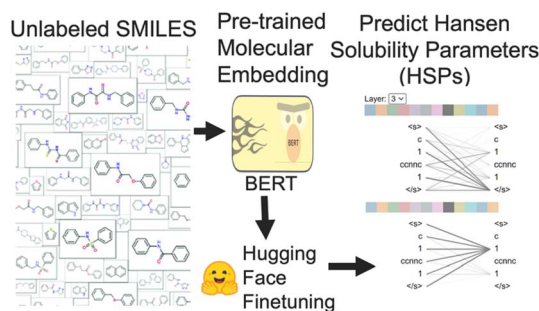


## PAPERS

145

## Using natural language processing (NLP)-inspired molecular embedding approach to predict Hansen solubility parameters

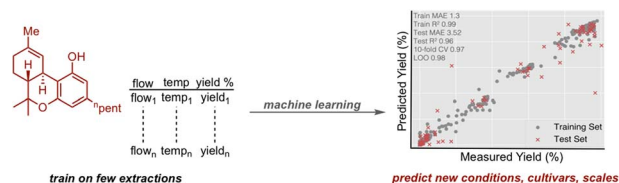
Jiayun Pang,\* Alexander W. R. Pine and Abdulai Sulemana



155

## Extraction yield prediction for the large-scale recovery of cannabinoids

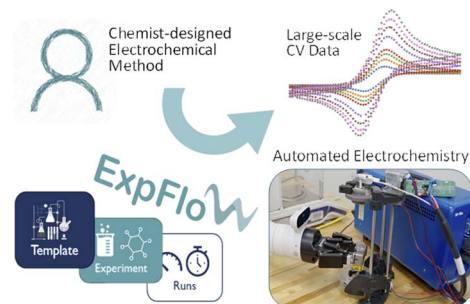
Hart Plommer, Isaiah O. Betinol, Tom Dupree, Markus Roggen\* and Jolene P. Reid\*



163

## ExpFlow: a graphical user interface for automated reproducible electrochemistry

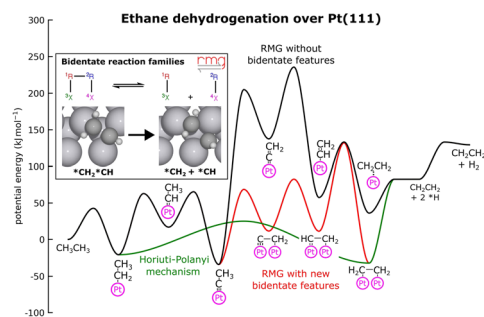
Rebekah Duke, Siamak Mahmoudi, Aman Preet Kaur, Vinayak Bhat, Ian C. Dingle, Nathan C. Stumme, Scott K. Shaw, David Eaton, Asmund Vego and Chad Risko\*



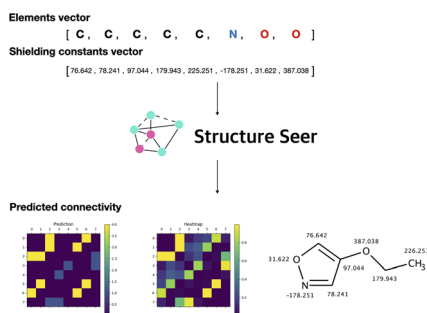
173

## Automatic mechanism generation involving kinetics of surface reactions with bidentate adsorbates

Bjarne Kreitz,\* Katrín Blöndal, Kirk Badger, Richard H. West and C. Franklin Goldsmith\*



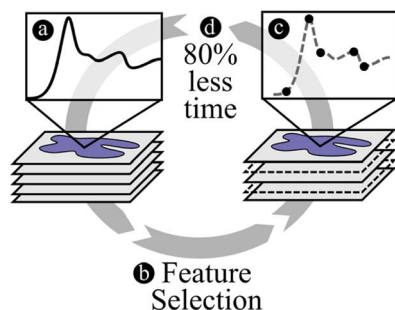
186



## Structure Seer – a machine learning model for chemical structure elucidation from node labelling of a molecular graph

Denis Andzheevich Sapegin\* and Joseph C. Bear

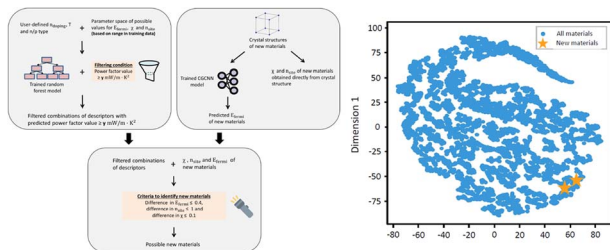
201



## Accelerating nano-XANES imaging *via* feature selection

Samantha Tetef, Ajith Pattammattel, Yong S. Chu, Maria K. Y. Chan\* and Gerald T. Seidler\*

210



## Machine learning based feature engineering for thermoelectric materials by design

U. S. Vaiteswar, Daniil Bash, Tan Huang, Jose Recatala-Gomez, Tianqi Deng, Shuo-Wang Yang, Xiaonan Wang\* and Kedar Hippalgaonkar\*

