Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

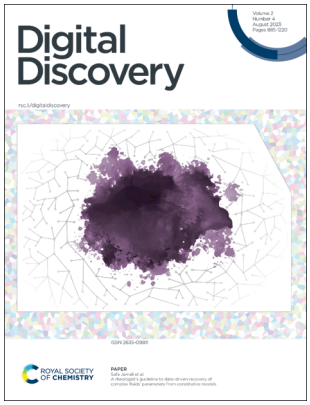
# Digital Discovery

## 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 2(4) 885–1220 (2023)



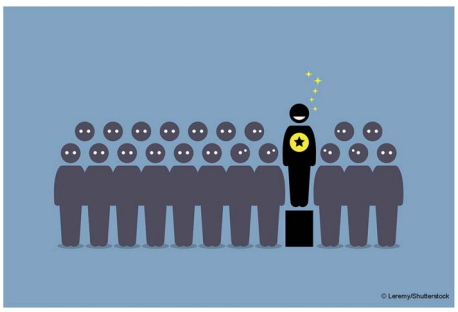

**Cover**
See Safa Jamali *et al.*, pp. 915–928. Image reproduced by permission of Safa Jamali and Milad Saadat from *Digital Discovery*, 2023, **2**, 915.

**Inside cover**
See María Victoria Gil, Berend Smit *et al.*, pp. 929–940. Image reproduced by permission of María Victoria Gil Matellanes from *Digital Discovery*, 2023, **2**, 929. This image was created by Kevin Jablonka using Midjourney.

## EDITORIAL

896

**Outstanding Reviewers for *Digital Discovery* in 2022**

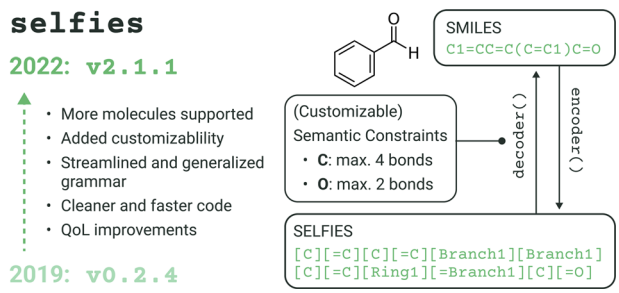


## TUTORIAL REVIEW

897

**Recent advances in the self-referencing embedded strings (SELFIES) library**

Alston Lo,* Robert Pollice, AkshatKumar Nigam, Andrew D. White, Mario Krenn and Alán Aspuru-Guzik



Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

## Editorial Staff

Editor
Anna Rulka

Deputy Editor
Audra Taylor

Editorial Production Manager
Viktoria Titmus

Assistant Editors
Angelica-Jane Kechinyere Onyekwere, Shwetha Krishna, Michael Whitelaw, Alexander Whiteside

Editorial Assistant
Samantha Campos

Publishing Assistant
Brittany Hanlon

Publisher
Neil Hammond

For queries about submitted articles please contact Viktoria Titmus, Editorial Production Manager in the first instance. E-mail digitaldiscovery@rsc.org

For pre-submission queries please contact Anna Rulka, Editor.
Email digitaldiscovery-rsc@rsc.org

Digital Discovery (electronic: ISSN 2635-098X) is published 6 times a year by the Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, UK CB4 0WF.

Digital Discovery is a Gold Open Access journal and all articles are free to read. Please email orders@rsc.org to register your interest or contact Royal Society of Chemistry Order Department, Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, CB4 0WF, UK Tel +44 (0)1223 432398; E-mail: orders@rsc.org

Whilst this material has been produced with all due care, the Royal Society of Chemistry cannot be held responsible or liable for its accuracy and completeness, nor for any consequences arising from any errors or the use of the information contained in this publication. The publication of advertisements does not constitute any endorsement by the Royal Society of Chemistry or Authors of any products advertised. The views and opinions advanced by contributors do not necessarily reflect those of the Royal Society of Chemistry which shall not be liable for any resulting loss or damage arising as a result of reliance upon this material. The Royal Society of Chemistry is a charity, registered in England and Wales, Number 207890, and a company incorporated in England by Royal Charter (Registered No. RC000524), registered office: Burlington House, Piccadilly, London W1J 0BA, UK, Telephone: +44 (0) 207 4378 6556.

Advertisement sales:
Tel +44 (0) 1223 432246; Fax +44 (0) 1223 426017;
E-mail advertising@rsc.org

For marketing opportunities relating to this journal, contact marketing@rsc.org

# Digital Discovery

rsc.li/digitaldiscovery

Digital Discovery is a gold open access journal publishing top research at the intersection of chemistry, materials science and biotechnology. Blurring the barriers between computation and experimentation, we focus on the integration of digital and automation tools with science, putting data first to ensure reproducibility and faster progress.

## Editorial Board

Editor in Chief
Alán Aspuru-Guzik, University of Toronto, Canada

Associate Editors
Jason E. Hein, University of British Columbia, Canada
Linda Hung, Toyota Research Institute, USA
Joshua Schrier, Fordham University, USA
Kedar Hippalgaonkar, Nanyang Technological University, Singapore
Cesar de la Fuente, University of Pennsylvania, USA

Members
Yousung Jung, KAIST, South Korea
Anat Milo, Ben-Gurion University of the Negev, Israel
Lilo D. Pozzo, University of Washington, USA
Ekaterina Skorb, ITMO University, Russia

## Advisory Board

Abigail Doyle, University of California Los Angeles, USA
Ola Engkvist, AstraZeneca and Chalmers University of Technology, Sweden
Pablo Carbonell, University of Valencia, Spain
Ian Foster, University of Chicago, USA
Cecilia Clementi, Freie Universität Berlin, Germany
Heather Kulik, MIT, USA
Silvana Botti, Friedrich Schiller University Jena, Germany
Marwin Segler, Microsoft, Germany
Jan Jensen, University of Copenhagen, Denmark
Berend Smit, EPFL, Switzerland
Conor Coley, MIT, USA
Koji Tsuda, The University of Tokyo, Japan
Isao Tanaka, Kyoto University, Japan
Shuye Ping Ong, University of California San Diego, USA
Alexandre Tkatchenko, University of Luxembourg, Luxembourg
Juan Alegre, Colorado State University, USA

## Information for Authors

Full details on how to submit material for publication in Digital Discovery are given in the Instructions for Authors (available from http://www.rsc.org/authors). Submissions should be made via the journal's homepage: rsc.li/digitaldiscovery

Authors may reproduce/republish portions of their published contribution without seeking permission from the Royal Society of Chemistry, provided that any such republication is accompanied by an acknowledgement in the form: (Original Citation)–Reproduced by permission of the Royal Society of Chemistry.

This journal is © The Royal Society of Chemistry 2023.
Apart from fair dealing for the purposes of research or private study for non-commercial purposes, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988 and the Copyright and Related Rights Regulation 2003, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the Publishers or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency in the UK. US copyright law is applicable to users in the USA.

Registered charity number: 207890

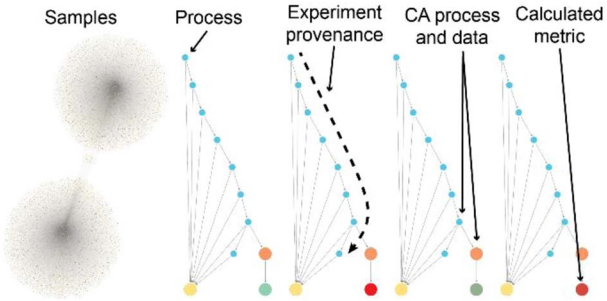
ROYAL SOCIETY OF CHEMISTRY

## COMMUNICATION

909

**The materials experiment knowledge graph**

Michael J. Statt,* Brian A. Rohr,* Dan Guevarra, Ja'Nya Breeden, Santosh K. Suram and John M. Gregoire*

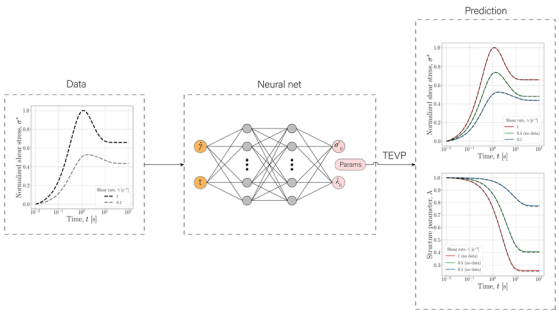



## PAPERS

915

**A rheologist's guideline to data-driven recovery of complex fluids' parameters from constitutive models**

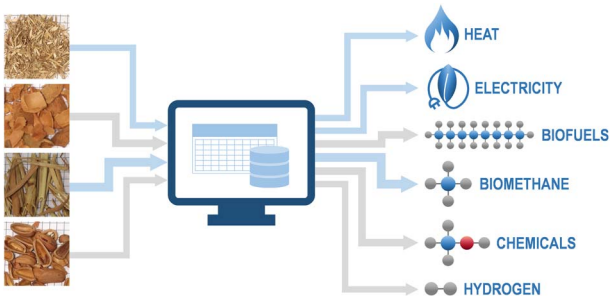
Milad Saadat, Deepak Mangal and Safa Jamali*




929

**Biomass to energy: a machine learning model for optimum gasification pathways**

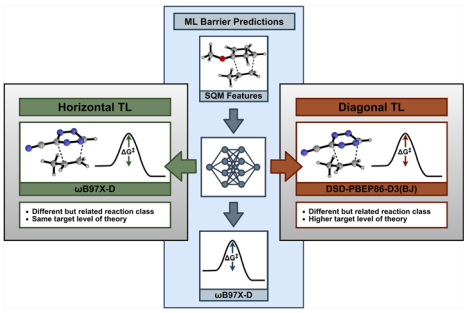
María Victoria Gil,* Kevin Maik Jablonka, Susana Garcia, Covadonga Pevida and Berend Smit*




941

**Machine learning reaction barriers in low data regimes: a horizontal and diagonal transfer learning approach**

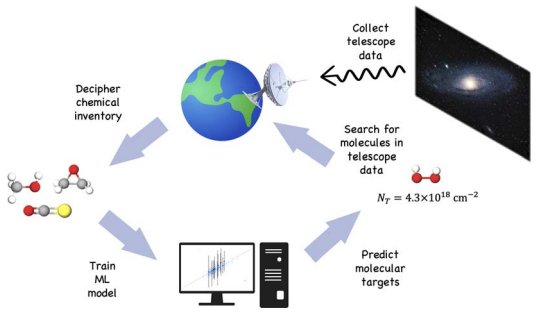
Samuel G. Espley, Elliot H. E. Farrar, David Buttar, Simone Tomasi and Matthew N. Grayson*




Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

PAPERS

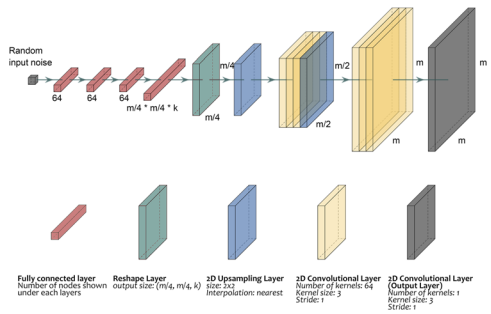
952




## Implementation of rare isotopologues into machine learning of the chemical inventory of the solar-type protostellar source IRAS 16293-2422

Zachary T. P. Fried,* Kin Long Kelvin Lee, Alex N. Byrne and Brett A. McGuire*

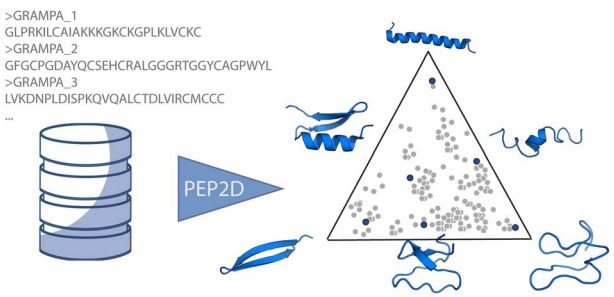
967




## A scalable neural network architecture for self-supervised tomographic image reconstruction

Hongyang Dong, Simon D. M. Jacques, Winfried Kockelmann, Stephen W. T. Price, Robert Emberson, Dorota Matras, Yaroslav Odarchenko, Vesna Middelkoop, Athanasios Giokaris, Olof Gutowski, Ann-Christin Dippel, Martin von Zimmermann, Andrew M. Beale, Keith T. Butler* and Antonis Vamvakeros*

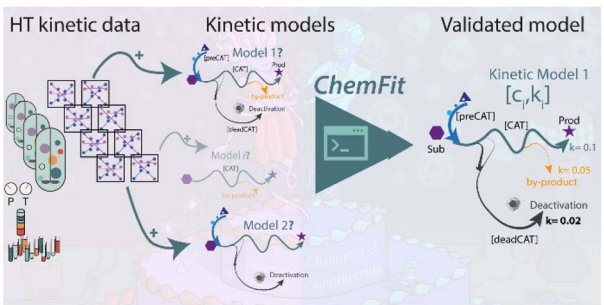
981




## Benchmarking protein structure predictors to assist machine learning-guided peptide discovery

Victor Daniel Aldas-Bulos and Fabien Plisson*

994




## Model-based evaluation and data requirements for parallel kinetic experimentation and data-driven reaction identification and optimization

Nathan Jiscoot, Evgeny A. Uslamin* and Evgeny A. Pidko*

Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.



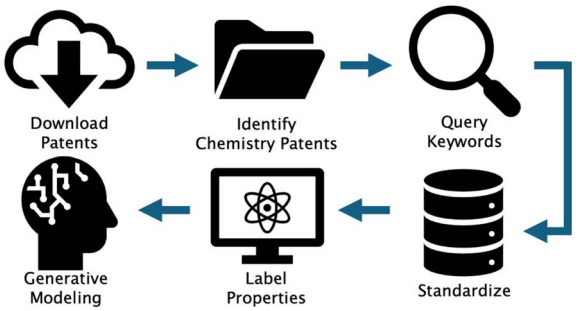
This journal is © The Royal Society of Chemistry 2023

PAPERS

1006

**Automated patent extraction powers generative modeling in focused chemical spaces**

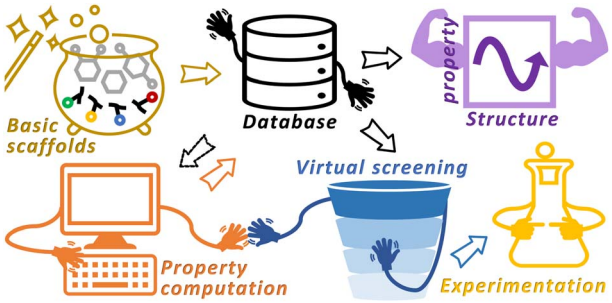
Akshay Subramanian, Kevin P. Greenman, Alexis Gervaix, Tzuhsiung Yang and Rafael Gómez-Bombarelli*




1016

**Discovery of lead quinone cathode materials for Li-ion batteries**

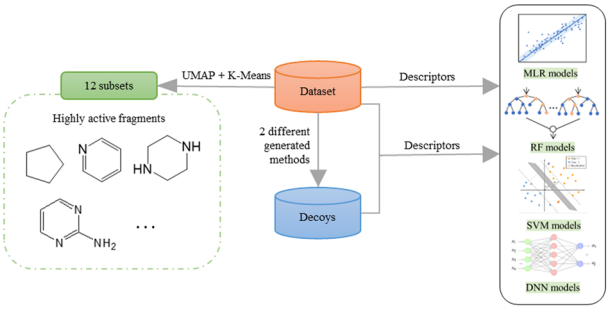
Xuan Zhou, Abhishek Khetan, Jie Zheng, Mark Huijben, René A. J. Janssen and Süleyman Er*




1026

**A SAR and QSAR study on cyclin dependent kinase 4 inhibitors using machine learning methods**

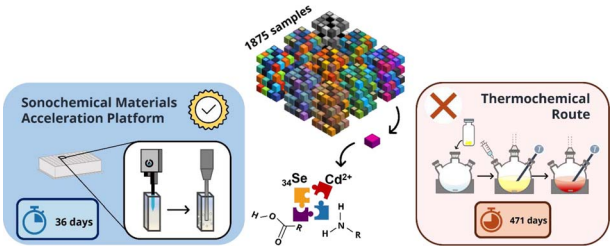
Xiaoyang Pang, Yunyang Zhao, Guo Li, Jianrong Liu* and Aixia Yan*




1042

**A high-throughput workflow for the synthesis of CdSe nanocrystals using a sonochemical materials acceleration platform**

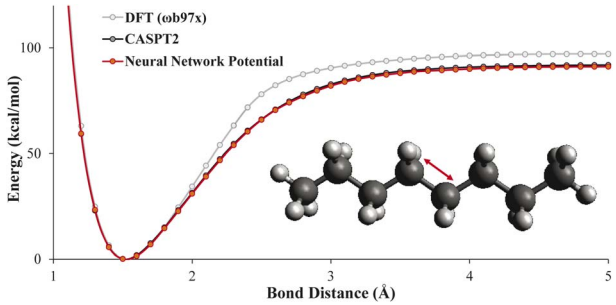
Maria Politi, Fabio Baum, Kiran Vaddi, Edwin Antonio, Joshua Vasquez, Brittany P. Bishop, Nadya Peek, Vincent C. Holmberg and Lilo D. Pozzo*




Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

PAPERS

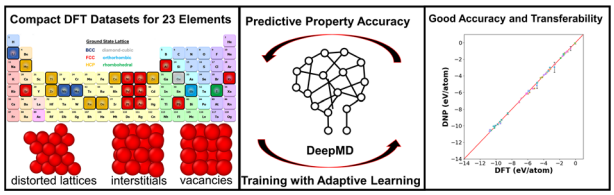
1058




**Neural network potentials for reactive chemistry: CASPT2 quality potential energy surfaces for bond breaking**

Quin H. Hu, Andrew M. Johannesen, Daniel S. Graham and Jason D. Goodpaster*

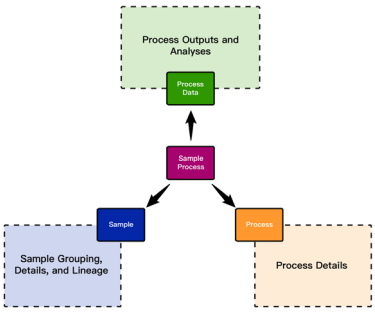
1070




**Highly transferable atomistic machine-learning potentials from curated and compact datasets across the periodic table**

Christopher M. Andolina and Wissam A. Saidi*

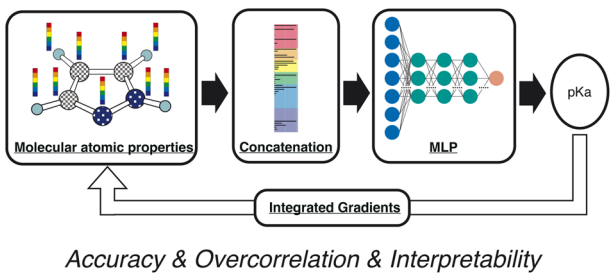
1078




**ESAMP: event-sourced architecture for materials provenance management and application to accelerated materials discovery**

Michael J. Statt,* Brian A. Rohr, Kris Brown, Dan Guevarra, Jens Hummelshøj, Linda Hung, Abraham Anapolsky, John M. Gregoire* and Santosh K. Suram*

1089




**Feature selection in molecular graph neural networks based on quantum chemical approaches**

Daisuke Yokogawa* and Kayo Suda

Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.



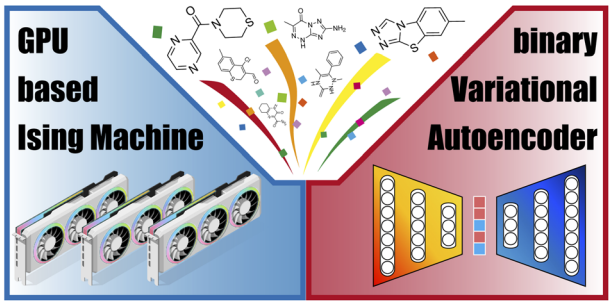
This journal is © The Royal Society of Chemistry 2023

PAPERS

1098

**Chemical design with GPU-based Ising machines**

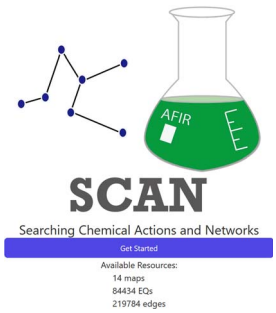
Zetian Mao, Yoshiki Matsuda, Ryo Tamura* and Koji Tsuda*




1104

**Searching chemical action and network (SCAN): an interactive chemical reaction path network platform**

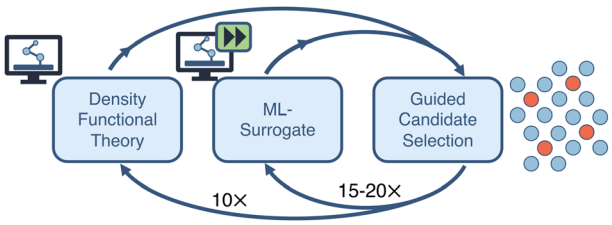
Mikael Kuwahara, Yu Harabuchi, Satoshi Maeda,* Jun Fujima* and Keisuke Takahashi*




1112

**By how much can closed-loop frameworks accelerate computational materials discovery?**

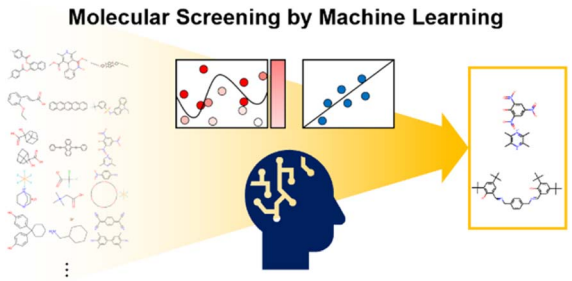
Lance Kavalsky, Vinay I. Hegde, Eric Muckley, Matthew S. Johnson, Bryce Meredig* and Venkatasubramanian Viswanathan*




1126

**Molecular screening for solid–solid phase transitions by machine learning**

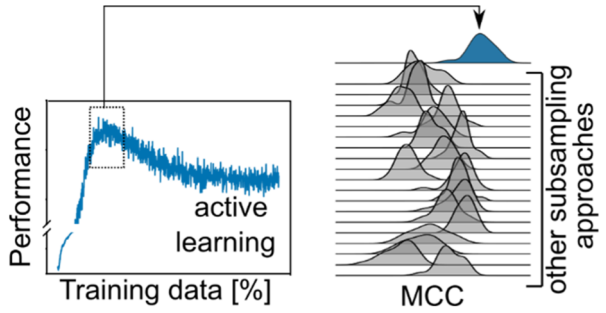
Daisuke Takagi, Kazuki Ishizaki, Toru Asahi and Takuya Taniguchi*




Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

PAPERS

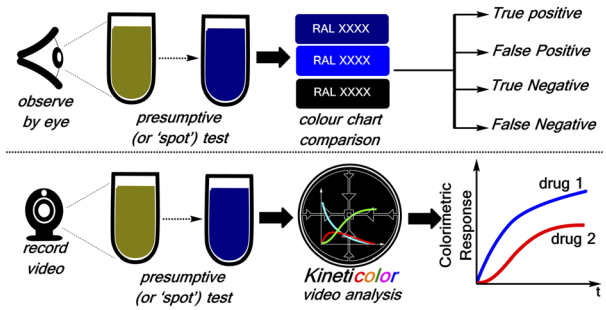
1134




**Improving molecular machine learning through adaptive subsampling with active learning**

Yujing Wen, Zhixiong Li, Yan Xiang and Daniel Reker*

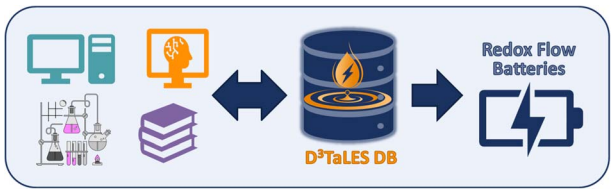
1143




**Teaching old presumptive tests new digital tricks with computer vision for forensic applications**

Nathalie Bugeja, Cameron Oliver, Nicole McGrath, Jake McGuire, Chunhui Yan, Felicity Carlysle-Davies and Marc Reid*

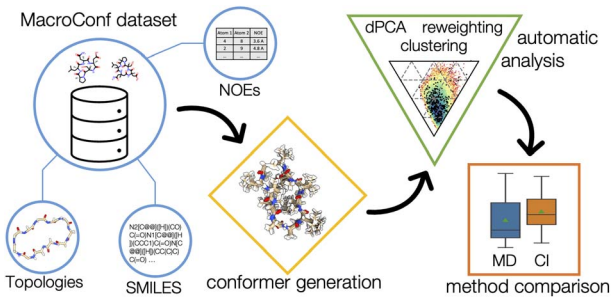
1152




**Towards a comprehensive data infrastructure for redox-active organic molecules targeting non-aqueous redox flow batteries**

Rebekah Duke, Vinayak Bhat, Parker Sornberger, Susan A. Odom and Chad Risko*

1163




**MacroConf – dataset & workflows to assess cyclic peptide solution structures**

Daniel Crusius, Jason R. Schnell, Flaviu Cipcigan and Philip C. Biggin*

Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.



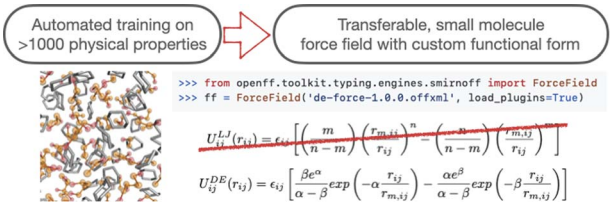
This journal is © The Royal Society of Chemistry 2023

PAPERS

1178

## A transferable double exponential potential for condensed phase simulations of small molecules

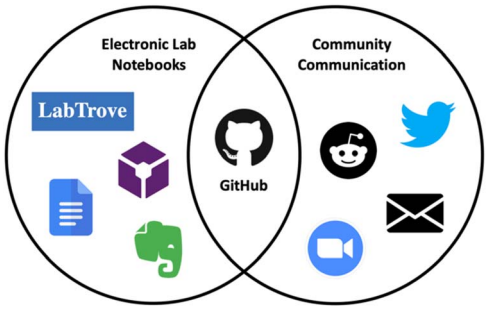
Joshua T. Horton, Simon Boothroyd, Pavan Kumar Behara, David L. Mobley and Daniel J. Cole*




1188

## GitHub as an open electronic laboratory notebook for real-time sharing of knowledge and collaboration

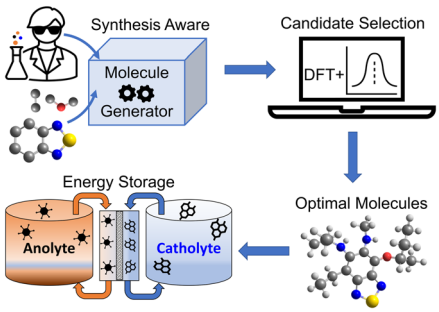
Kymberley R. Scroggie, Klementine J. Burrell-Sander, Peter J. Rutledge and Alice Motion*




1197

## *In silico* discovery of a new class of anolyte redoxmers for non-aqueous redox flow batteries

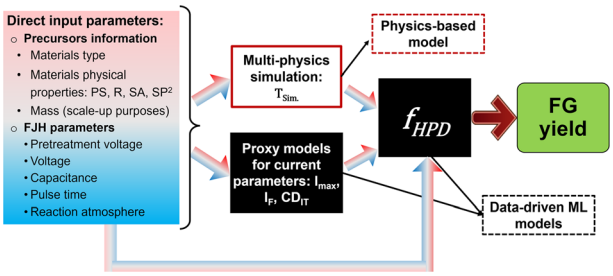
Akash Jain, Ilya A. Shkrob, Hieu A. Doan, Lily A. Robertson, Lu Zhang and Rajeev S. Assary*




1209

## A scientific machine learning framework to understand flash graphene synthesis

Kianoosh Sattari, Lucas Eddy, Jacob L. Beckham, Kevin M. Wyss, Richard Byfield, Long Qian, James M. Tour* and Jian Lin*




Open Access Article. Published on 08 agosto 2023. Downloaded on 18/07/2025 03:49:56.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.