Use of Materials Informatics to Discover New Materials: Theory and Successes

David A. Winkler
La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Australia
Monash Institute of Pharmaceutical Science, Monash University, Parkville, Australia
School of Pharmacy, University of Nottingham, Nottingham, UK
Email: d.winkler@latrobe.edu.au

ID: PLN01
Plenary Talk
Aug. 23, 9.00 - 9.40

Keywords: Materials design, computational chemistry, machine learning, artificial intelligence, molecular modelling

Currently, we are experiencing amazing, paradigm shifting scientific developments. We understand that the size of small molecule and materials spaces is essentially, infinite, representing an inexhaustible supply of potential drugs and materials with useful properties. This has seen a rapid increase in automation and robotics, allowing synthesis of new molecules and materials and measurement of properties orders of magnitude faster. This has created massive databases of complex genetic, structural, chemical, property, and biological information.

Finding "islands of chemical utility" in a vast palette of possibilities and extracting meaning from massive databases has driven a rise in applications of AI and machine learning, and development of methods. There has been a parallel rise in applications to most aspects of modern life – medicine, finance, manufacturing, social media. Recently, we saw development of rapid, accurate quantum machine learning methods, generative methods to suggest new molecules or materials with improved properties, prediction of protein structures from sequence (AlphaFold), the beginning of general AI (ChatGPT), massive "make on demand" chemical libraries (ZINC-22), seminal work on autonomous chemical discovery, and increasing use of other AI methods (evolutionary algorithms) to discover molecules and materials with improved properties.

This presentation will discuss drivers for these developments, summarize contributions my teams have made to adapting and applying machine learning, and provide examples of applications to biomaterials and regenerative medicine, drug design, 2D and porous materials, nanomaterials, surface science, cancer diagnostics, corrosion control, and sustainable energy sources.

References

1. Li *et al.*, "Rational Atom Substitution to Obtain Efficient, Lead-Free Photocatalytic Perovskites Assisted by Machine Learning and DFT Calculations", Angew. Chem. Int. Ed. **62**, e202315002 (2023).
2. Surmiak *et al.*, "Machine Learning-Enhanced High-Throughput Fabrication and Optimization of Quasi-2D Ruddlesden-Popper Perovskite Solar Cells", Adv. Energy Mater. **13**, 2203859 (2023).
3. Mai *et al.*, "Machine Learning in the Development of Adsorbents for Clean Energy Applications", Adv. Sci. **9**, 2203899 (2022).
4. Mai *et al.*, "Machine Learning for Electrocatalyst and Photocatalyst Design and Discovery", Chem. Rev. **122**, 13478 (2022).
5. Muratov *et al.*, "A critical overview of computational approaches employed for COVID-19 drug discovery", Chem. Soc. Rev. **50**, 9121 (2021).
6. Vasilevich *et al.*, "Evolutionary design of optimal surface topographies for biomaterial", Sci. Rep. **10**, 22160 (2020).
7. Meftahi *et al.*, "Machine Learning Property Prediction for Organic Photovoltaic Devices", npj Comput. Mater. **6**, 166 (2020).
8. Muratov, *et al.*, "QSAR without Borders", Chem. Soc. Rev. **49**, 3525 (2020).