# ****De Novo Drug Design: Combining Pharmacophore Modeling, 3D Shape Matching, and Generative AI****

***Jozef Fülöp¹, Martin Šícho¹, Daniel Svozil***

Department of Informatics and Chemistry & CZ-OPENSCREEN: National Infrastructure for Chemical Biology, Faculty of Chemical Technology, University of Chemistry and Technology, Technická 5, 16628, Prague, Czech Republic

2 Leiden Academic Centre for Drug Research, Leiden University, 55 Einsteinweg, 2333 CC, Leiden, The Netherlands

Recent work shows that shape‑based generative reinforcement‑learning models enable scaffold hopping and diverse lead discovery [1], while voxel‑based shape autoencoders can uncover novel scaffolds from 3‑D inputs [2]. Pharmacophore modeling that exploits shape‑and‑colour Tanimoto scores (e.g., ROCS) reliably detects compounds sharing key 3‑D features beyond close analogues [3]. Building on this, we introduce an open‑source pipeline that (i) aligns ligands by shape and pharmacophoric “colour,” (ii) filters for high Tanimoto shape/colour similarity and synthetic accessibility, and (iii) uses DrugEx to propose new chemotypes meeting both criteria [4]. Using the chemokine receptor CCR2, an allosteric target involved in inflammation and metastasis, as a case study [5], we benchmark ROCS/OMEGA against RDKit/CDPKit [6, 7] and find speed‑versus‑optimisation trade‑offs on large libraries, yet comparable overlay accuracy. This integrated workflow therefore provides a practical open‑source alternative for next‑generation lead discovery.

References  
[1] Papadopoulos et al., Bioorg. Med. Chem., 2021 [5] Zheng et al., Nature, 2016   
[2] Skalic et al., J. Chem. Inf. Model., 2019 [6] RDKit: [www.rdkit.org](http://www.rdkit.org/)  
[3] Kearnes & Pande, J. Comput. Aided Mol. Des., 2016 [7] CDPKit: github.com/molinfovienna/CDPKit   
[4] Šícho et al., J. Chem. Inf. Model., 2023