From *In Silico* Design to Automated Synthesis: An AI-Driven Framework for Late-Stage Functionalization

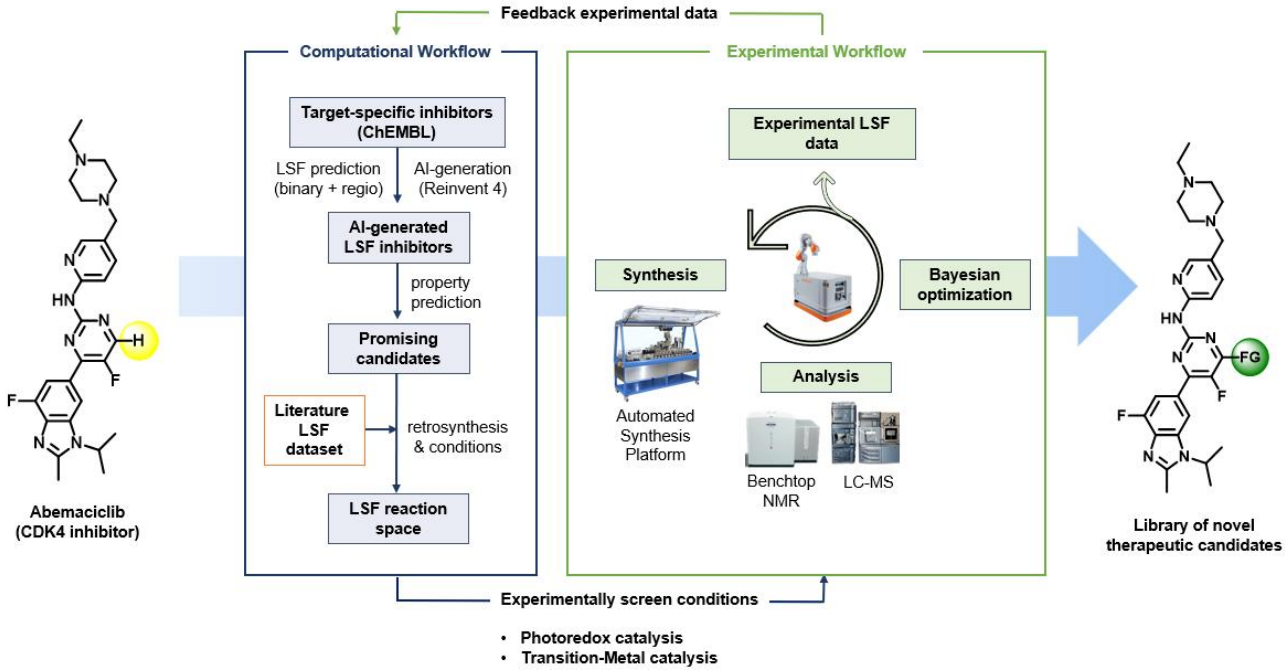
Lyubomir Kotopanov[1] Anthony Bradley[1] Jean-Francois Ayme[1] Kenneth Clark[2] John W. Ward[1] Andrew I. Cooper[1]

[1][University of Liverpool, Department of Chemistry and Materials Innovation Factory, University of Liverpool, 51 Oxford Street, Liverpool, L7 3NY, UK.] [2][AstraZeneca, Macclesfield Campus, Mill Court, Charter Way, Silk Road Business Park, Macclesfield, Cheshire, SK10 2NA, UK.]

Correspondence to: [Andrew I. Cooper] aicooper@liverpool.ac.uk

1. Introduction

Self-driving laboratories guided by machine learning algorithms hold the potential to significantly accelerate discoveries in the chemical sciences. However, to fully exploit the opportunities offered by autonomous laboratories, medicinal chemistry faces a major bottleneck – not only in the physical preparation of molecules, but also in computer-based retrosynthetic planning. Current algorithms favour well-established chemistries, whereas newer, potentially more efficient reactions are under-prioritised because they are less explored in the literature. We present a flexible data-driven framework to guide high-throughput late-stage functionalization of drug compounds. This pipeline generates diverse therapeutic candidates *in silico* with promising properties and synthesizability via user-defined late-stage functionalization reactions with predicted conditions. Our results suggest that data-driven methods can be utilized to establish a robust and standardized workflow for the autonomous generation of promising drug-like molecules that can be efficiently accessed via innovative reaction pathways.

Acknowledgments

The authors acknowledge funding from both the Leverhulme Trust via the Leverhulme Research Centre for Functional Materials Design and AstraZeneca UK.