DSMC.2: A Monte Carlo Tree-Based Search Model for Multi-Objective Drug Design

Yali Lyu[1], Filippo Prischi[2*], Tristan Cazenave[3*], Olivier E. Pardo[4*]

[1] *Faculty of Medicine, Imperial College, Du Cane Road, London W12 0NN, United Kingdom*
[2] *Randall Centre for Cell and Molecular Biophysics, School of Basic and Medical Biosciences, King's College London, London SE1 1UL, United Kingdom* filippo.prischi@kcl.ac.uk
[3] *LAMSADE, Université Paris-Dauphine, Pl. du Maréchal de Lattre de Tassigny, 75016 Paris, France* tristan.cazenave@lamsade.dauphine.fr
[4] *Division of Cancer, Department of Surgery and Cancer, Imperial College, Du Cane Road, London W12 0NN, United Kingdom* o.pardo@imperial.ac.uk

Despite recent advances in Machine Learning (ML) and Artificial Intelligence (AI), the drug discovery process remains an expensive and inefficient bottleneck within the design-make-test-analyze (DMTA) cycle. Moreover, most de novo drug design models are single-objective, focusing on optimizing only one desired property.

We introduce DSMC.2, a multi-objective search model based on Nested Monte Carlo Search[1] (NMCS) for *de novo* drug design and optimization. NMCS is a recursive extension of Monte Carlo Tree Search (MCTS) that enables the use of different scoring functions at various levels of the search hierarchy. Through case studies targeting the androgen receptor (AR) and human epidermal growth factor receptor 2 (HER2), we demonstrate that DSMC.2 effectively generates molecules that are both drug-like and protein-specific. For top-scoring molecules, interaction analysis successfully identifies extensive matching interactions with existing drug molecules, demonstrating DSMC.2's utility in readily applicable scenarios.

Additionally, DSMC.2 leverages the previously developed DSMC[2] module to ensure 100% chemical validity and employs Natural Language Processing (NLP) techniques such as n-grams to achieve higher internal chemical diversity compared to existing search models. Furthermore, we develop a DSMC.Opt module that utilizes substructures from previously generated DSMC.2 molecules to create new compounds with desirable drug-likeness and binding specificity scores, highlighting further enhancement potential. Taken together, DSMC.2 is a robust and versatile search model for de novo drug design and optimization.

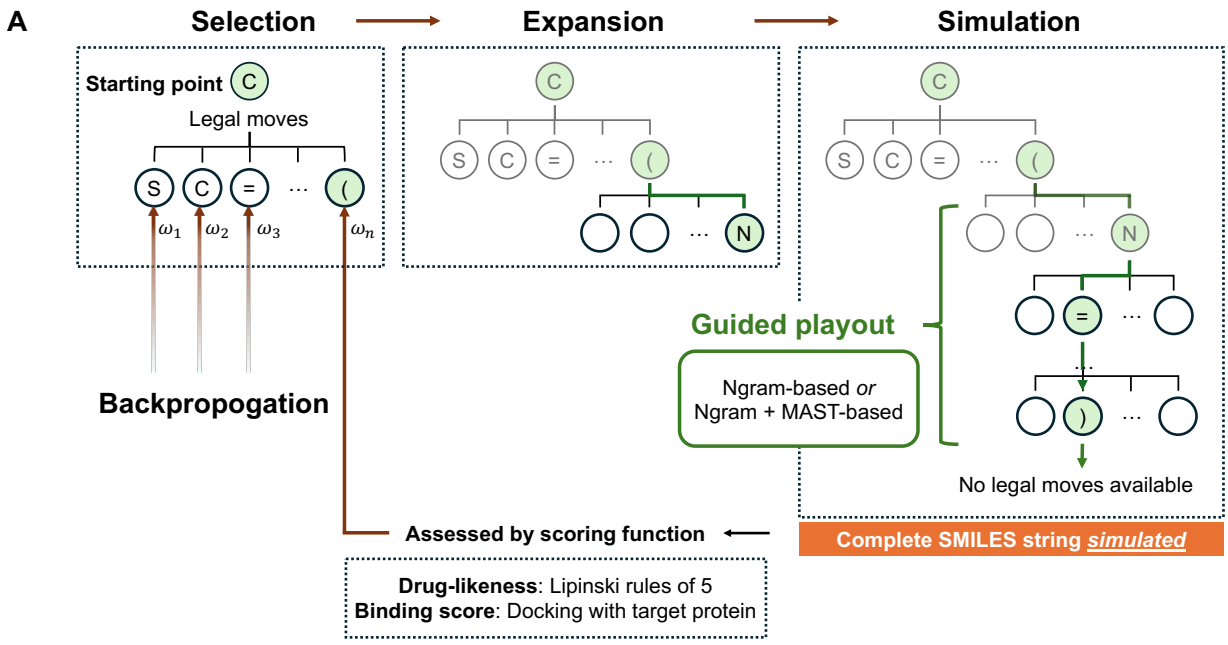

Figure 1 Workflow of NMCS-enabled DSMC.2

[1] Cazenave, T. Nested Monte-Carlo Search. in *IJCAI* 456–461 (Pasadena, 2009).
[2] Roucairol, M., Georgiou, A., Cazenave, T., Prischi, F. & Pardo, O. E. DrugSynthMC: An Atom-Based Generation of Drug-like Molecules with Monte Carlo Search. *J Chem Inf Model* **64**, 7097–7107 (2024).