

Encoding Stereochemistry in Molecular Fingerprints

M. Orsi¹, D. Probst², J. Reymond^{3*}

¹University of Bern, Department of Chemistry, Biochemistry and Pharmacy, ²École Polytechnique Federale de Lausanne, ³University of Bern

Molecular fingerprints are important tools enabling fast similarity comparisons between molecules¹. These fingerprints encode different molecular properties, such as pharmacophores², atom distances^{3,4} and substructures⁴⁻⁶, and are particularly useful in the context of virtual screening, where one carries out millions of comparisons to detect potential hits. However, a serious limitation of the most well-known fingerprints is their inability to account for chirality in a chemically meaningful manner. Specifically, the challenge lies in encoding chirality to distinguish enantiomers and diastereomers as highly similar yet distinct entities. In this study, we introduce MHFP6* and MAP4*, enhanced versions of our MHFP6⁶ and MAP4⁴ fingerprints, which correctly encode CIP chirality in a chemically meaningful manner. Both fingerprints demonstrate the capability to correctly differentiate isomers, thereby overcoming the previously mentioned limitations.



[1] Peter Willet, *Drug Discov. Today* **2006**, 11 (23-24), 1046-1053.

[2] Alice Capecchi, Mahendra Awale, Daniel Probst, Jean-Louis Reymond, *Mol. Inform.* **2019**, 38 (5), 1900016.

[3] Raymond E. Carhart, Dennis H. Smith, Rengachari Venkataraghavan, *J. Chem. Inf. Comput. Sci.* **1985**, 25 (2), 64-73.

[4] Alice Capecchi, Daniel Probst, Jean-Louis Reymond, *J. Cheminformatics* **2020**, 12 (1), 43.

[5] David Rogers, Mathew Hahn, *J. Chem. Inf. Model.* **2010**, 50 (5), 742-754.

[6] Daniel Probst, Jean-Louis Reymond, *J. Cheminformatics* **2018**, 10 (1), 66.