Received: 7 February 2024 | Revised: 3 April 2024 | Accepted: 4 April 2024

DOI: 10.1002/minf.202400050

RESEARCH ARTICLE

molecular informatics

Navigating pharmacophore space to identify activity discontinuities: a case study with BCR-ABL

Maroua Lejmi[1, 2] | **Damien Geslin**[3] | **Ronan Bureau**[3] | **Bertrand Cuissart**[1] | **Ilef Ben Slima**[4, 5] | **Nida Meddouri**[6] | **Amel Borgi**[2] | **Jean-Luc Lamotte**[1] | **Alban Lepailleur**[3]

[1]Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen, UNICAEN, ENSICAEN, CNRS - UMR GREYC, Normandie Univ, Caen, France

[2]Laboratoire en Informatique, Programmation Algorithmique et Heuristique, LIPAH, Université de Tunis El Manar, Tunis, Tunisia

[3]Centre d'Etudes et de Recherche sur le Médicament de Normandie, UNICAEN, CERMN, Normandie Univ, Caen, France

[4]ISMAI, University of Kairouan, Kairouan, Tunisia

[5]Laboratory of Signals, systeMs, aRtificial Intelligence and neTworkS, SM@RTS, Digital Research Center of Sfax, Sfax, Tunisia

[6]Laboratoire de Recherche de l'EPITA, LRE, Le Kremlin-Bicêtre, Paris, France

Correspondence
Alban Lepailleur, Centre d'Etudes et de Recherche sur le Médicament de Normandie, UNICAEN, CERMN, Normandie Univ, Caen, France
+33231566822 +33231566803.
Email: alban.lepailleur@unicaen.fr

Abstract

The exploration of chemical space is a fundamental aspect of chemoinformatics, particularly when one explores a large compound data set to relate chemical structures with molecular properties. In this study, we extend our previous work on chemical space visualization at the pharmacophoric level. Instead of using conventional binary classification of affinity (active vs inactive), we introduce a refined approach that categorizes compounds into four distinct classes based on their activity levels: super active, very active, active, and inactive. This classification enriches the color scheme applied to pharmacophore space, where the color representation of a pharmacophore hypothesis is driven by the associated compounds. Using the BCR-ABL tyrosine kinase as a case study, we identified intriguing regions corresponding to pharmacophore activity discontinuities, providing valuable insights for structure-activity relationships analysis.

KEYWORDS
activity discontinuity, chemical space, pharmacophore, structure-activity relationships

1 | INTRODUCTION

The concept of chemical space is central to many of the tasks in chemoinformatics where large sets of compounds must be explored and chemical structures must be related to molecular properties and molecular activities [1]. Molecular space is frequently represented as a coordinate-based model providing an appealing representation but it suffers from a number of limitations, such as its high dimensionality. The visualization of the high-dimensional chemical space clearly requires the application of dimensionality reduction techniques. The most commonly utilized methods include principal component analysis (PCA), self-organizing maps (SOM), stochastic proximity embedding (SPE), t-distributed stochastic neighbor embedding (t-SNE), and

Maroua Lejmi and Damien Geslin contributed equally to the work.

