



**HAL**  
open science

## Towards lattice-based interactive pharmacophore exploration

Etienne Lehembre, Jean-Luc Lamotte, Abdelkader Ouali, Christophe Couronne, Bruno Crémilleux, Alban Lepailleur, Jean-Luc Manguin, Ronan Bureau, Albrecht Zimmermann, Bertrand Cuissart

► **To cite this version:**

Etienne Lehembre, Jean-Luc Lamotte, Abdelkader Ouali, Christophe Couronne, Bruno Crémilleux, et al.. Towards lattice-based interactive pharmacophore exploration. 10th SFCi, Sep 2021, Lille, France. hal-03764647

**HAL Id: hal-03764647**

**<https://hal.science/hal-03764647>**

Submitted on 30 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# TOWARDS LATTICE-BASED INTERACTIVE PHARMACOPHORE EXPLORATION

Etienne Lehembre<sup>a</sup>, Jean-Luc Lamotte<sup>b</sup>, Abdelkader Ouali<sup>a</sup>, Christophe Couronne<sup>a</sup>, Bruno Crémilleux<sup>a</sup>, Alban Lepailleur<sup>b</sup>, Jean-Luc Manguin<sup>a</sup>, Ronan Bureau<sup>b</sup>, Albrecht Zimmermann<sup>a</sup>, Bertrand Cuissart<sup>a</sup>

<sup>a</sup> : GREYC, CNRS UMR 6072, UNICAEN, Normandy Univ. Caen, France

<sup>b</sup> : CERMN, EA 4258 FR CNRS 3038 INC3M SF 4206 ICORE, UNICAEN, Normandy Univ. Caen, France

corresponding author : etienne.lehembre@unicaen.fr

Computing key structural configurations from ligands is a major challenges in drug design. In the **InvolvD project**<sup>1</sup>, we designed a **pharmacophore mining method** to discover structural layouts potentially linked to a biological activity.

## 1. Method overview

- 1) Defining **pharmacophoric features**.
- 2) Representing molecular data-set as **pharmacophore graphs**.
- 3) Extracting **pharmacophores** by mining the pharmacophore graphs [3] via Norns<sup>a</sup>.
- 4) Structuring pharmacophore set into a **concept lattice** [1].
- 5) Computing quality of each pharmacophore by applying the **Pharmacophore Activity Delta (PAD)** measure, the latter being based on the *Growth Rate*.

a. <https://forge.greyc.fr/projects/norns>

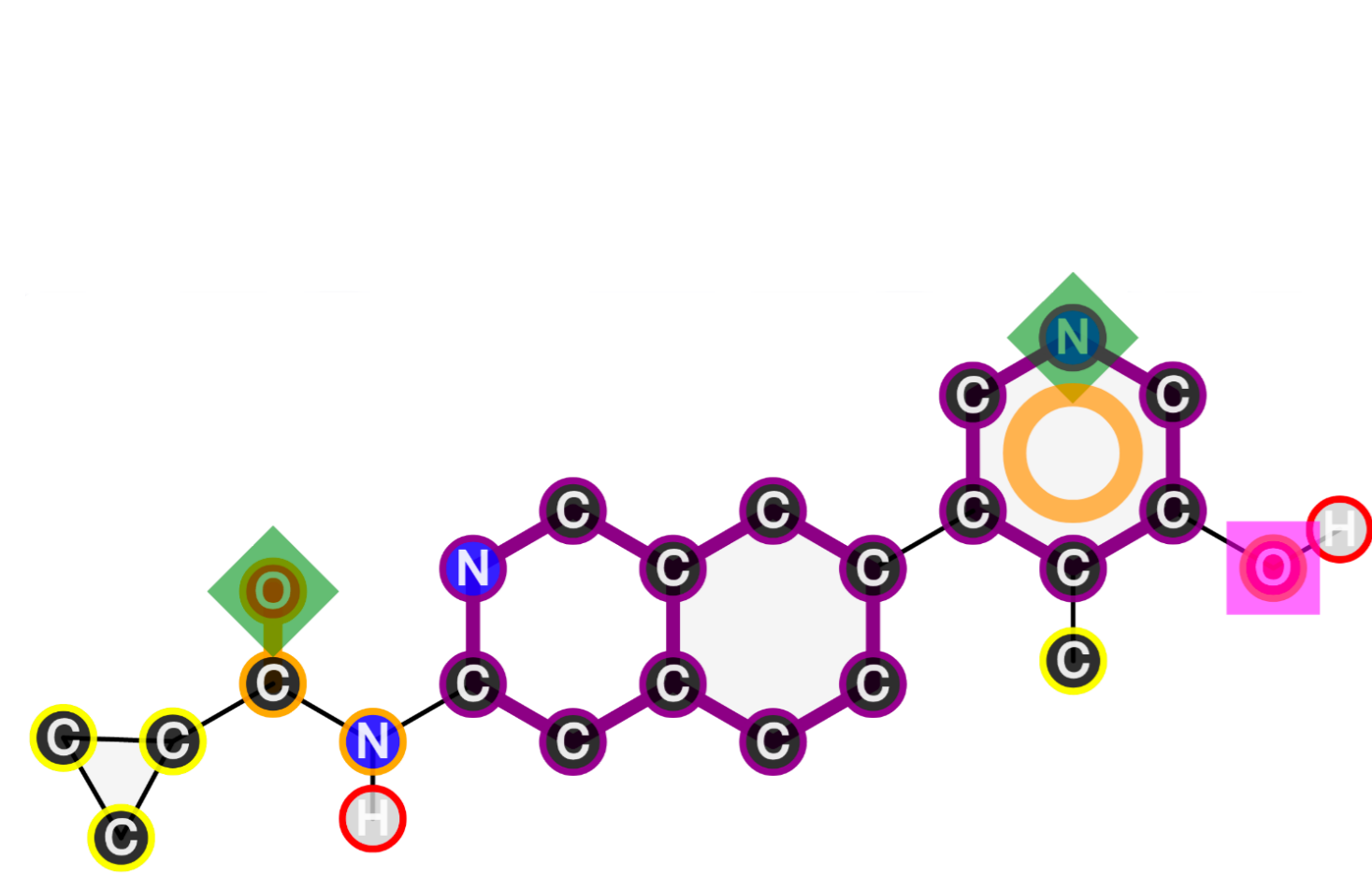


FIGURE 1: Molecule and its pharmacophoric features

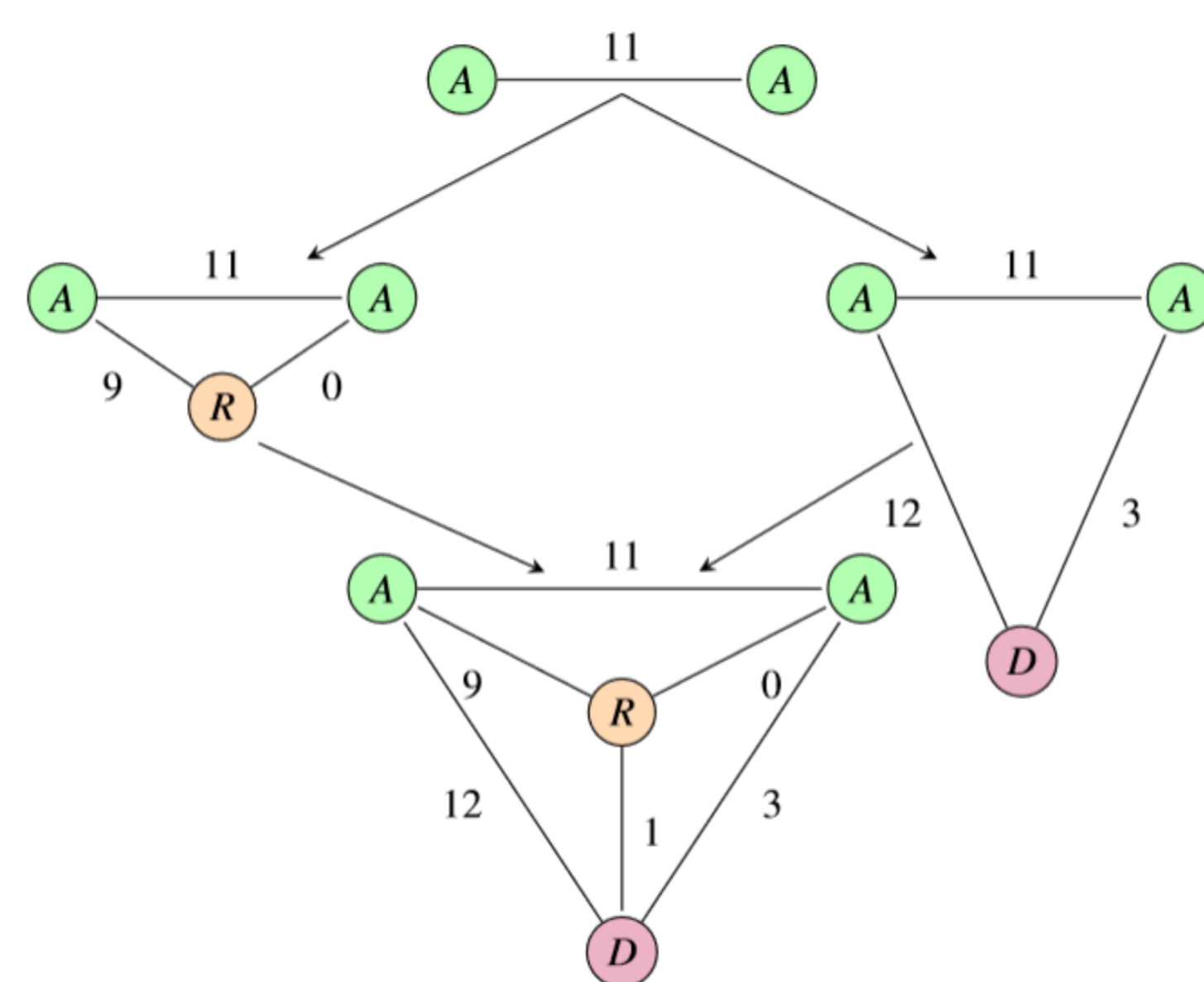


FIGURE 2: Pharmacophore inclusions

## 2. From a pharmacophore lattice to a condensed lattice

In order to structure the pharmacophores of interest composing our molecules, we use a **lattice structure**

- 1) Order relation : sub-graph inclusions.
- 2) Object / Attributes : pharmacophores / molecules.

Then, **the lattice is condensed** by building **equivalence classes** which group pharmacophores occurring in the same set of molecules as soon as they constitute a connected part.

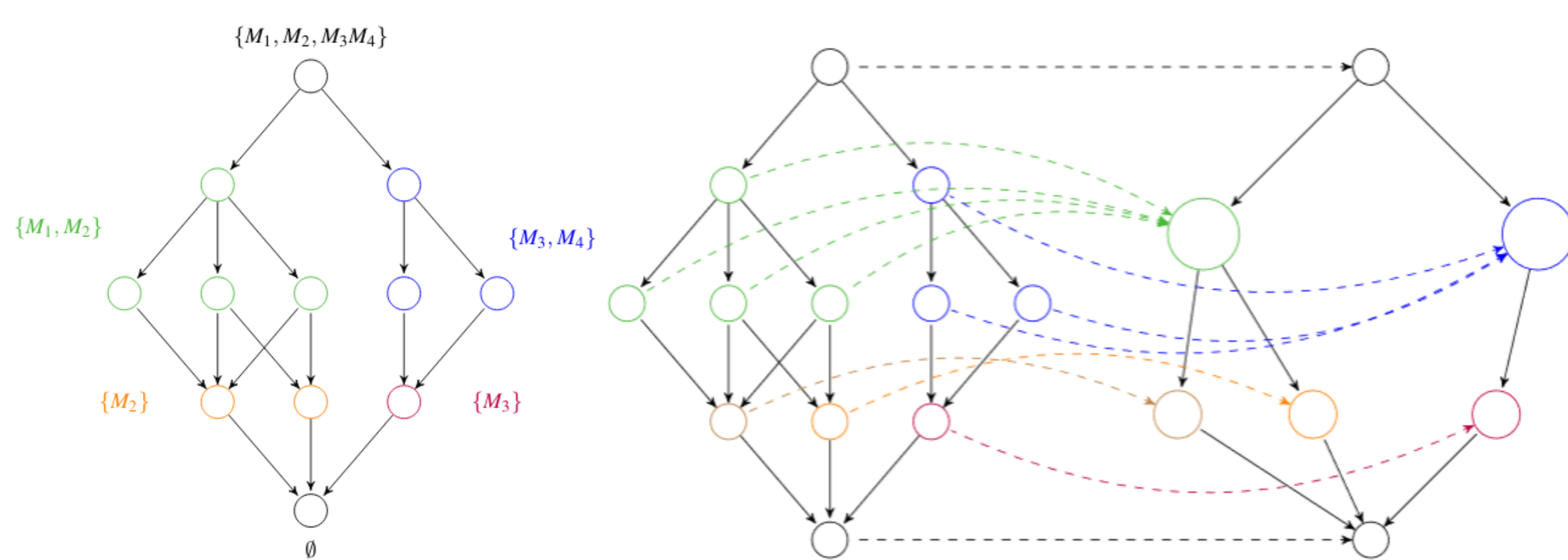
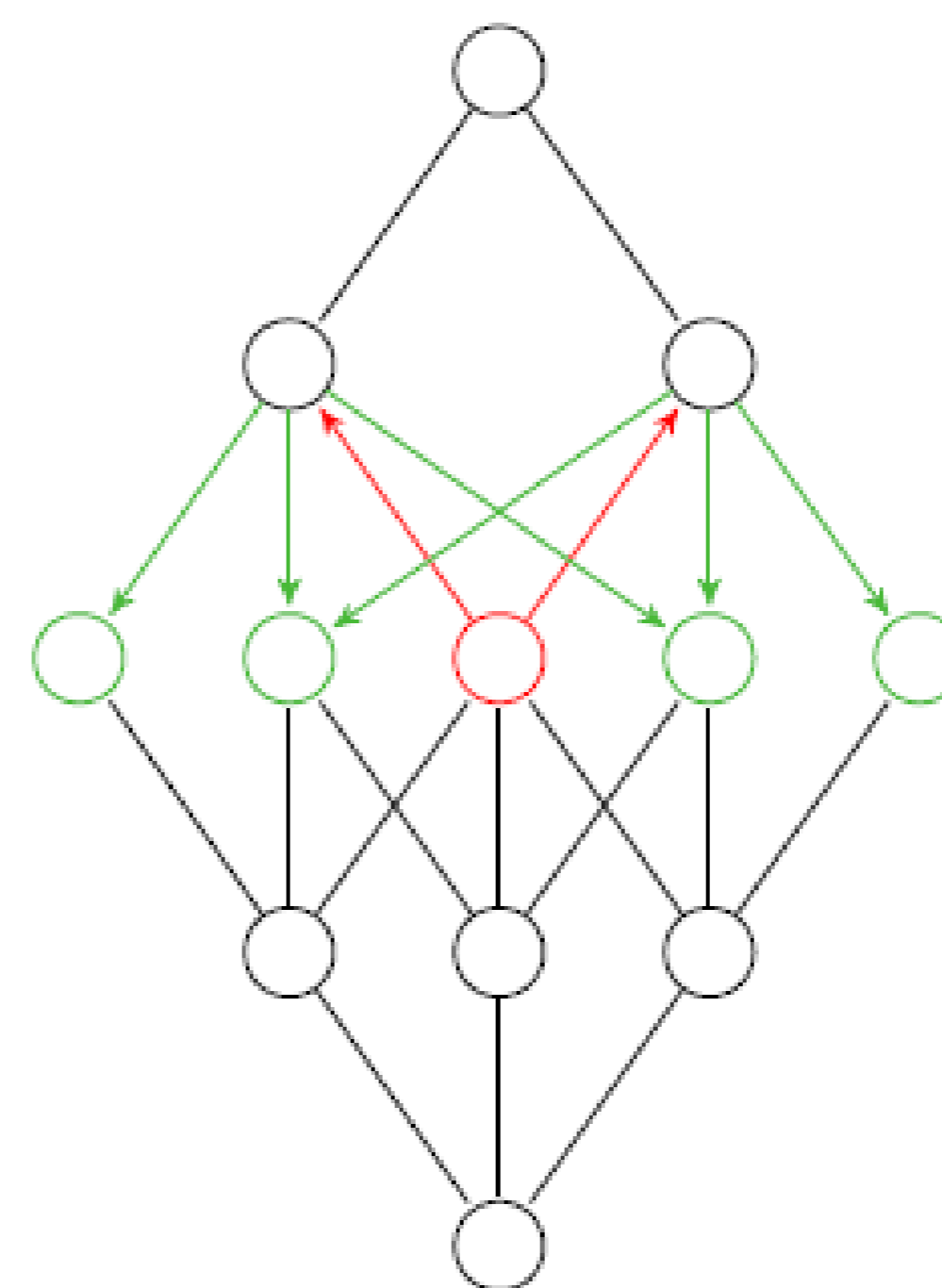


FIGURE 3: Pharmacophore lattice (left) and its corresponding condensed lattice (right)

Orange nodes from left figure have been separated because :

- 1) They are not connected by a common direct ancestor nor a common direct descendant.
- 2) One is not the ancestor nor the descendant of the other.

## 3. PAD : a new quality measure



By using the **condensed lattice** and the **Growth Rates of the pharmacophores**, one computes the **Pharmacophore Activity Delta** :

- 1) Select a node in the lattice (red node).
- 2) Identify the ancestors (red arrows).
- 3) Identify the siblings (green nodes).
- 4) Compute the family standard deviation.
- 5) Compute node delta from its quality and the standard deviation.

PAD has been inspired by the notion of

an **Activity Cliff** [2]

## 4. Results

Data-set : 1,485 ligands from ChEMBL-ABL.

Lattice-set : 122,363 pharmacophores extracted from the data-set.

Pharmacophores have been extracted with the following parameters :

- 1) Minimal frequency : 10.
- 2) Size : 1 to 7 features.

PAD have been computed on the condensed lattice with a minimum selecting threshold of 2 times the family standard deviation.

	Order 1	Order 2	Order 3	Order 4	Order 5	Order 6	Total
<b>Selected</b>	6	6	<b>208</b>	<b>212</b>	48	12	<b>492</b>
<b>Discarded</b>	0	301	<b>5,181</b>	<b>8,064</b>	1,487	28	<b>15,061</b>
<b>Total</b>	6	307	<b>5,389</b>	<b>8,276</b>	1,535	40	<b>15,553</b>

TABLE 1: Equivalence classes and PAD selection

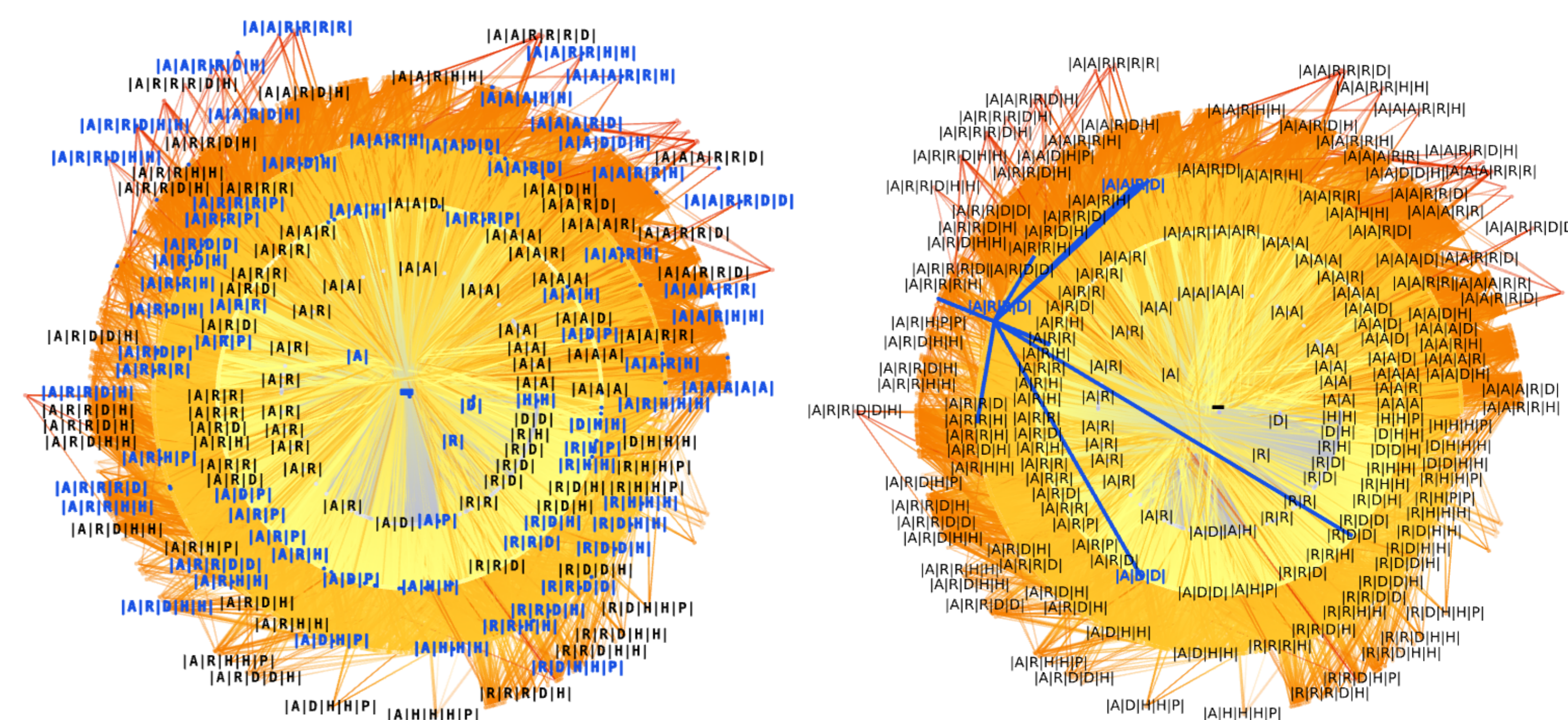


FIGURE 5: Equivalence classes network (left) and equivalence class connection (right)

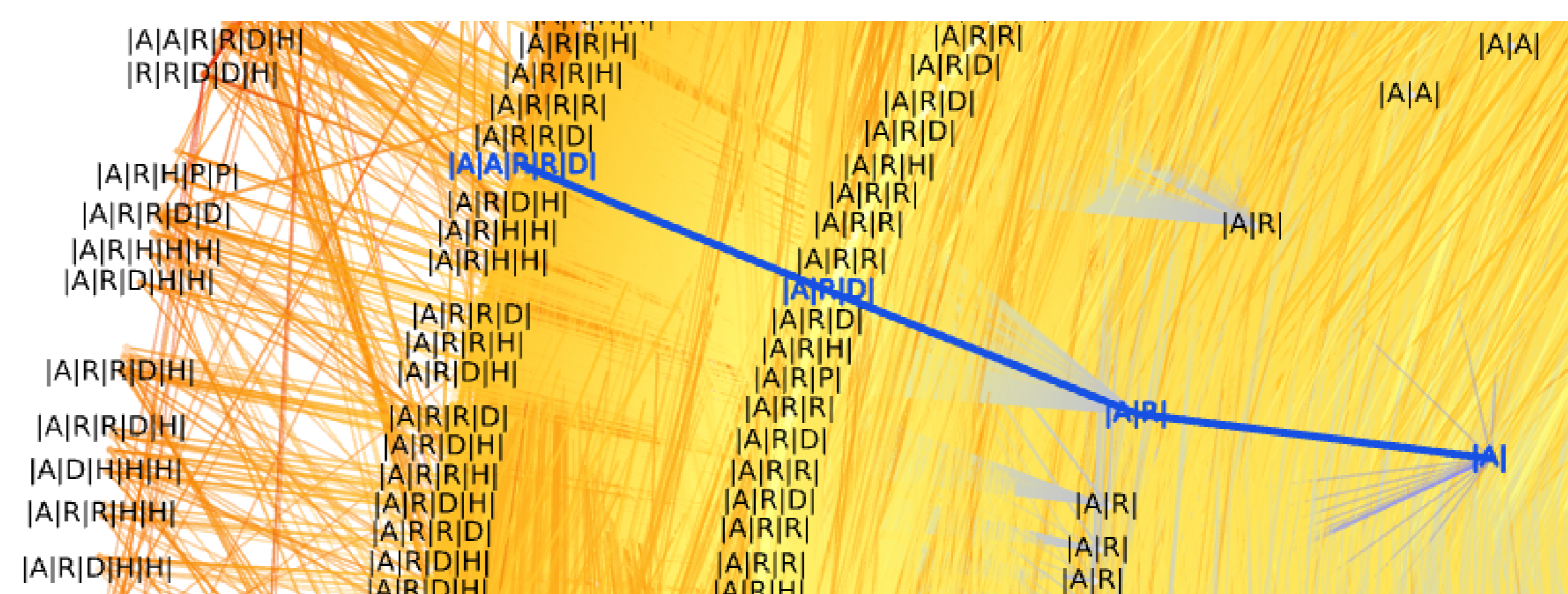


FIGURE 6: Path from a simple class to a complex selected class with Tulip<sup>a</sup>

a. <https://tulip.labri.fr/site/>

## Références

- [1] Sergei O. KUZNETSOV et Sergei A. OBIEDKOV. "Algorithms for the Construction of Concept Lattices and Their Diagram Graphs". In : sous la dir. de Luc De RAEDT et Arno SIEBES.
- [2] Jose L. MEDINA-FRANCO et al. "Characterization of Activity Landscapes Using 2D and 3D Similarity Methods : Consensus Activity Cliffs". In : *Journal of Chemical Information and Modeling* 49.2 (2009), p. 477-491.
- [3] Jean-Philippe MÉTIVIER et al. "The Pharmacophore Network : A Computational Method for Exploring Structure–Activity Relationships from a Large Chemical Data Set". In : *Journal of Medicinal Chemistry* 61.8 (2018), p. 3551-3564.

1. <https://involv.d.greyc.fr/>