



**HAL**  
open science

# Protein Interaction Explorer (PIE): A Web Platform for Exploring Protein-Protein Interactions

Fabien Mareuil, Alexandra Moine-Franel, Anuradha Kar, Michael Nilges,  
Constantin Bogdan Caimbur, Olivier Sperandio

## ► To cite this version:

Fabien Mareuil, Alexandra Moine-Franel, Anuradha Kar, Michael Nilges, Constantin Bogdan Caimbur, et al.. Protein Interaction Explorer (PIE): A Web Platform for Exploring Protein-Protein Interactions. JOBIM2024: Journées Ouvertes en Biologie, Informatique, et Mathématiques - 2024, Jun 2024, Toulouse (FRANCE), France. pasteur-04619520

**HAL Id: pasteur-04619520**

**<https://pasteur.hal.science/pasteur-04619520>**

Submitted on 20 Jun 2024

**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.



# Protein Interaction Explorer (PIE): A Web Platform for Exploring Protein-Protein Interactions

Fabien Mareuil<sup>1</sup>, Alexandra Moine-Franel<sup>2,3</sup>, Anuradha Kar<sup>2</sup>, Michael Nilges<sup>2</sup>, Constantin Bogdan Caimbur<sup>2</sup>, and Olivier Sperandio<sup>2</sup>

<sup>1</sup> Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub F-75015 Paris, France

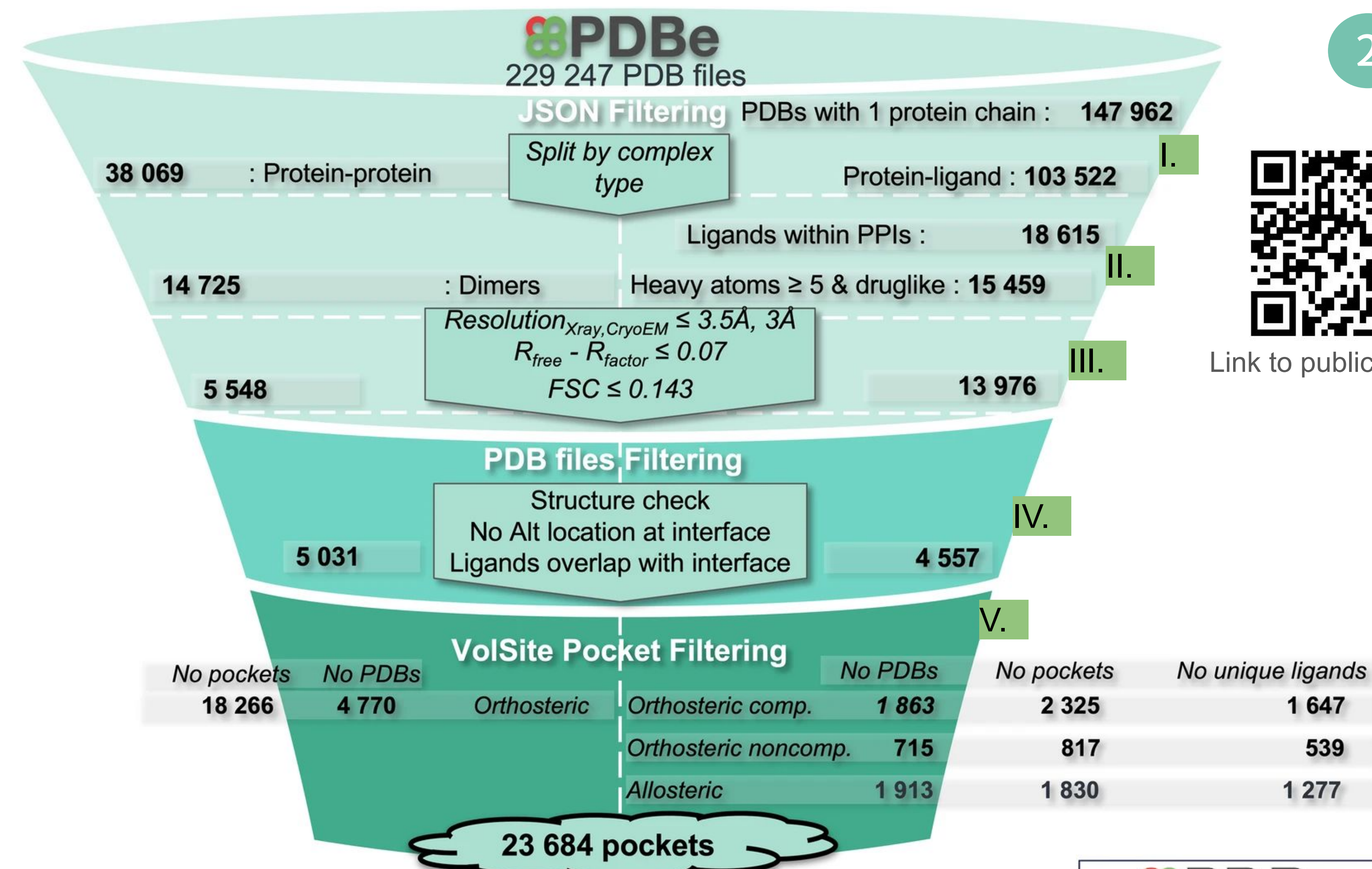
<sup>2</sup> Institut Pasteur, Université Paris Cité, Structural Bioinformatics Unit, Department of Structural Biology and Chemistry, CNRS UMR3528, Paris F-75015, France

<sup>3</sup> Collège Doctoral, Sorbonne Université, Paris F-75005, France

## 1 Context

Protein-protein interactions (PPI) are a key focus in drug discovery. Their 3D structures contain extensive information that can be utilized to prioritize drug target selection, identify druggable binding pockets, and ultimately discover new PPI inhibitors (iPPI). To leverage these 3D data, we present Protein Interaction Explorer (PIE) a web-based tool specifically crafted to help structure-based drug discovery focused on protein-protein interactions (PPIs).

Constructed upon extensive structural data encompassing thousands of heterodimer complexes, including those successfully liganded, PIE provides a comprehensive suite of tools dedicated to aid decision-making in PPI drug discovery. PIE enables biologists/chemists/bioinformaticians to identify and characterize crucial factors such as the presence of binding pockets or functional binding sites at the interface, predicting hot spots, and foreseeing similar protein-embedded pockets for potential repurposing efforts.



## 2 Create a dataset of heterodimer complexes

We setup a quality control pipeline to mine only pertinent structural data:

- I. Select heterodimer complex and protein/ligand complex associated with heterodimer(s)
- II. Select only drug-like PPI-related ligands
- III. Select only complexes with state-of-the-art structural biology criteria
- IV. Select proteins with ligand(s) associated with heterodimer(s) in which one of the ligands is located at the heterodimer interface (distance threshold: 6 Å)
- V. After the pocket detection (VolSite [2]), select proteins with a ligand positioned at less than 1Å of a pocket of a superimposed heterodimer partner.

The retrieved data with all calculated pocket properties (Volsite) are saved in our database and used to evaluate a pocket similarity (PSI) based on the Euclidean distance of normalized properties.

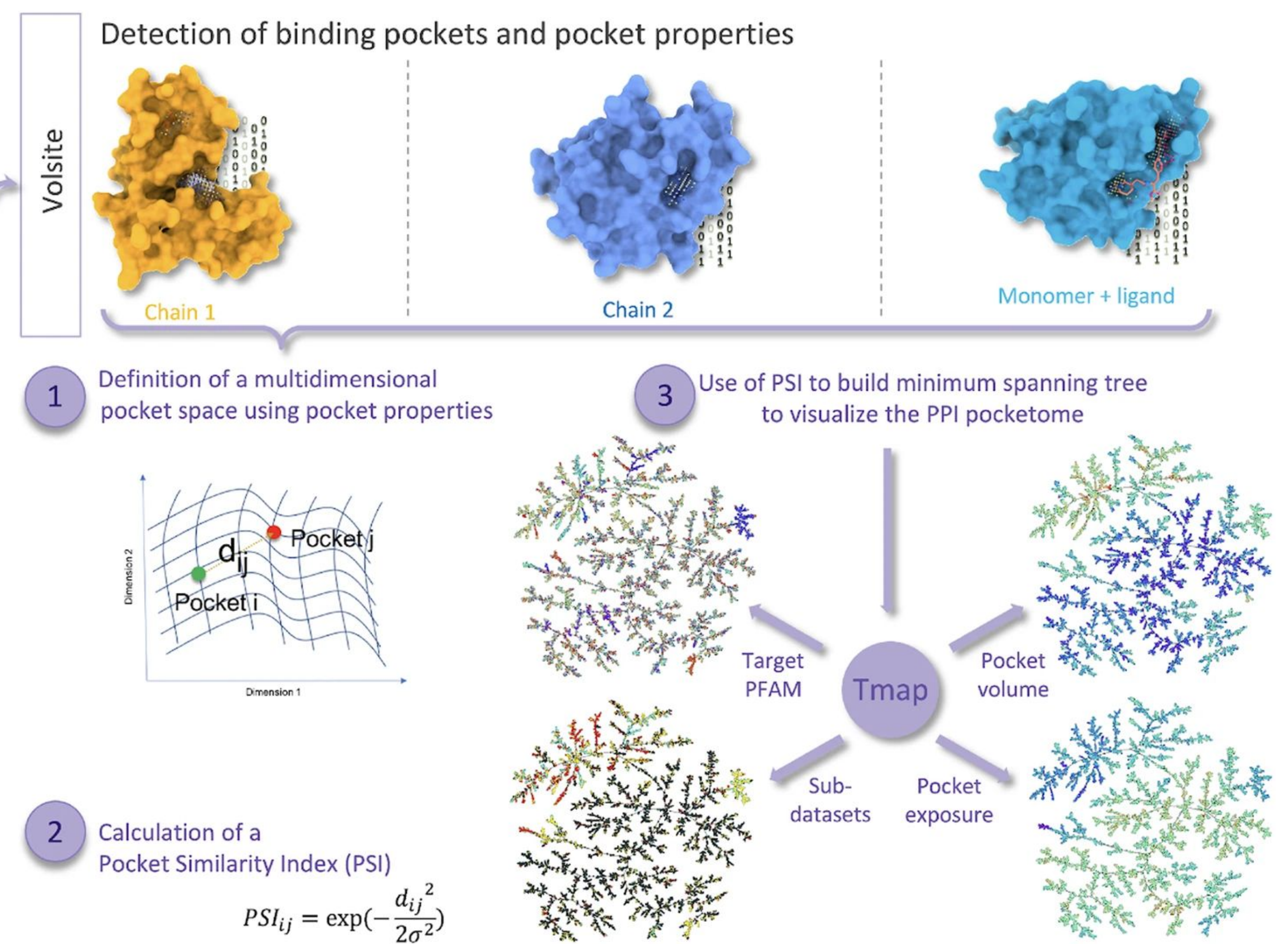
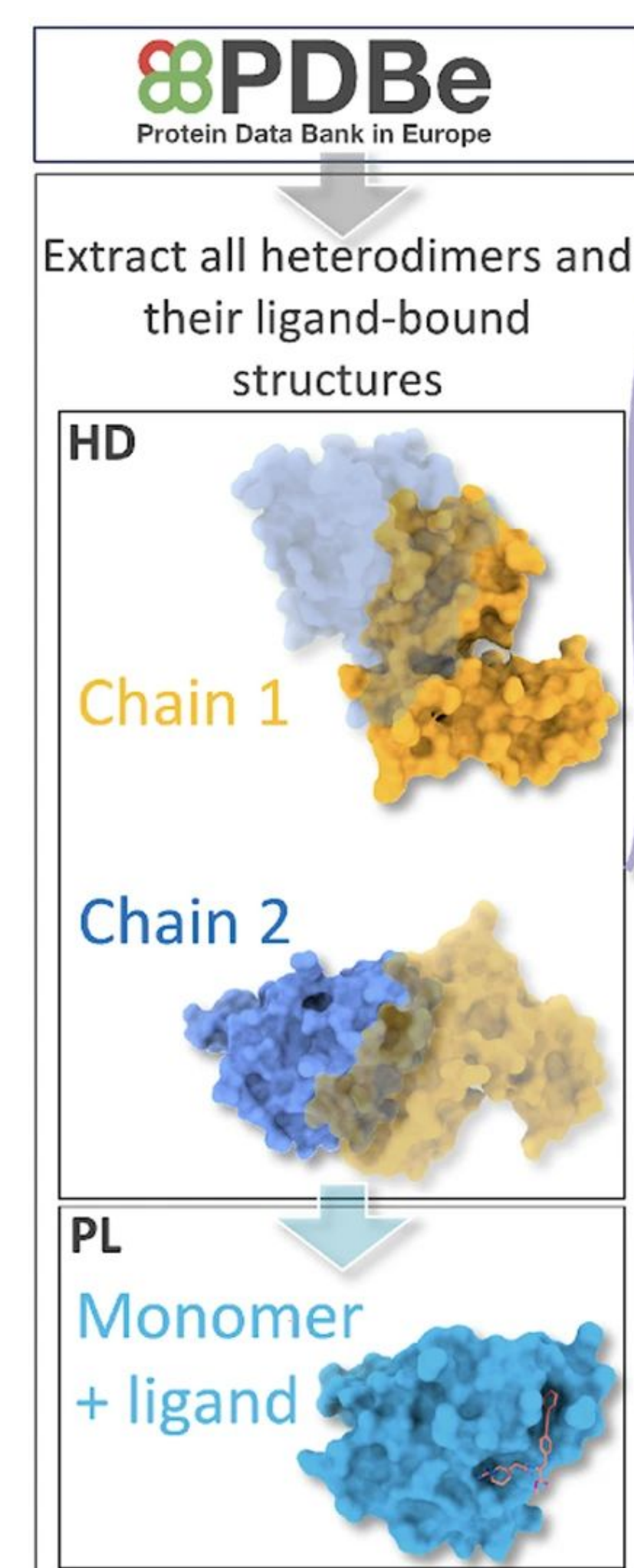
## 3 iPPI-DB PIE: Navigate the pocket

We supplemented our dataset with:

- **Hot Spots Prediction**, FoldX(v5) [3] software was employed to assess the impact of mutations on the stability of protein complexes. Residues exhibiting a significant change in free energy are categorized as critical or warm hot spots.
- **Functional Binding Site Prediction**, InDeep [4] was utilized for predicting functional binding sites.
- **Pocketome Visualization**, represented as a minimum spanning tree using the TMAP [5] tool. This method selectively chooses pairs from the complete PSI matrix to span the entire dataset optimally, minimizing the total distance. The resulting "tree" provides a simplified yet powerful visual representation reflecting local proximity (pocket similarity) in the high-dimensional pocketome space.



Link to publication [4]



## 4 Conclusion

Protein Interaction Explorer allows to :

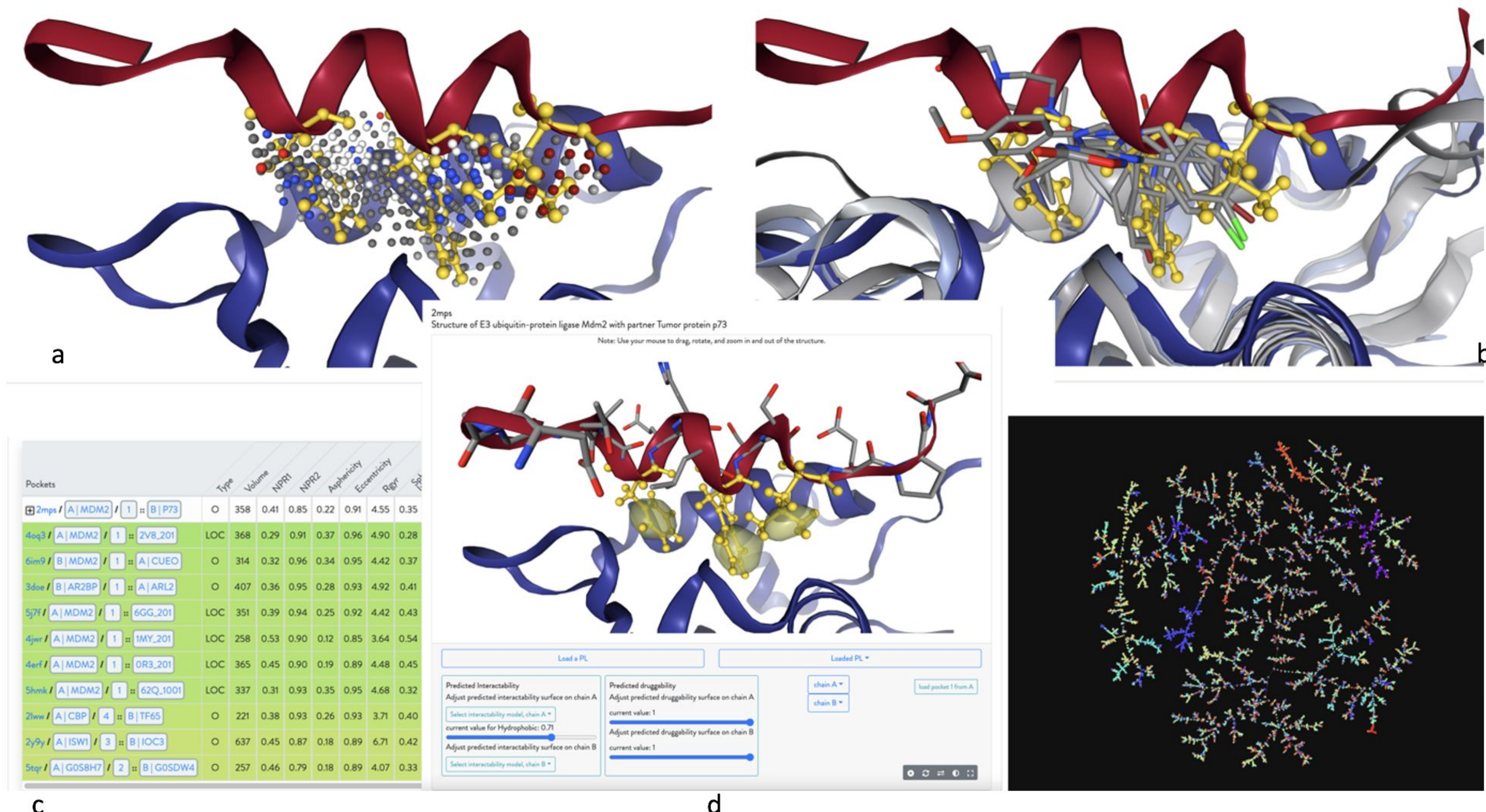
- a. Query PDB or UniProt to visualize pockets
- b. Superimpose all available heterodimer structures and liganded structures.
- c. List similar binding pockets and assess possible partner repurposing
- d. Visualize InDeep annotations
- e. Navigate an interactive PPI-related pocketome using TMAP

With the inclusion of this new dataset in the iPPI-DB database, we allow the user to go beyond the query capabilities already available, and to query it from a pocket perspective. Users can now query the dataset and, based on the aforementioned pocket properties, explore the pocket space in addition to the PPI chemical space.

Come and explore



Link to iPPI-DB web site



[1]: Moine-Franel, A., Mareuil, F., Nilges, M., Caimbur, C. B., & Sperandio, O. (2024). A comprehensive dataset of protein-protein interactions and ligand binding pockets for advancing drug discovery. *Scientific Data*, 11(1), 402. <https://doi.org/10.1038/s41597-024-03233-z>

[2]: Desaphy, J., Azdimousa, K., Kellenberger, E., & Rognan, D. (2012). Comparison and druggability prediction of protein-ligand binding sites from pharmacophore-annotated cavity shapes. <https://doi.org/10.1021/ci300184x>

[3]: Guerois, R., Nielsen, J. E., & Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *Journal of molecular biology*, 320(2), 369-387. [https://doi.org/10.1016/S0022-2836\(02\)00442-4](https://doi.org/10.1016/S0022-2836(02)00442-4)

[4]: Mallet, V., Checa Ruano, L., Moine Franel, A., Nilges, M., Duart, K., Bouvier, G., & Sperandio, O. (2022). InDeep: 3D fully convolutional neural networks to assist in silico drug design on protein-protein interactions. *Bioinformatics*, 38(5), 1261-1268. <https://doi.org/10.1093/bioinformatics/btab849>

[5]: Probst, D., & Reymond, J. L. (2020). Visualization of very large high-dimensional data sets as minimum spanning trees. *Journal of Cheminformatics*, 12(1), 12. <https://doi.org/10.1186/s13321-020-0416-x>