iPPI-DB: A community-based database of Protein-Protein Interaction modulators

Torchet, R., [1] Moine-Franel, A., [1,2,3] Druart, K.,* [1,2,3] Borges, H., [1,2,3] Doppelt-Azeroual, O.,[1] Mareuil, F.,[1] Ménager, H.,[1] and Sperandio, O.[1,2,3]

* karen.druart@pasteur.fr

1. C3BI, Institut Pasteur 28 rue du Dr Roux, 75015 Paris, France.

2. Structural Bioinformatics Unit, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France.

3. CNRS UMR 3528, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France.

With about 130,000 binary PPIs and possibly more just in humans[1,2], the development of drugs targeting Protein-Protein Interactions (PPIs), represents a significant step toward expanding the druggable genome[3] and a possible leverage on the pharmacological modulation of disease-associated cellular pathways. In 2012, we developed iPPI-DB[4,5], a database manually curated from the scientific literature that contains the structure, some physicochemical characteristics, the pharmacological data and the profile of the PPI targets of several hundred modulators of protein-protein interactions.

We have completely revisited the web interface to query the database (http://ippidb.pasteur.fr). iPPI-DB can now be queried through an interface that allows simultaneous combinations of many chemical and pharmacological criteria. Users can easily combine multiple filters to constitute complex queries, then share the URL corresponding to this query with collaborators, and download the corresponding data. While the chemical similarity search has been preserved in this version, other criteria have been added. Boolean criteria can be checked, such as the availability of certain types of data on the compounds, e.g X-ray crystallographic structure, cellular assay, pharmacokinetic data, or the compliance with chemistry rules like the Lipinski's RO5[6]. Threshold-based criteria can also be applied such as to comply with certain ranges of physico-chemical properties or activities. Query results can be displayed as thumbnails, as a list of cards, or as a table, all sortable. These different types of display, along with multiple sorting options, provide a wide range of accesses to the data from different perspectives while allowing the refinement of a given query. An individual card is available for each compound in iPPI-DB. It provides all available data on the compound. This includes its structure, its physicochemical/pharmacological profiles, but also the structures and therapeutic areas of its 15 most chemically similar marketed or investigated drugs.

The new version of iPPI-DB now provides a dedicated community-based graphical interface to enter new data in the database, constituting an important improvement over the previously complex and largely manual process. This interface has been designed with the aim of facilitating as much as possible the contributions, even from experts who are not familiar with the technical aspects of this database. Each contribution is based on the description of the content of a publication or a patent. Through a wizard-based web interface, users provide in a step-by-step process, the architecture of the PPI complex(es), the chemical compounds tested for modulation, and the various assays in which those compounds were tested. The contribution interface only requests minimal information from users in order to reduce the risks for errors and facilitate contributions: whenever contributors provide some information, the server automatically retrieves additional details from other reference databases, using either native HTTP requests or the BioServices package[8]. Following the sketching of compounds within an embedded Chemaxon Marvin Sketch or the copy of IUPAC name, the chemical structures are compiled and processed through an automated pipeline that calculates physico chemical properties and similarities with known drugs using the Chemaxon libraries.

[1] Khanna, V. and Ranganathan, S. J. Cheminform, 2011, 3, 30 1-14.
[2] Venkatesan, K. *et al.* Nat. Methods, 2009, 6(1), 83-90.
[3] Laraia, L., McKenzie, G., *et al.* Chem. Biol, 2015, 22, 689–703.
[4] Labbé, CM., Laconde, G., *et al.* Drug Discovery Today, 2013, 18, 958–968.
[5] Labbé, CM., Kuenemann MA., *et al. Nucleic Acids Res.* 44, D542–D547 (2016).
[6] Lipinski, C. A., Lombardo, F., *et al.* Adv Drug Deliv Rev, 2001, 46, 3–26.
[7] Bateman, A. *Nucleic Acids Res.,* 2019, 47, D506–D515.
[8] Cokelaer, T., Pultz, D., *et al. Bioinformatics,* 2013, 29(24), 3241-3242.