

# MACHINE LEARNING-GUIDED DE NOVO DESIGN OF A PROPRIETARY PPI MODULATOR LIBRARY

Iolanda Micco<sup>1</sup>, Davide Sala<sup>1</sup>, Serena Meini<sup>2</sup>, Federico Munafo'<sup>2</sup>, Maria Pia Catalani<sup>1</sup>

<sup>1</sup> Axxam S.p.A., Via Meucci 3, 20091 Bresso (Milan), Italy

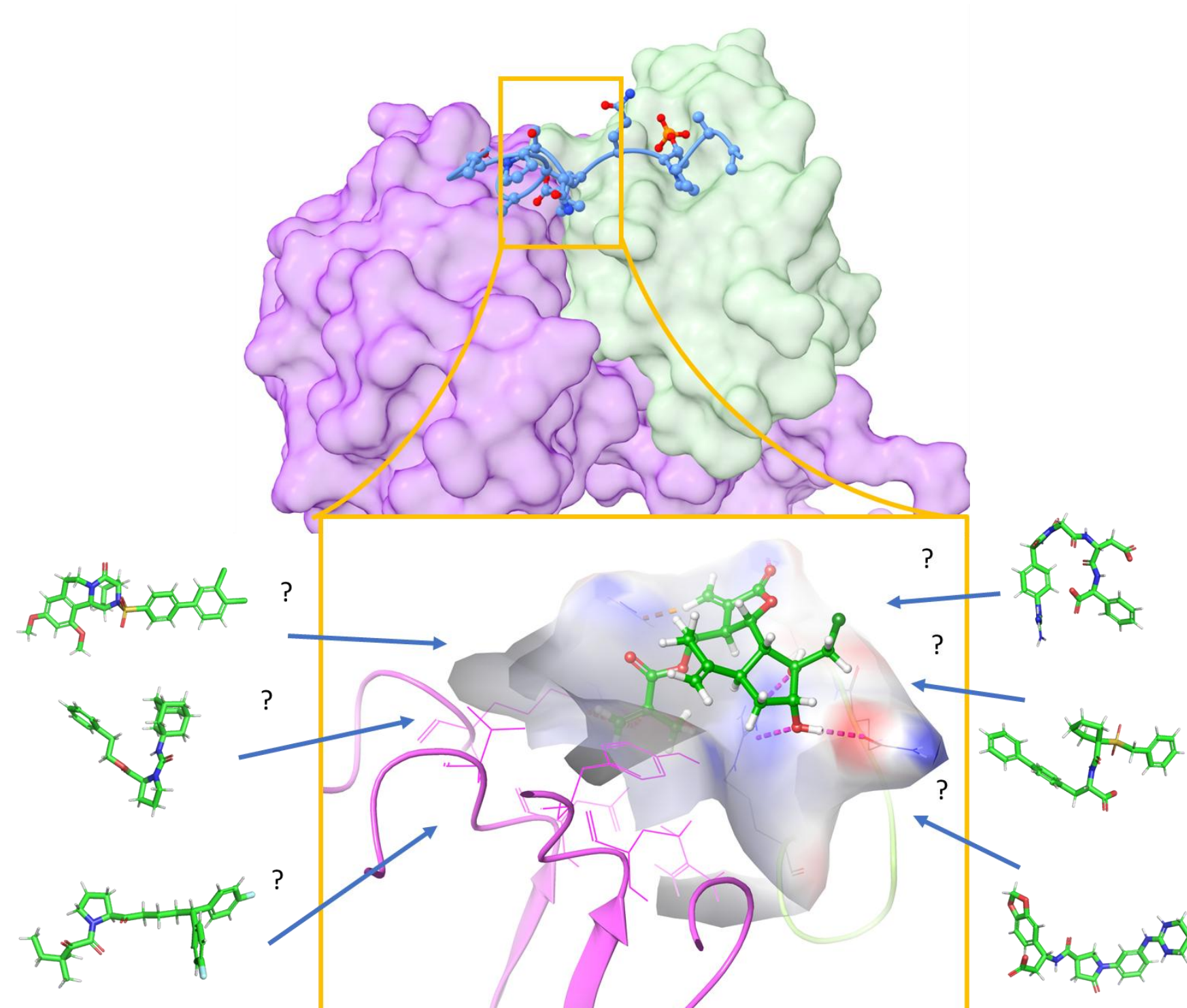
<sup>2</sup> Axxam S.p.A., Via Provinciale Schito 131, 80058, Torre Annunziata (NA, Italy)

Protein-Protein Interactions (PPI) play a crucial role in disease modulation yet targeting them with small molecules presents significant challenges due to the complex and often dynamic nature of PPI interfaces, which differ from typical ligand-binding sites targeted by conventional drugs. Therefore, small molecules targeting PPI often deviate from representative drug-like characteristics. **A multi-step approach, involving machine learning-guided methods was used for the de novo design of a library of PPI modulators with improved drug-like properties.**

## Protein-Protein Interactions

Targeting **protein-protein interactions** (PPIs) has recently emerged as an attractive strategy for designing innovative drugs for highly complex diseases.<sup>[1]</sup> Despite the advancements in the field, the identification of PPI inhibitors remains a challenging task, mainly due to the intrinsic characteristics of the protein-protein interface<sup>[2,3]</sup>: generally **flat, hydrophobic and deficient in polar residues**. For this reason, **small molecules targeting PPI often deviate from representative drug-like characteristics**, and most of them satisfy the so-called rule-of-4:<sup>[4,5]</sup>

- **MW > 400**
- **LogP > 4**
- **Num of rings > 4**
- **Num of HBA > 4**



## Generative Machine Learning in MedChem

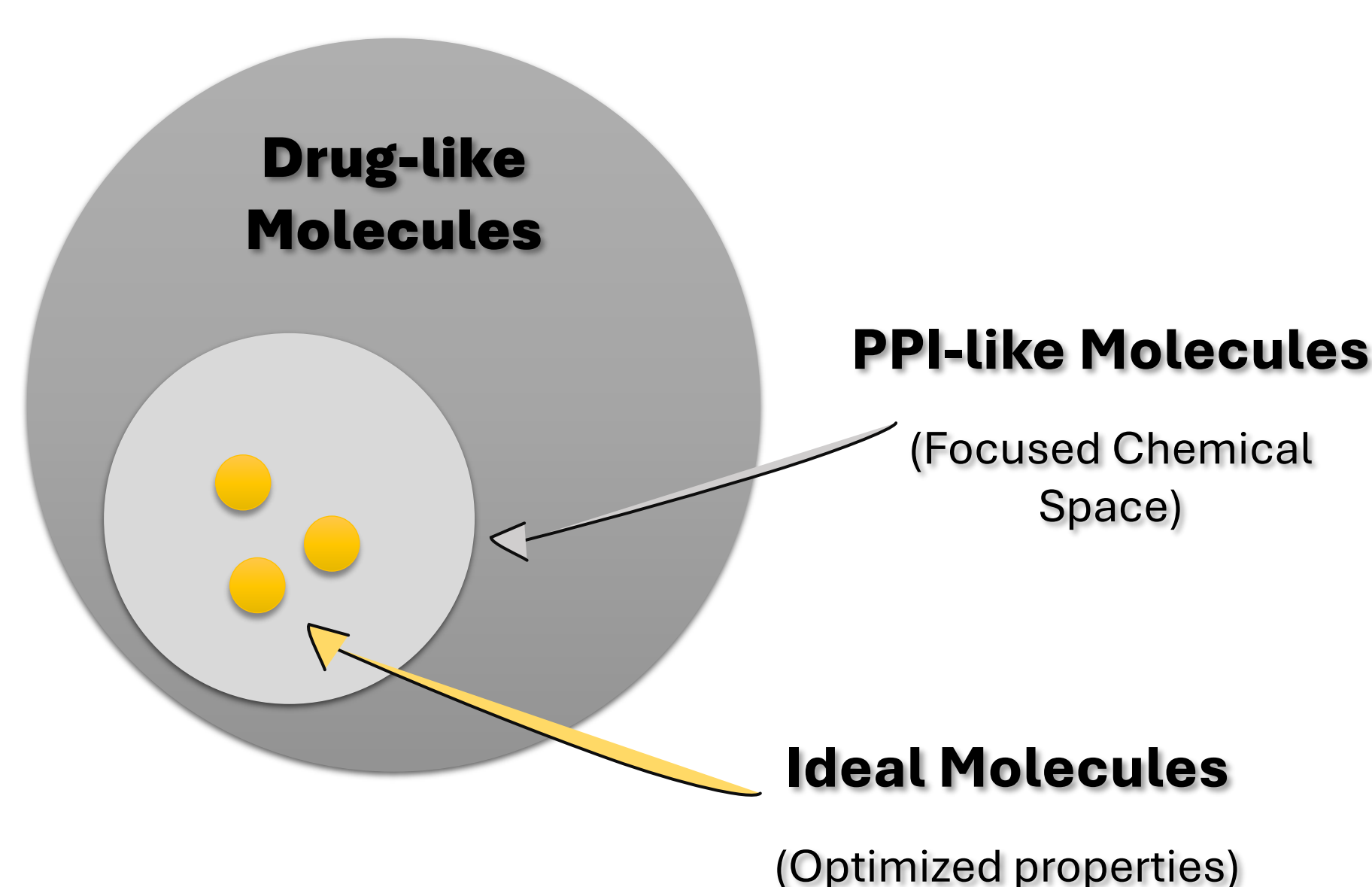
**Generative machine learning** is revolutionizing drug discovery by enabling the *de novo* design of small molecules with specific molecular and therapeutic properties.

By using innovative models like **recurrent neural networks (RNNs)** and **transformers**, it is now possible to generate novel molecular structures from scratch. These models can be fine-tuned and enhanced with other optimization techniques, such as **reinforcement learning** and **transfer learning** to obtain molecules with the desired drug characteristics.

While challenges remain, the integration of generative models and algorithms is significantly accelerating the process of discovering new drug candidates.

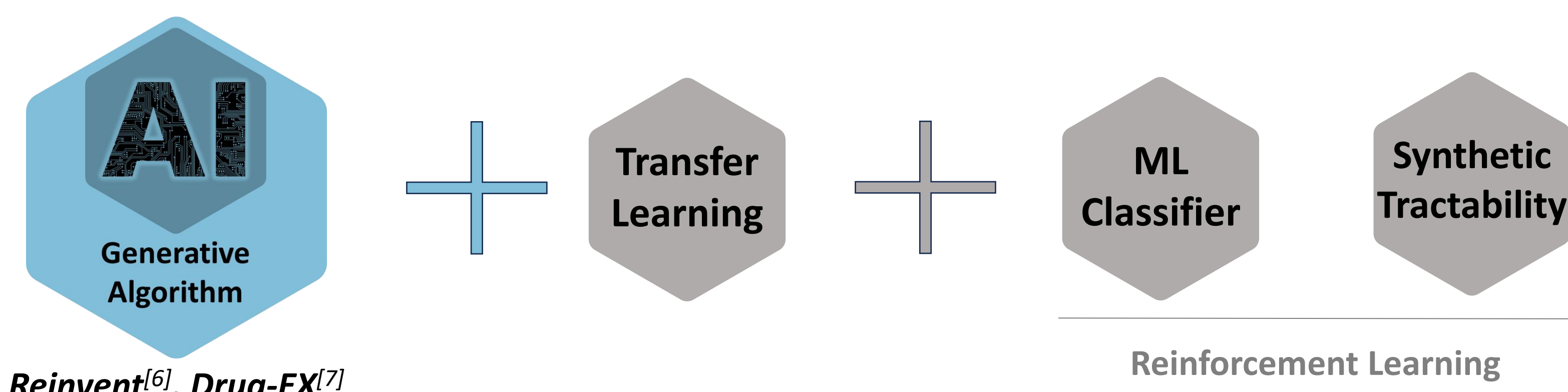
## Data Curation

**Data is crucial** to the success of any Machine Learning approach. We started from the collection of **molecules annotated to be PPI modulators** from different sources: 2P2ldb, iPPI-DB, Timbal, ChEMBL, and literature.



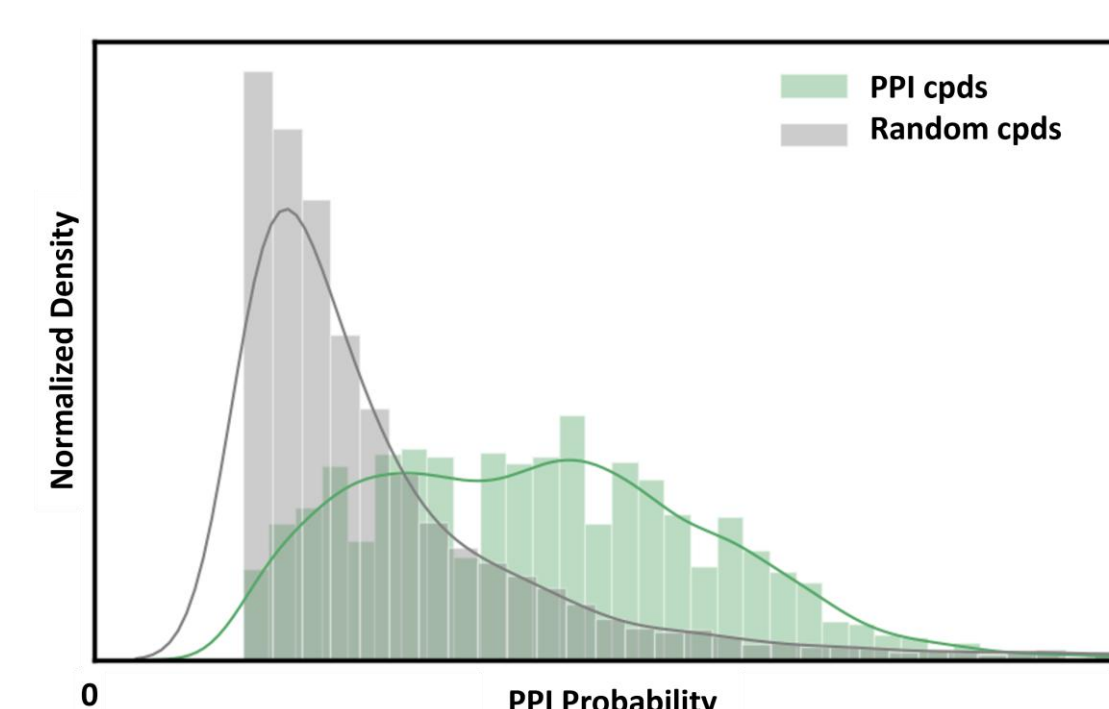
Schematic Representation of Ideal Molecules (Yellow Spheres) within Focused (Light Grey) and General (Dark Grey) Chemical Space.

## De Novo Generation and Optimization



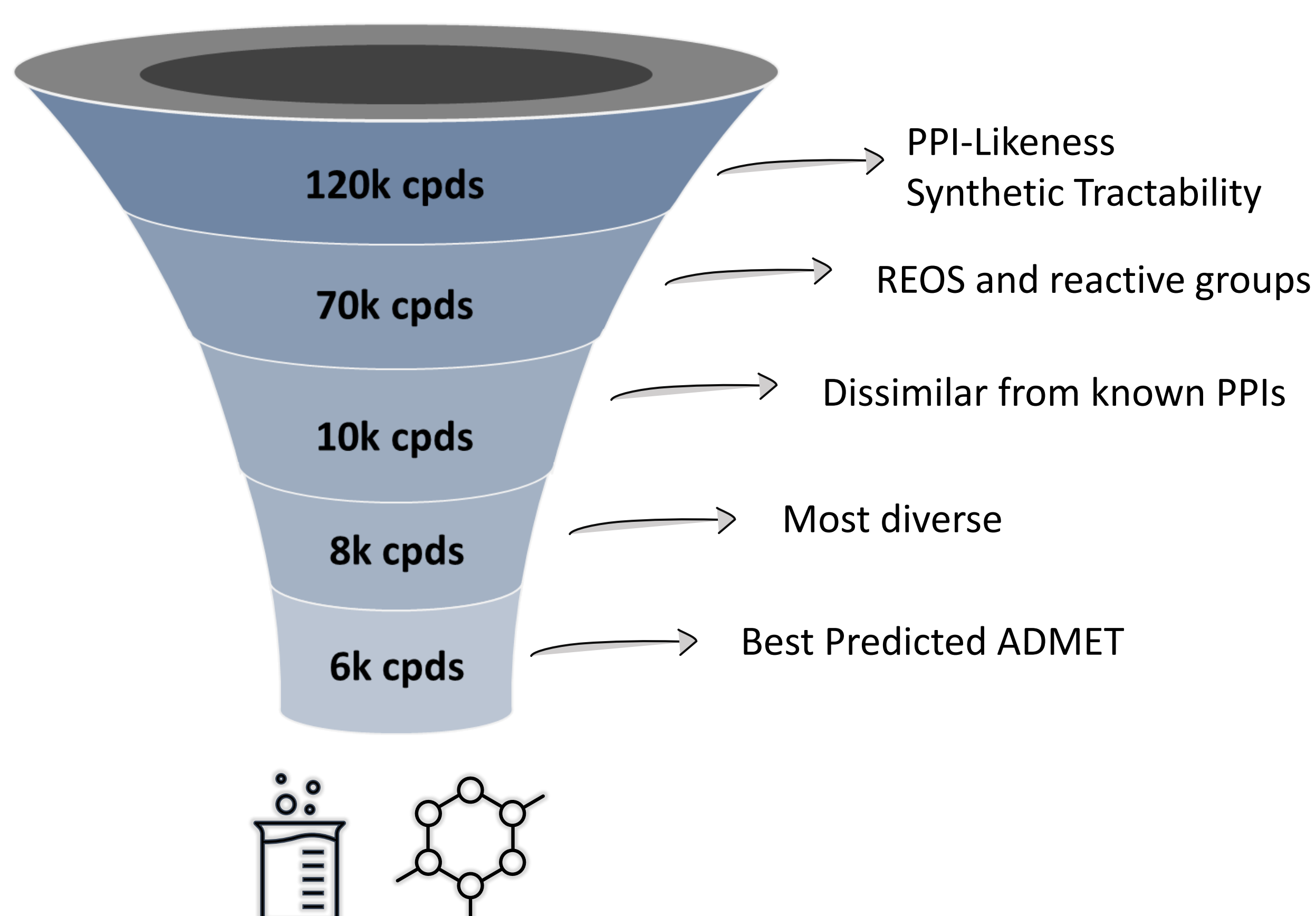
Collected Data was used to optimize the Generative Model in two ways:

1. Transfer Learning: PPI molecules were used to narrow down the chemical space to be explored by the generative algorithm
2. A Machine Learning classifier was trained to distinguish between PPI and non-PPI molecules.
3. Moreover, an algorithm for synthetic tractability was used during the Reinforcement Learning phase.

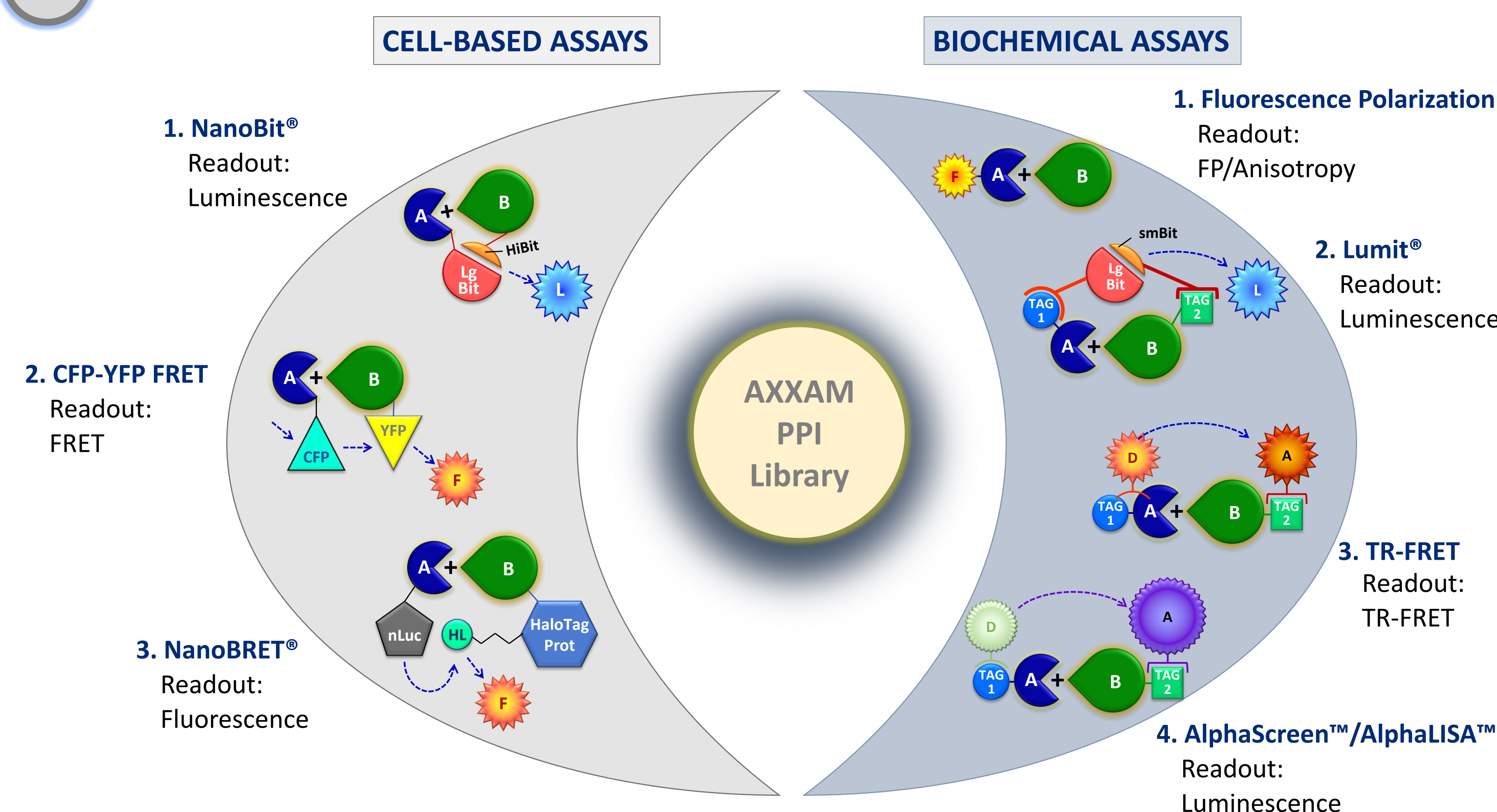


Evaluation of the Machine Learning Classifier's ability to discriminate known PPIs (not used during the training process - Light Green) from random compounds (Light Grey).

## Filtering, Evaluation and Synthesis



## PPI in Axxam



The Axxam Integrated PPI Engine for Hit Discovery

## References

- [1] Nat Rev Drug Discov 2016, 15, 533.  
 [2] Brief Bioinform. 2022 Jul; 23(4): bbac165.  
 [3] Adv Appl Bioinform Chem. 2020 Nov 12;13:11-25  
 [4] Curr. Opin. Chem. Biol. 2011, 15, 475-481.  
 [5] RSC Med. Chem., 2021, 12, 1731-1749.  
 [6] J. Chem. Inf. Model. 2020, 60, 12, 5918-5922.  
 [7] J. Chem. Inf. Model. 2023, 63, 12, 3629-3636.

