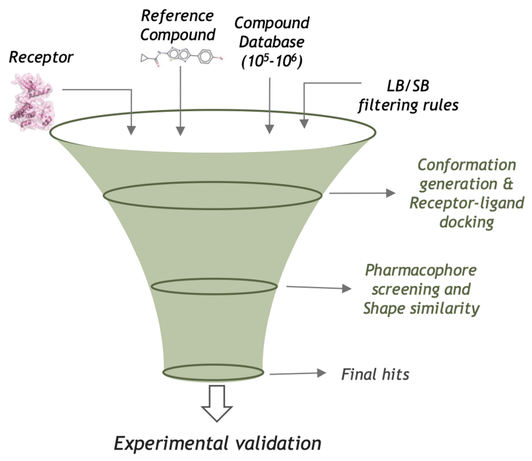
# DockAI: Efficient Exploration of Ultra-large Chemical Spaces Using Active Learning

Maxime Laugeois, Brice Hoffmann, Paul Join-Lambert, Anna Kriukova, Maud Jusot, Ennys Gheyouche, Stéphanie Labouille, Juan Sanz García, Christopher Housseman, Nicolas Brosse, Clarisse Descamps, Guillaume Plum, Nicolas Drizard, Cédric Thao, Hamza Tajmouati, Brian Atwood, Nicolas Do Huu, Quentin Perron, Yann Gaston-Mathé

*Iktos SAS (65 Rue de Prony, 75017 Paris, France) & Iktos Inc (50 Milk Street, Boston MA 02109)*

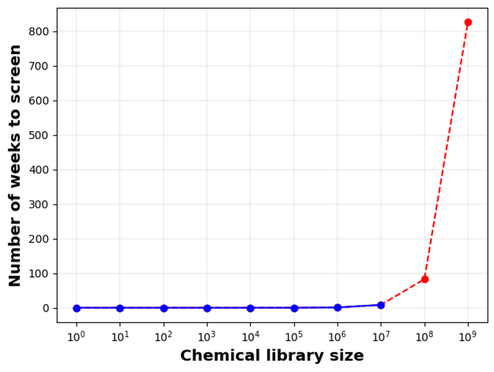
## BACKGROUND

- **Hit discovery** is a crucial step in the drug discovery process, identifying new areas of chemical space for development
- Druglike chemical space is vast (~$10^{60}$); exploration is challenging
- The larger the chemical space to explore within a database, the higher the likelihood of finding hits




**Virtual Screening**: High throughput docking of chemical libraries; allows for rapid screening of large libraries, significantly reducing time & cost associated with experimental screening

- Efficient way to find hits against biological targets
- Several success stories reported over the last two decades

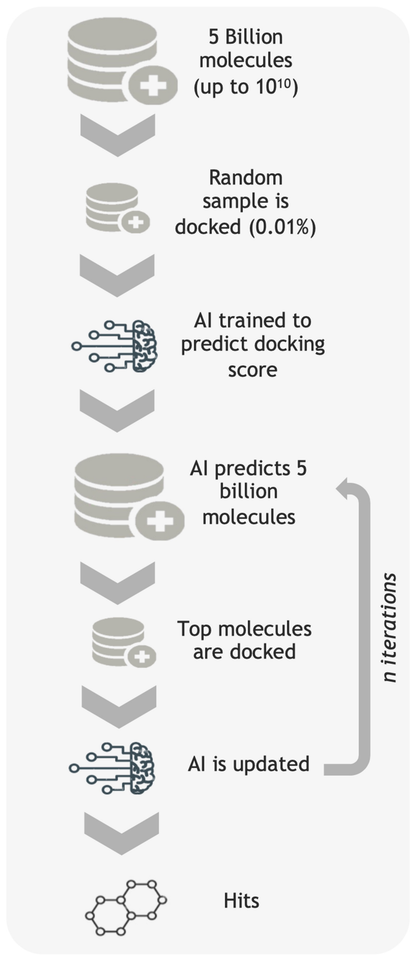



- Limited by the size of the compound database that can be docked
- Until recently, Virtual Screening campaigns typically processed compound databases of available building blocks (~$10^5$-$10^6$)
- **Ultra-large libraries** of make-on-demand building blocks has increased library size (~$10^8$-$10^{10}$ and increasing) making **Virtual Screening computationally intractable**

**Therefore, an algorithmically efficient approach to screen large libraries for hit discovery is desirable**

## IKTOS' SOLUTION: dockAI

***Efficiently explore virtual libraries of chemical providers to find new active compounds rapidly and at low cost***




### dockAI PIPELINE

Docking scores for **billions of molecules** obtained **by docking <1%** of data
Pipeline based on Active Learning approach
Parallel docking on the cloud (AWS)
Hits for further analysis in 1-2 days

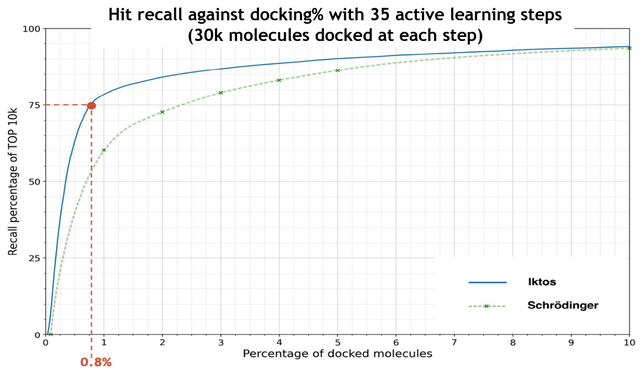
### COMPARISON WITH SCHRÖDINGER

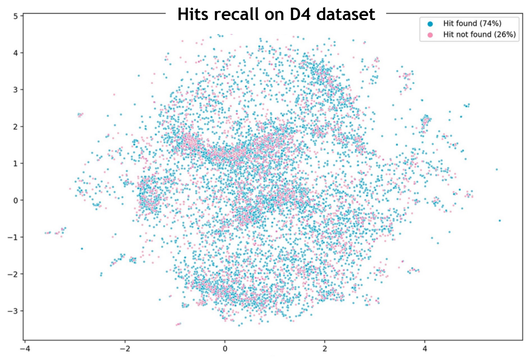
**Initial Dataset**: D4 Dopamine receptor (*n*=140M)
**Active learning iterations**: Iktos→35; Schrödinger→5
**Results**:

- ~75% of hit molecules retrieved after 1M dockings by dockAI
- High diversity; very good coverage of the chemical space

**Conclusion**: dockAI outperforms Schrödinger's comparable approach*

*Data extracted from: *J. Chem. Theory Comput.* **2021**, *17*, 7106-7119.




## dockAI IN COMMERCIAL PROJECTS

### *Project 1: South-Korean biotech*

#### Objective

- **Find druglike hits for a CNS target**; X-ray ligands are derived from natural compounds; very few druglike molecules from competitors
- 20 X-ray structures in PDB; resolution 1-3Å

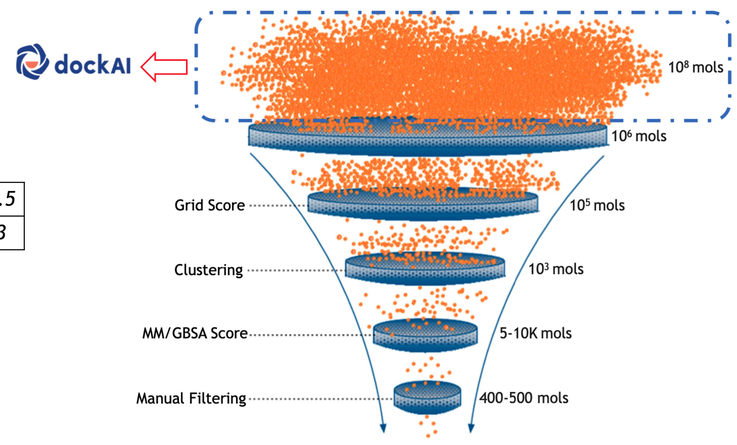
#### Selection criteria

- Docking score
- Key interactions

| | | |
|---|---|---|
| *0 < logP < 5* | *200 < MW < 500* | *QED >= 0.5* |
| *RotBonds <=5* | *HBA <= 7* | *HBD <= 3* |

#### Chemical library size

- WuXi: 100 Million
- Mcule: 160 Million




#### dockAI protocol

- 5 runs, each with a different setup, changing the reward function or X-ray structure
- Docking runs on Iktos' AWS infrastructure
- 2.7 Million dockings in total

#### RESULTS

- **62 molecules** have been synthesized and tested so far
- **5 different scaffolds** in potential hits identified
- **20 molecules <3 µM, up to 10 nM** (30% hit rate)

### *Project 2: Medicines for Malaria Ventures*

*(in partnership with AWS and Intel)*

#### Objective

- **Find druglike hits to target enzyme from Plasmodium;** known X-ray ligands are large and flexible molecules
- 3 X-ray structures from literature and 1 proprietary X-ray structure available

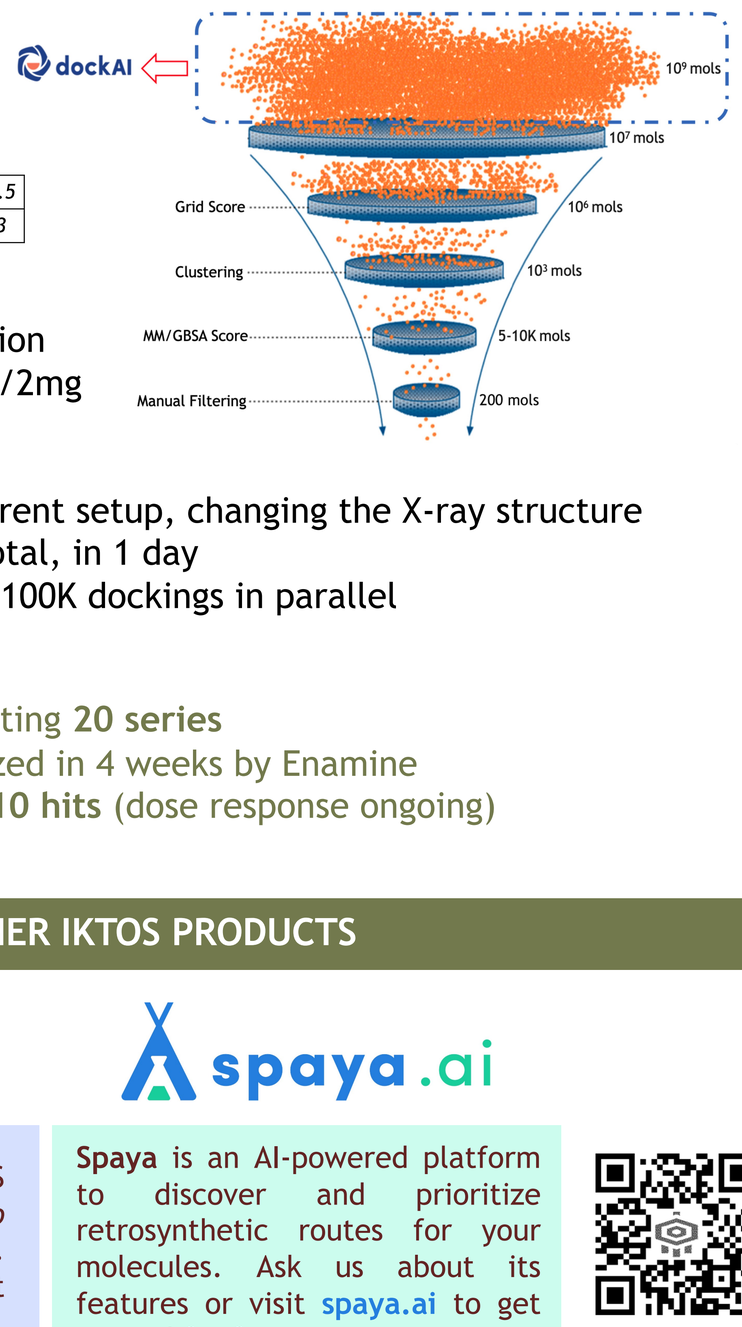
#### Selection criteria

- Docking score
- Key interactions

| | | |
|---|---|---|
| *0 < logP < 5* | *200 < MW < 500* | *QED >= 0.5* |
| *RotBonds <=5* | *HBA <= 7* | *HBD <= 3* |

#### Chemical library size

- Enamine REAL: 5.5 Billion
- Compounds cost <250€/2mg




#### dockAI protocol

- 2 runs, each with different setup, changing the X-ray structure
- 2 million dockings in total, in 1 day
- Docking on AWS, up to 100K dockings in parallel

#### RESULTS

- **200 molecules** representing **20 series**
- **176 molecules** synthesized in 4 weeks by Enamine
- Initial results indicate ~**10 hits** (dose response ongoing)

## OTHER IKTOS PRODUCTS

**Makya** is a chemist-friendly SaaS platform for AI-driven *de novo* 2D drug design focused on MPO. Ask us about its features or visit **makya.ai**

**Spaya** is an AI-powered platform to discover and prioritize retrosynthetic routes for your molecules. Ask us about its features or visit **spaya.ai** to get started for free

Visit **iktos.ai**