

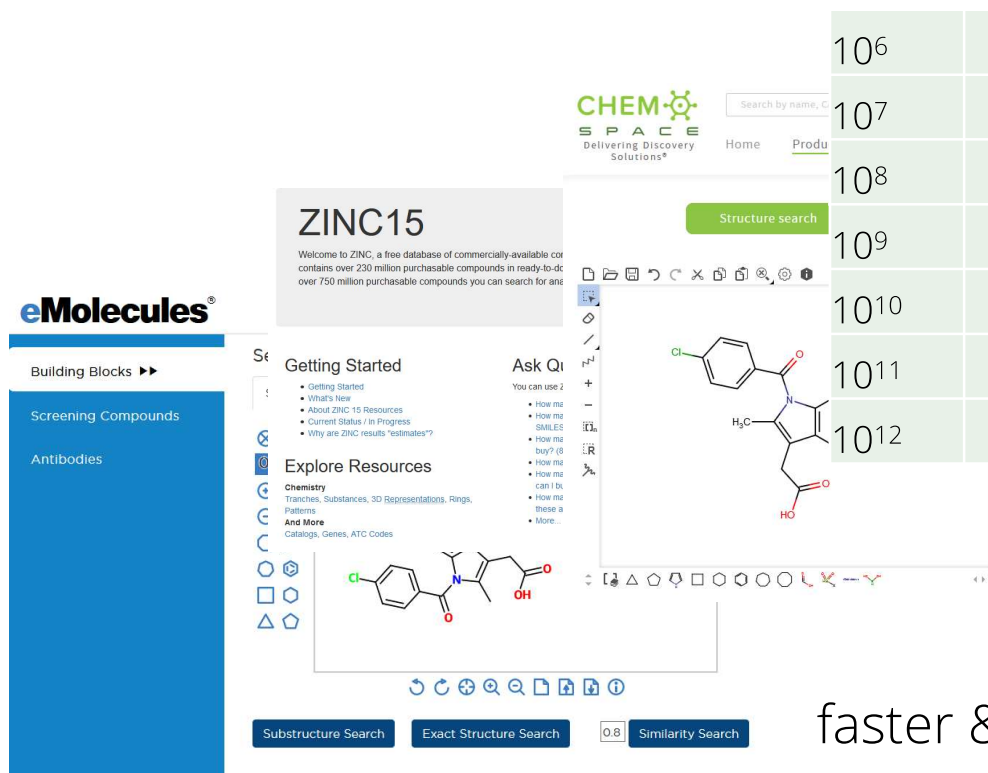
Christian Lemmen

# Efficient 3D exploration of multi-billion compound spaces

Stop searching – start discovering

# Libraries The Classical Way

- Enumeration of XXL-libraries requires
  - loads of memory – particularly for 3D conformers
  - special hardware and/or significant processing time



The screenshot shows the eMolecules ZINC15 database interface. On the left, there is a navigation menu with 'Building Blocks', 'Screening Compounds', and 'Antibodies'. The main content area includes a 'Getting Started' section, 'Explore Resources', and a search interface. The search interface has three buttons: 'Substructure Search', 'Exact Structure Search', and 'Similarity Search'. A chemical structure is displayed in the center, and a table on the right provides performance metrics for different search methods.

$10^6$	1.000.000	3,8 MB	1 sec
$10^7$	10.000.000	38 MB	10 sec
$10^8$	100.000.000	380 MB	1,7 min
$10^9$	1.000.000.000	3,8 GB	17 min
$10^{10}$	10.000.000.000	38 GB	2,8 hrs
$10^{11}$	100.000.000.000	380 GB	28 hrs
$10^{12}$	1.000.000.000.000	3,8 TB	12 days

Search options: Exact match, Substructure, Similarity

0.8 Similarity Search

faster & faster & faster...



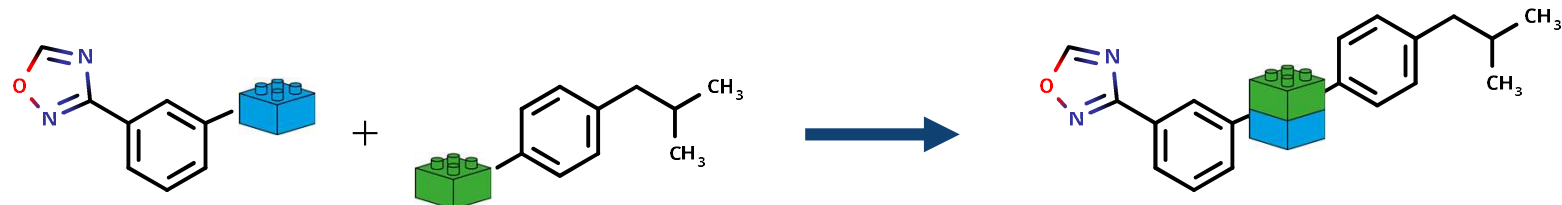
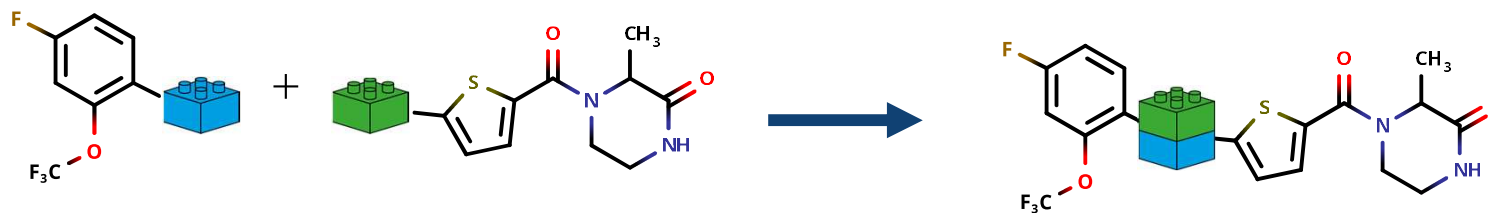
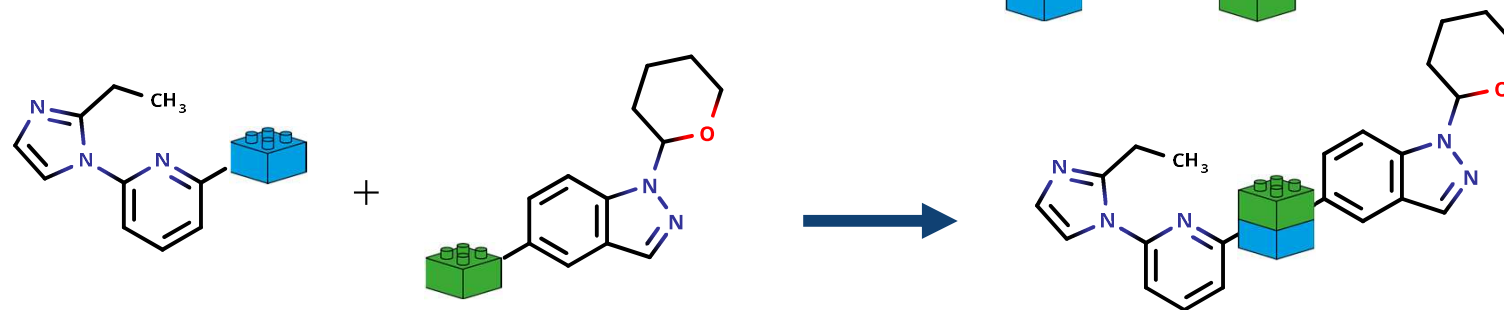
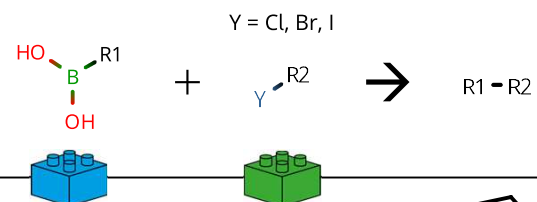
*“If I had asked my customers what they wanted,  
they would have said **faster horses**”*

- Henry Ford



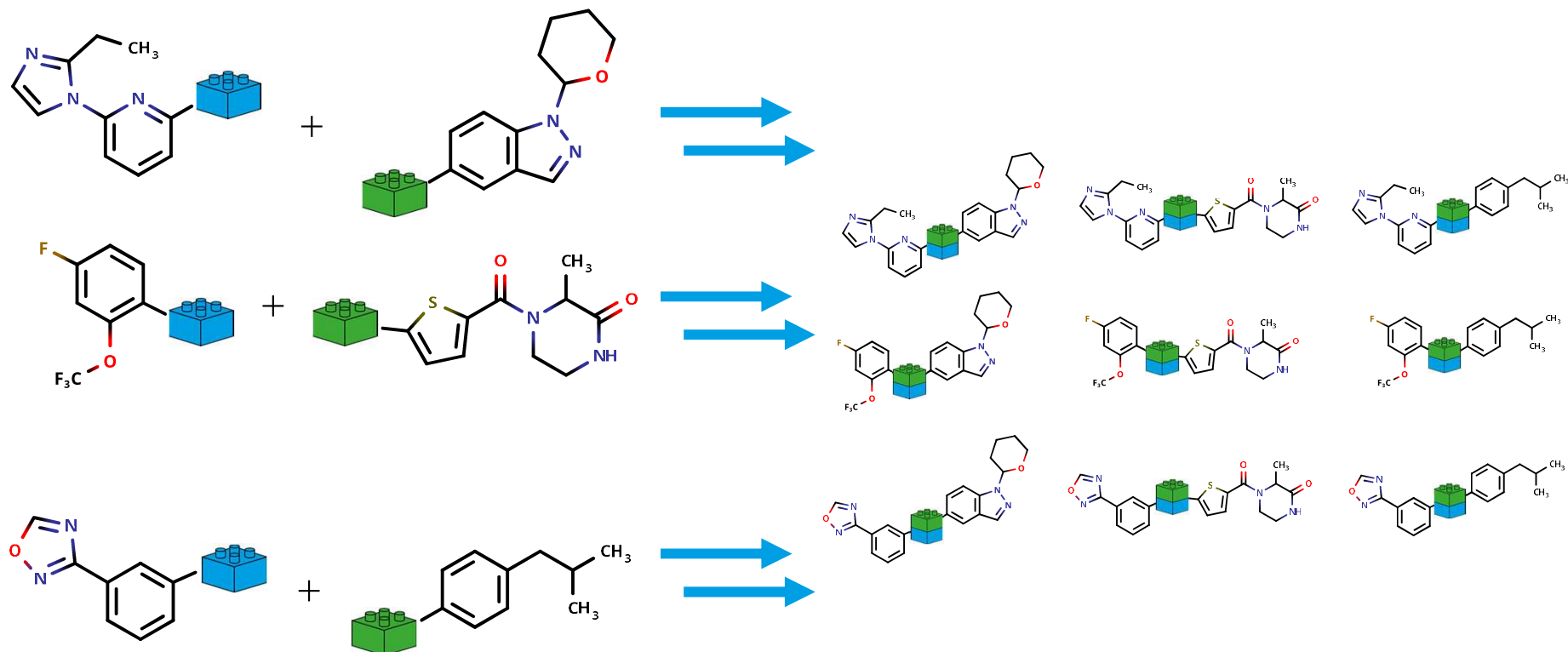
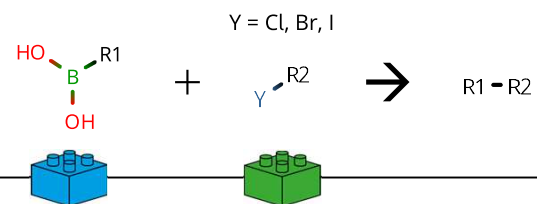
# Our "Idea" of a Car

Reactions, e.g. Suzuki coupling:













# Our "Idea" of a Car

Reactions, e.g. Suzuki coupling:

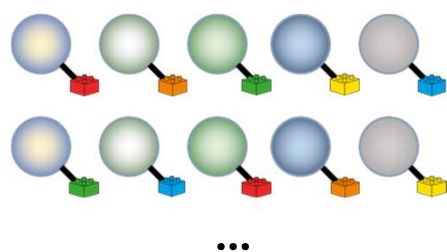


# Multiple Reactions → Chemical Space

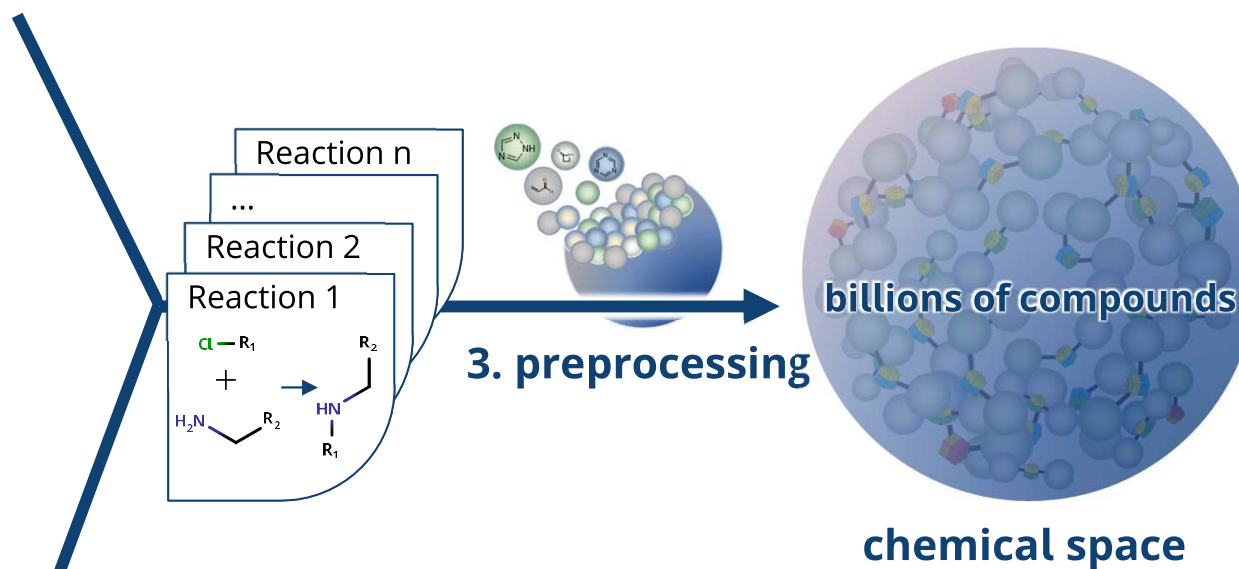
## 1. chemistry know-how

					
		✓	✓		
	✓			✓	
	✓				
		✓			✓
				✓	

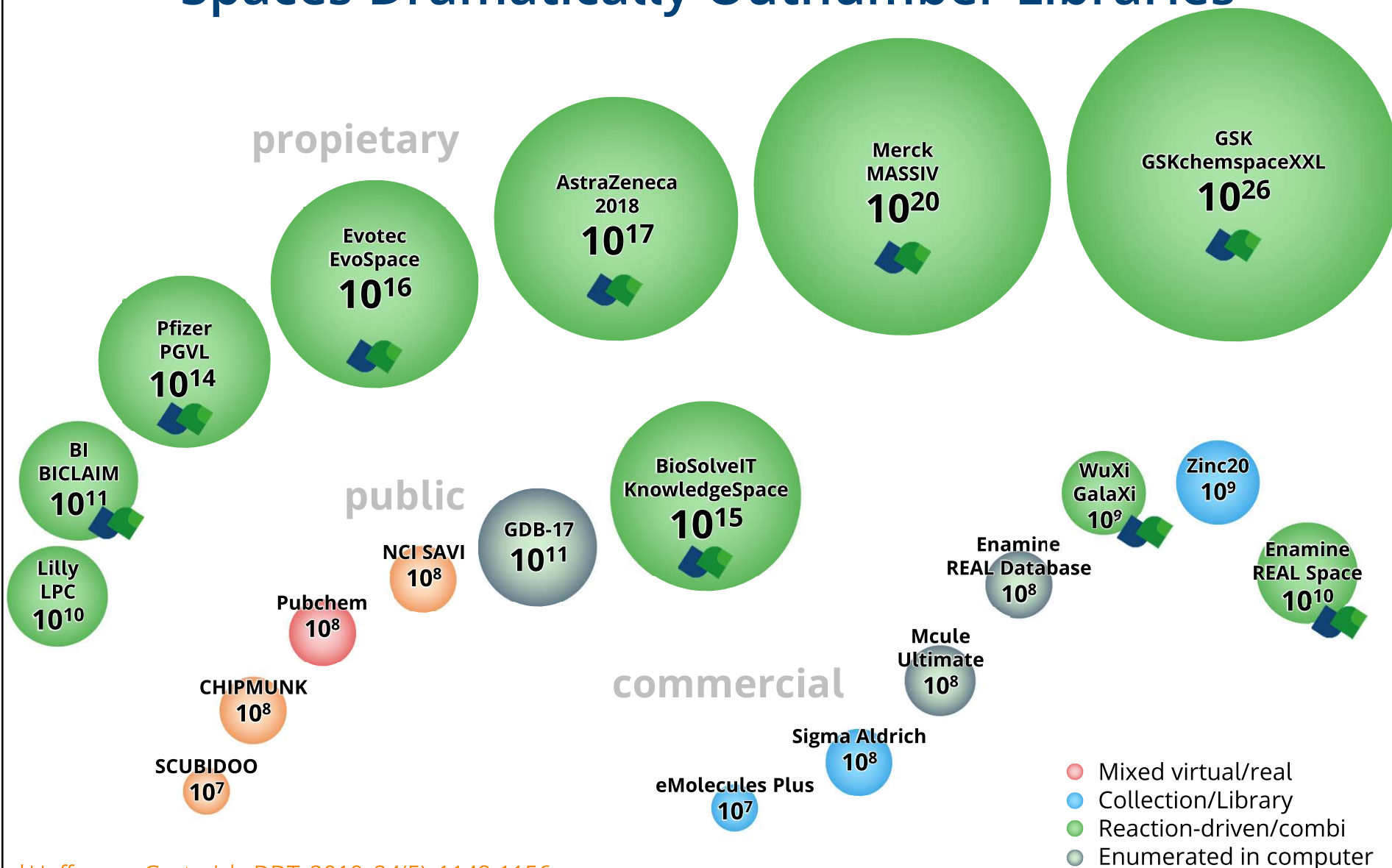
compatibility



## 2. building blocks



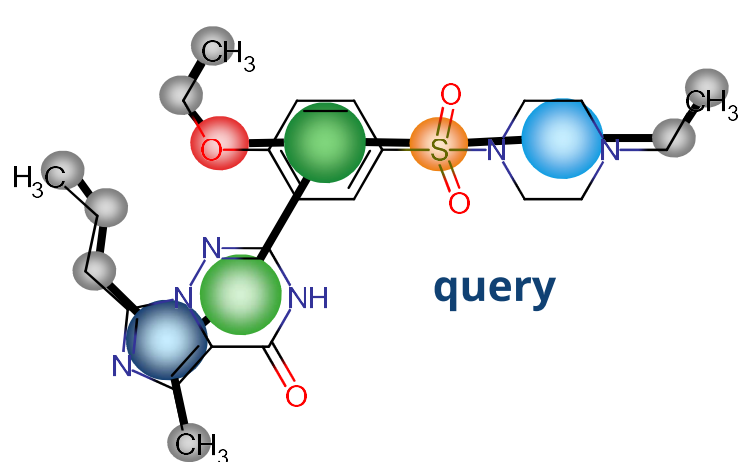
# Spaces Dramatically Outnumber Libraries



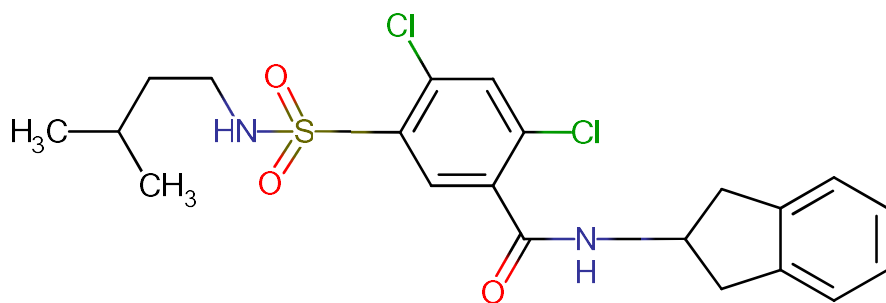
\*Hoffmann, Gastreich, DDT, 2019, 24(5), 1148-1156



# Similarity Searching without Enumeration



Dynamic Programming  
Algorithm



Rarey et al, JCAMD 2001, 15, 497





10<sup>20</sup> Molecules



MERCK

## MASSIV – It works!

- Applied to **12 drug discovery** projects with 1-4 experiments each
- Per experiment: 4-60 compounds with **> 80% feasibility**
- **Higher speed** (2x faster) & **lower costs** (10x cheaper)
- **High IP** by design

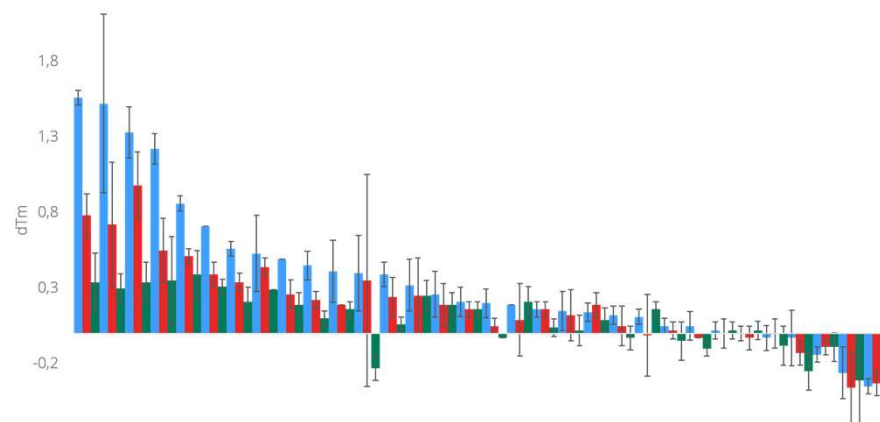
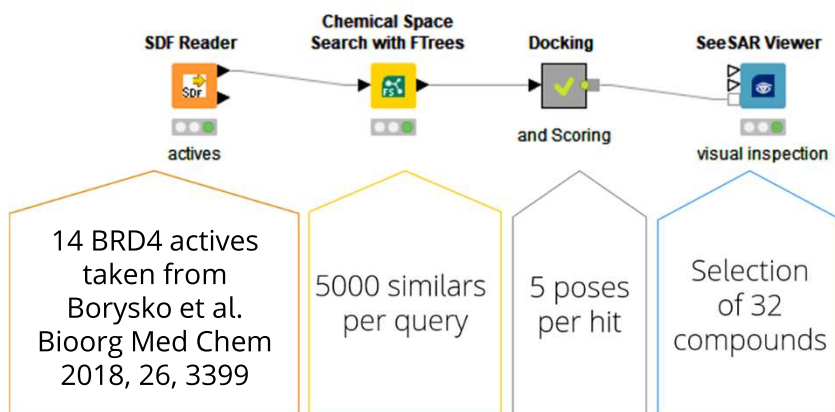
**CURIOUS2018**  
FUTURE INSIGHT CONFERENCE

From Krier & Klingler, curious2018 Darmstadt, Germany

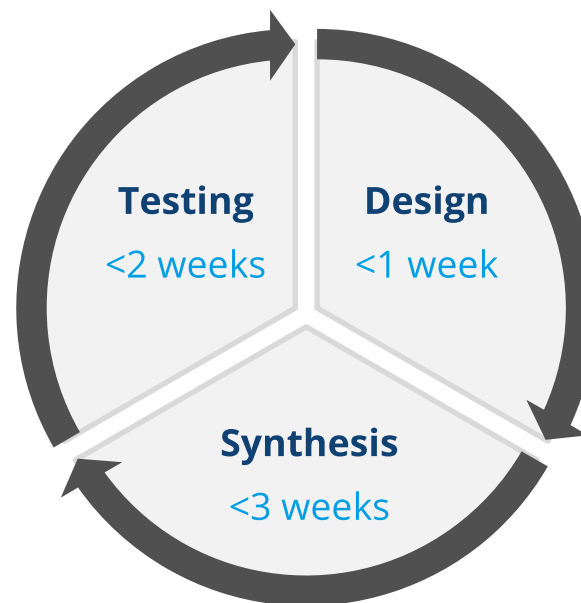


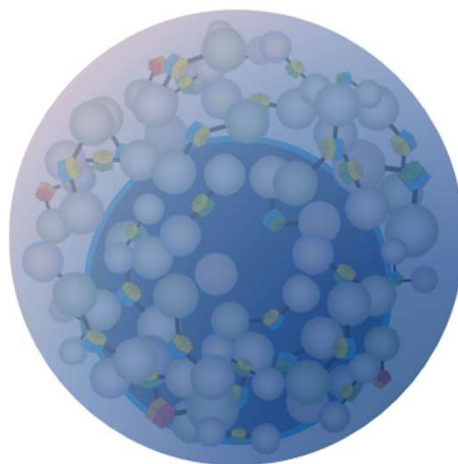
# Similarity Searching + 3D Works

Klingler et al, *Molecules* **2019**, 24(17), 3096



Query	Hit	Tanimoto sim	FTrees sim	$\Delta T_m$ (@40 $\mu$ M)	IC <sub>50</sub> ( $\mu$ M)
		0.456	0.956	0.7	10
		0.277	0.920	1.5	26
		0.323	0.933	1.3	44
		0.333	0.953	1.5	68
		0.356	0.932	0.5	141

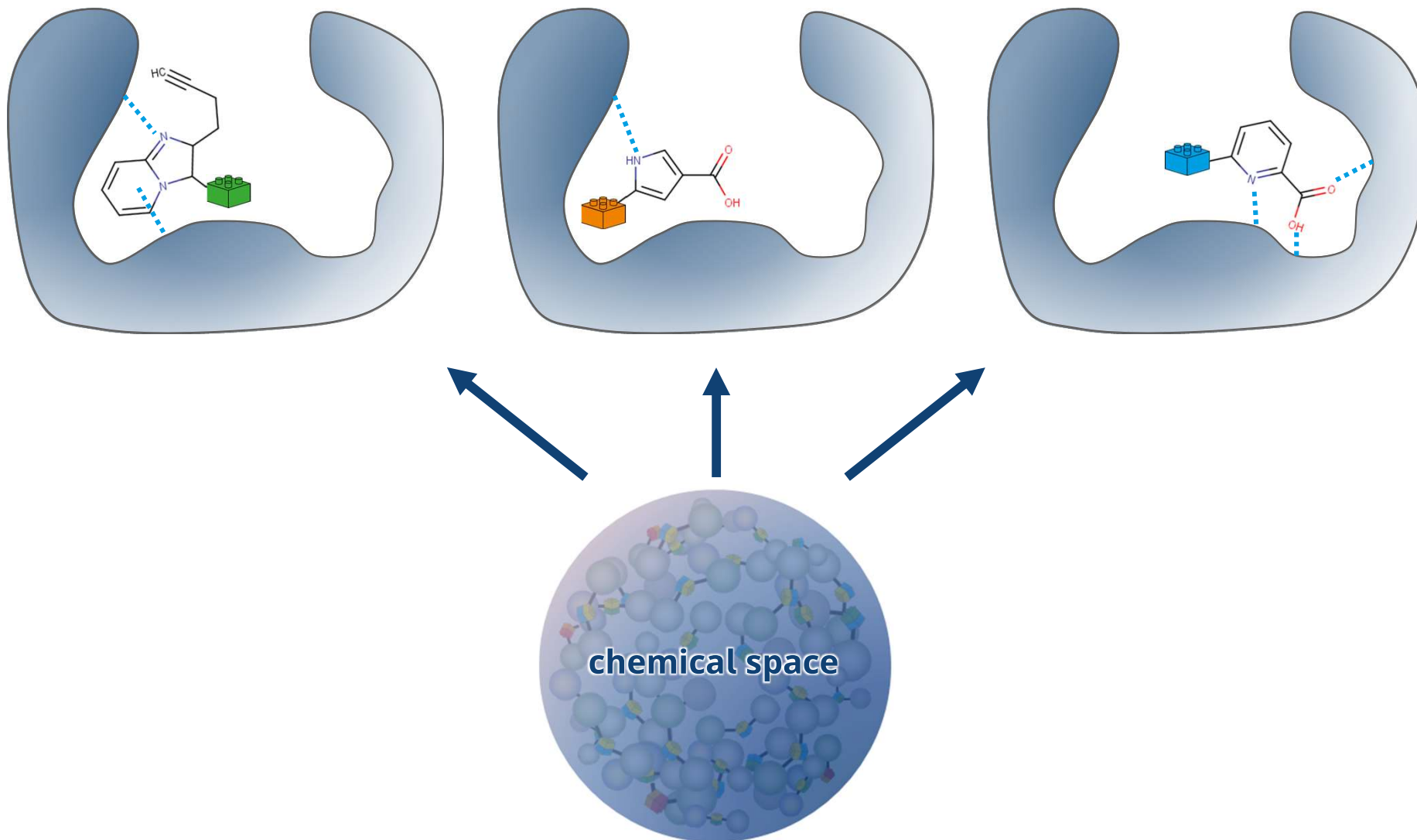




Can we use these vast resources also  
directly in 3D-structure-based design?  
→ Chemical Space Docking

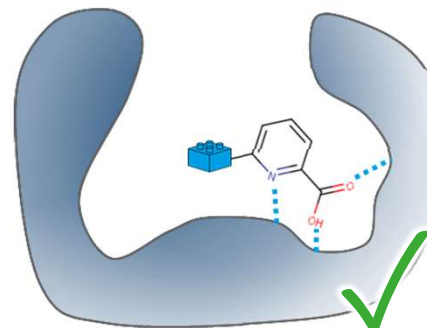
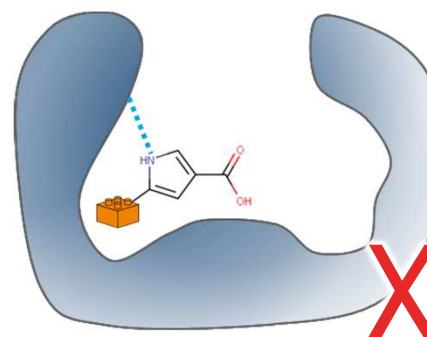
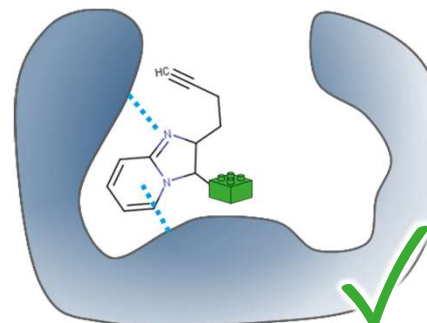


# Building Block Placement

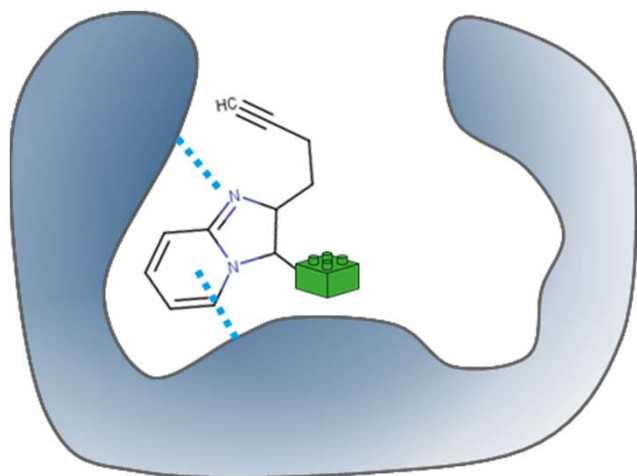












# Building Block Filtering

- Automated
  - unwanted linker positions
  - low scores
  - few interactions
- Manually
  - pharmacophores
  - unspecific binding

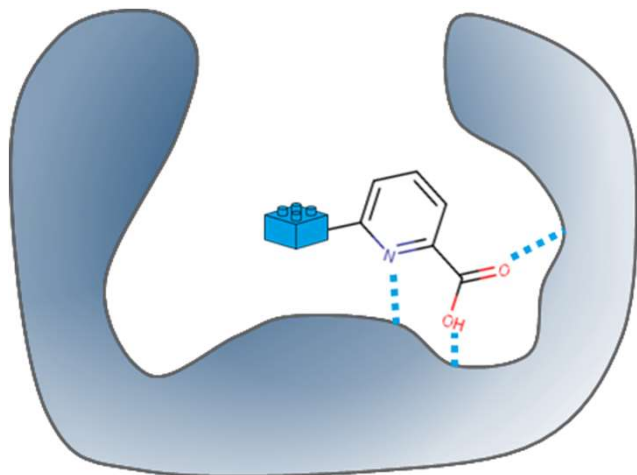


# Combinatorial Expansion



					
		✓	✓		
	✓			✓	
	✓				
		✓			✓
				✓	

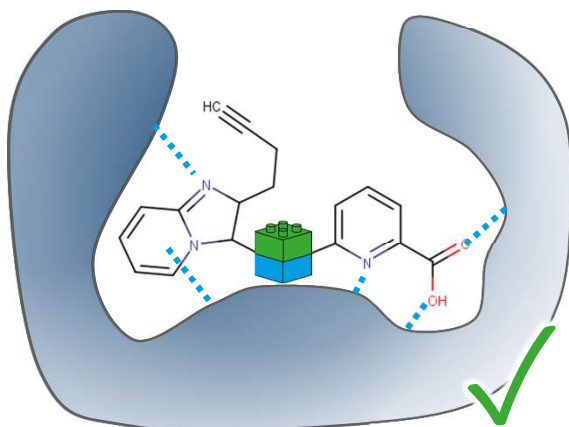
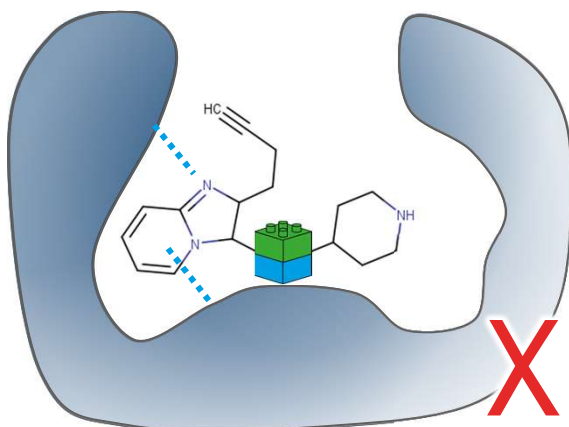
compatibility



- Enumerate libraries with compatible reagents for chosen fragments



## Placement and More Filtering



### Template-based Docking:

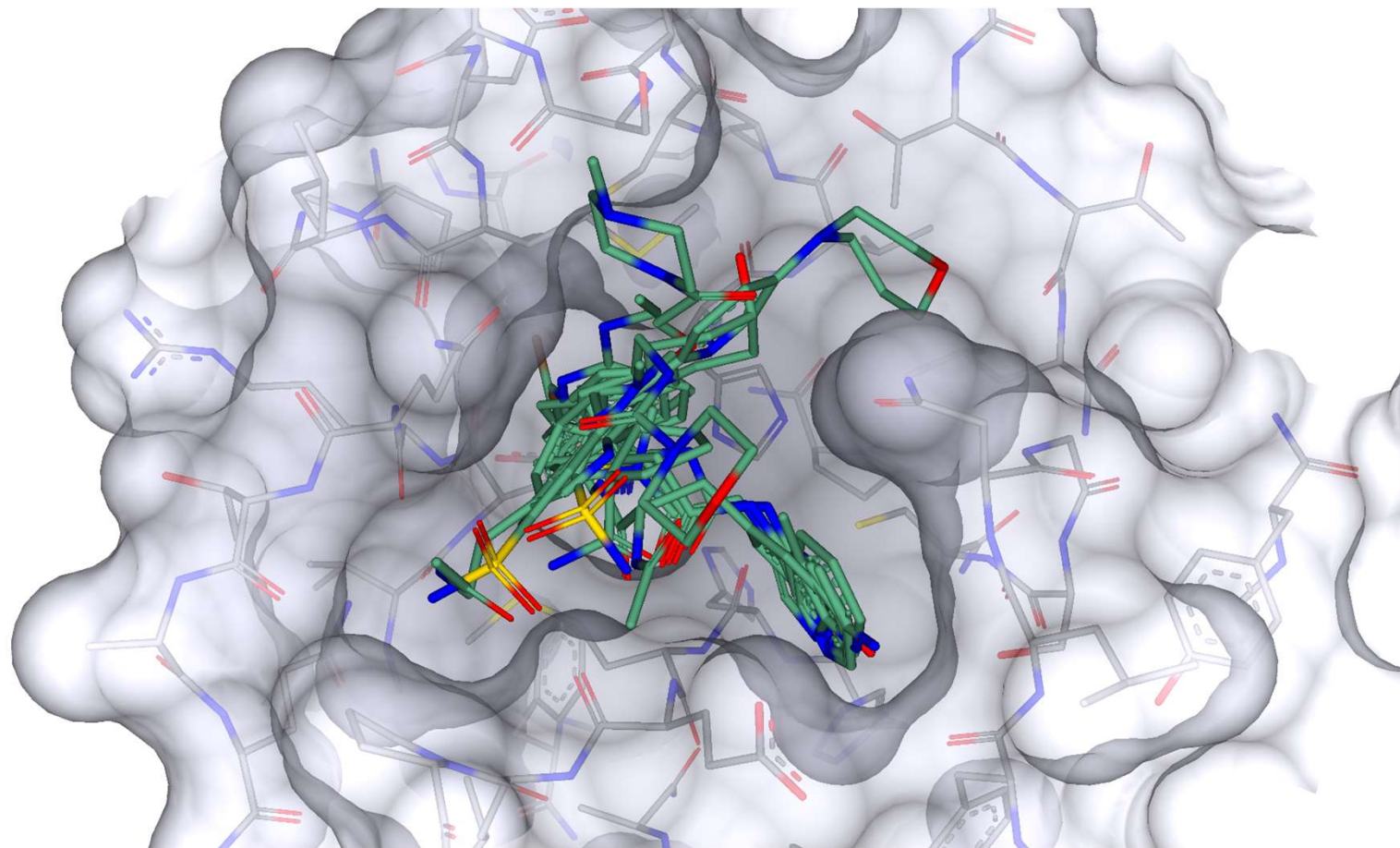
- Fragment from step 1 stays in place
- Added reagent is flexibly attached
- 2<sup>nd</sup> round of filtering and eyeballing to extract the top of the list





# SARS-CoV-2 main protease (M<sup>pro</sup>)

- Starting point: 17 non-covalent crystal structures by X-Chem



<https://covid.postera.ai/covid>





# Chemical Space Docking Workflow

Docking all REAL Space building blocks

Selection of the best 100

Enumeration of 1,731,819 & Docking

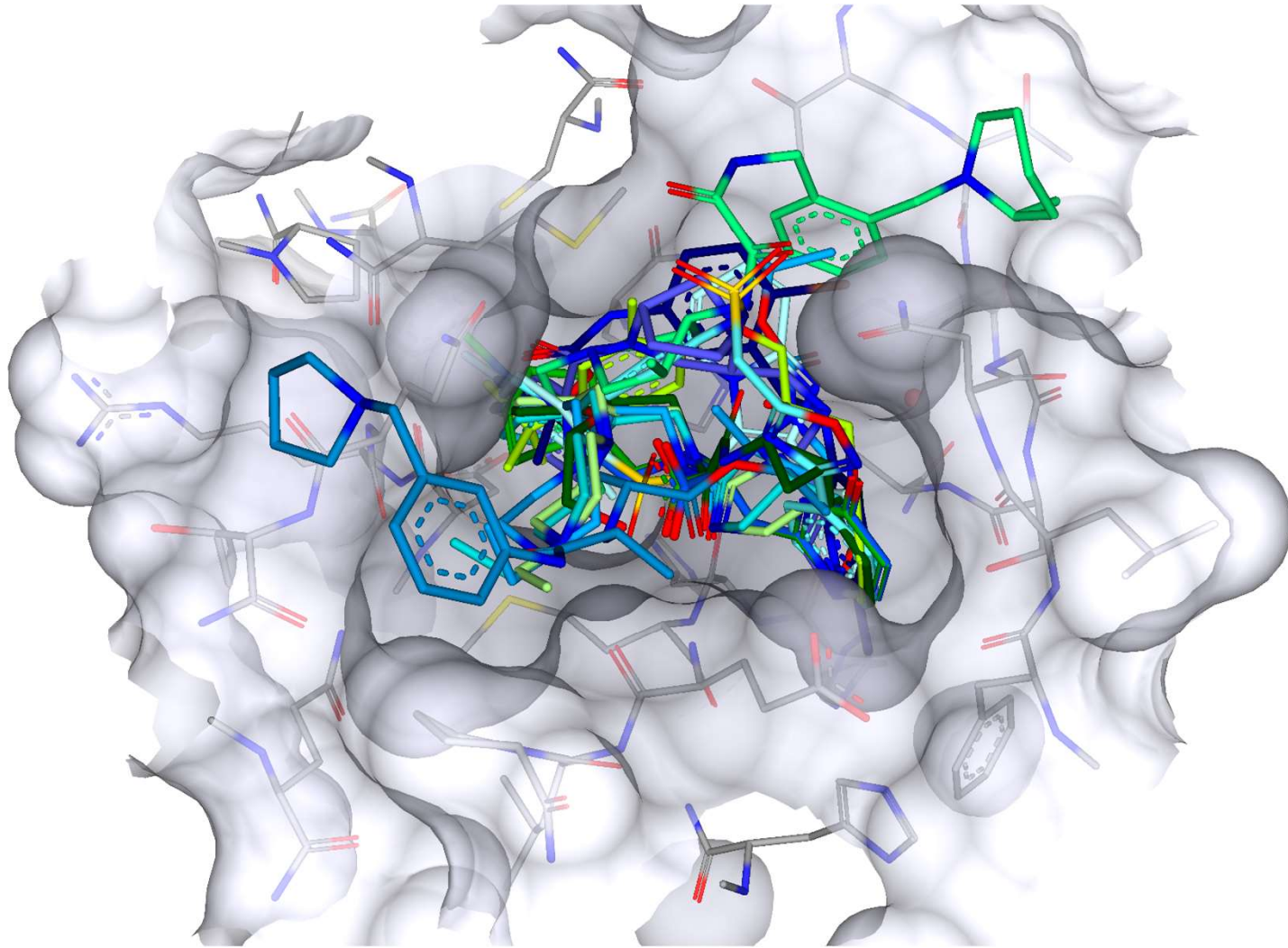
Scoring 2,292,975 poses

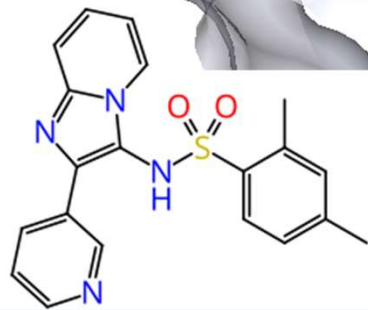
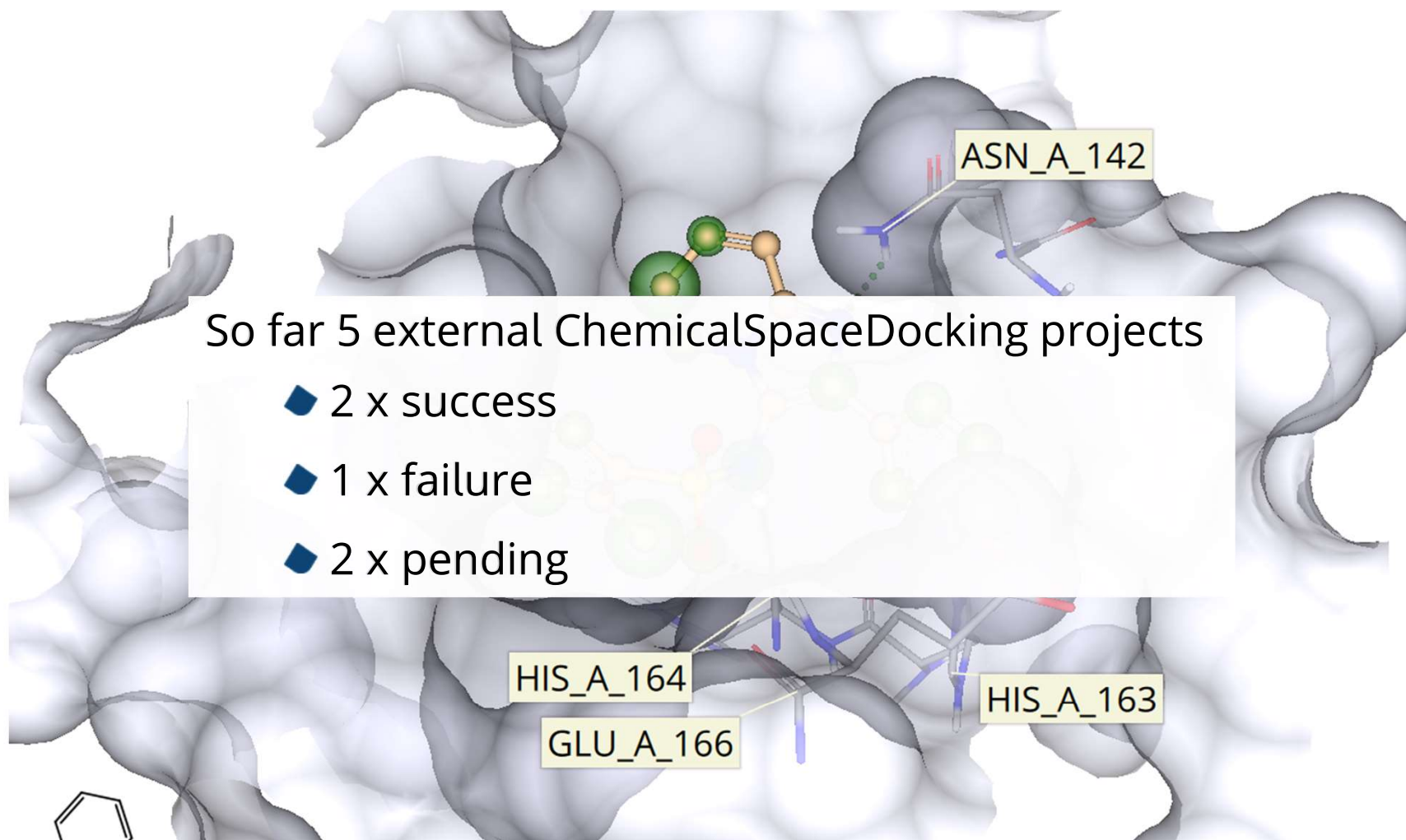
Inspection of the best 50,000

Selection of 13 candidates



# Candidates





s232682\_6950814\_3005078



# Unique Advantages of this Approach

- ◆ Explores billions of compounds via docking
- ◆ Virtual hits become real: e.g. through Enamine or WuXi
- ◆ Saves a lot of time and money



# The Future – What's in it for me?

More & bigger Spaces  
Chemical Similarity, Substructure & 3D Searching



## Interested? – Join the Chemical Space Club!



- ◆ Keep in touch with the key players
- ◆ Discussions & presentations
- ◆ Early announcements
- ◆ [linkedin.com/groups/9004052/](https://www.linkedin.com/groups/9004052/)

