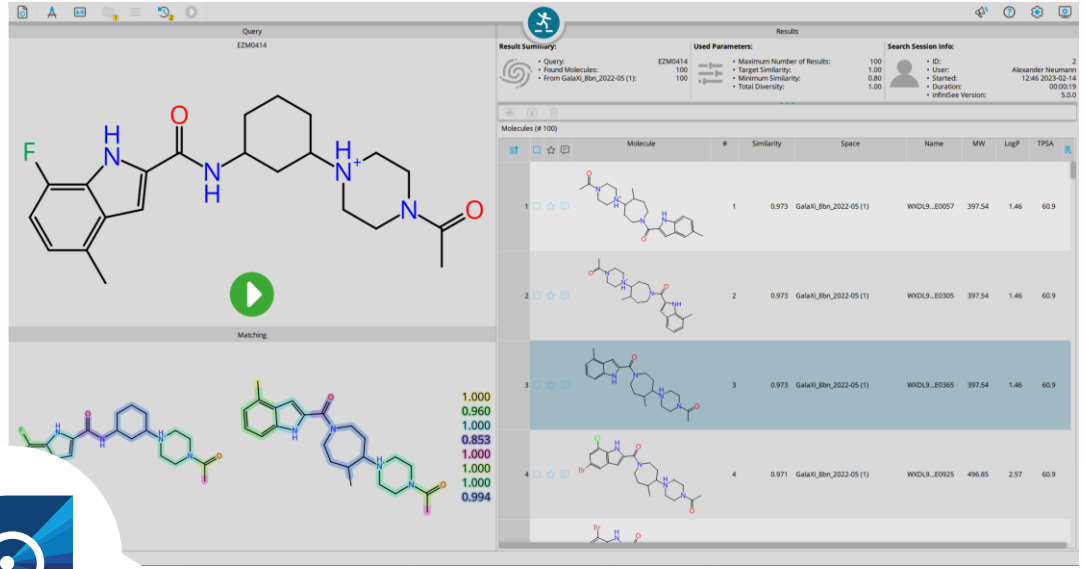
BioSolveIT
expect actives!



infiniSee

Chemical Space Navigation Platform

Screen massive Chemical Spaces

Find accessible compounds

Discover scaffolds for your needs

See how molecules are related

New to Chemical Spaces?

An Overview on the Combinatorial Marvel

Chemical Spaces are your access point to billions or more of easily accessible and tangible compounds. Our partner's Spaces are the perfect source for commercially available results.

Chemical Spaces

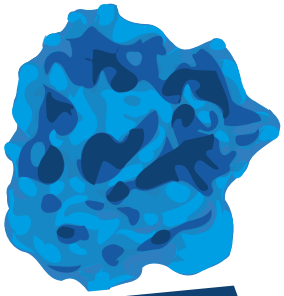
Chemical Spaces are **vast compound catalogs** consisting of building blocks and reactions to combine them.

In comparison to enumerated libraries, they generally contain **more compounds** and can be searched **faster** and **more efficiently** for relevant chemistry.

We at BioSolveIT have dedicated ourselves to develop methods that efficiently mine these Chemical Spaces for relevant compounds.

Search methods

Comparison: Enumerated vs. Combinatorial



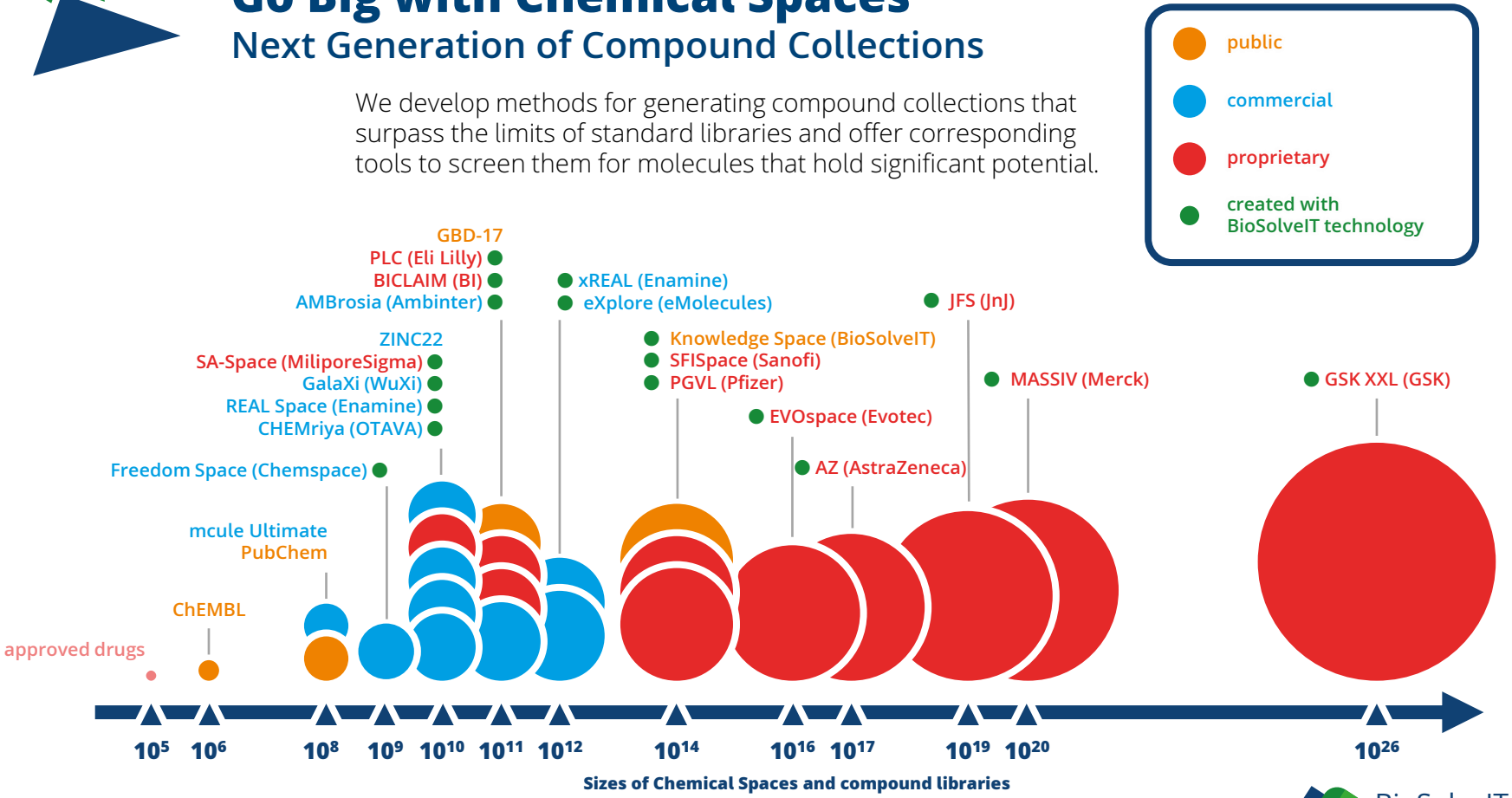
	Compound Library	Chemical Space
Compounds	each entry explicitly listed	Information on building blocks and reactions only
Size	~10^6 (up to 10^{10})	starting at 10^9
Data	TB-PB	MB-GB
Type	enumerated	combinatorial
Processing time	hours-weeks	seconds to minutes*
File type	.sdf, .mol2, .SMILES	.space
Generation	curated/mixed	building blocks + reactions
Results	accessible (?)	accessible!

An extensive list of all contained compounds does not exist.
Compounds are generated on the fly during a search.

Go Big with Chemical Spaces

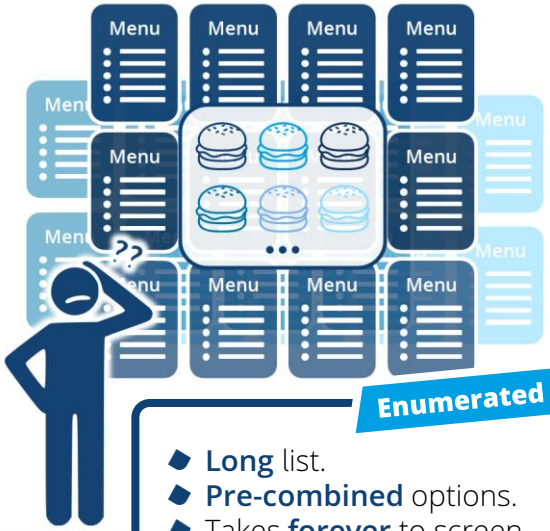
Next Generation of Compound Collections

We develop methods for generating compound collections that surpass the limits of standard libraries and offer corresponding tools to screen them for molecules that hold significant potential.



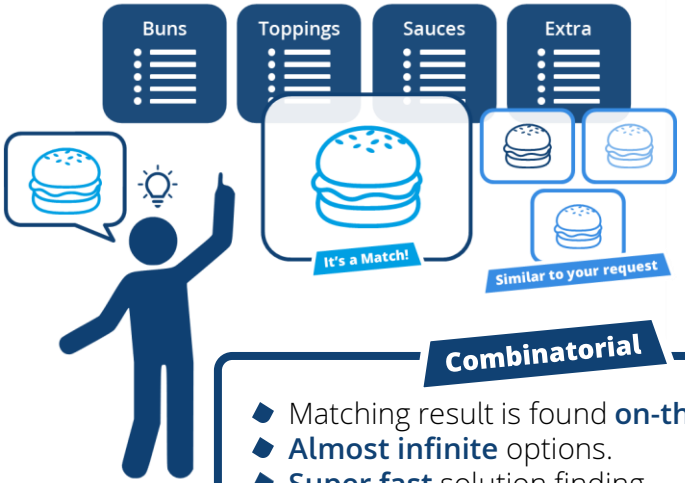
Explained: Enumerated vs. Combinatorial

An Approach to Handle Large Numbers



Enumerated

- **Long** list.
- **Pre-combined** options.
- Takes **forever** to screen.



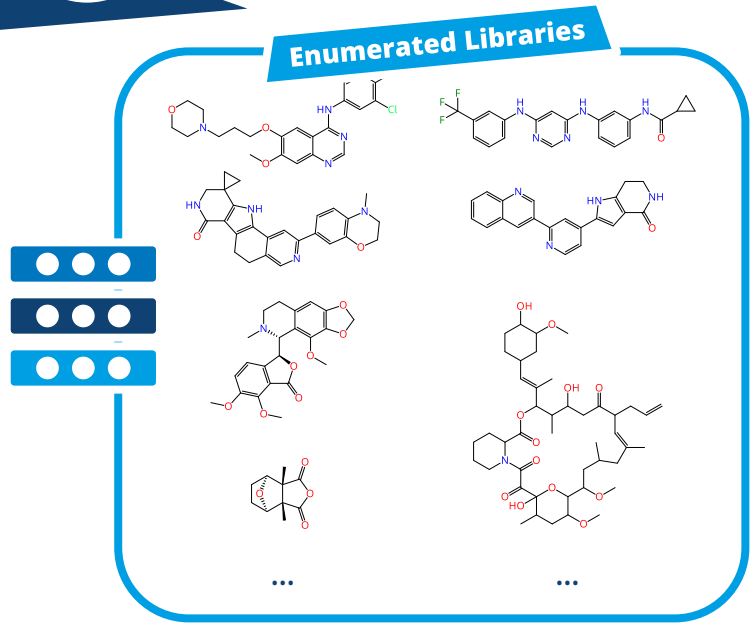
Combinatorial

- Matching result is found **on-the-fly**.
- **Almost infinite** options.
- **Super fast** solution finding.

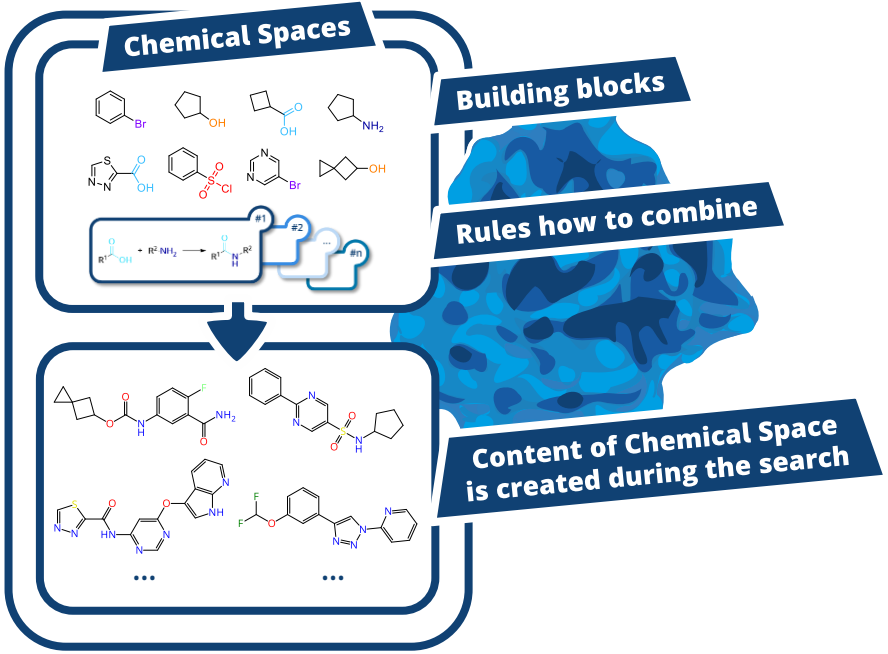
- Likely **not all possible** combinations of ingredients are included.
- All options must be reviewed to come to a conclusion, even the ones **you're not interested in**.
- **Rigid system** that may not address your preferences.

- Features **ingredients to all possible combinations**.
- Accelerated browsing: Combinations that **are not similar enough** to your request are not explored.
- **All results can be found**, not a single possibility is discarded.

Content of Libraries and Chemical Spaces



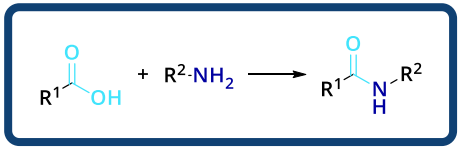
- Each entry is explicitly listed.
- No synthetic information involved.
- Content can come from different sources: in-house synthesized compounds, natural products from extracts, computationally generated molecules, ...



- Content is generated on-the-fly based on the task.
- Chemistry-aware. Results can be synthesized.
- Set up allows faster screening and more coverage.

How to “Combinatorial Chemistry”

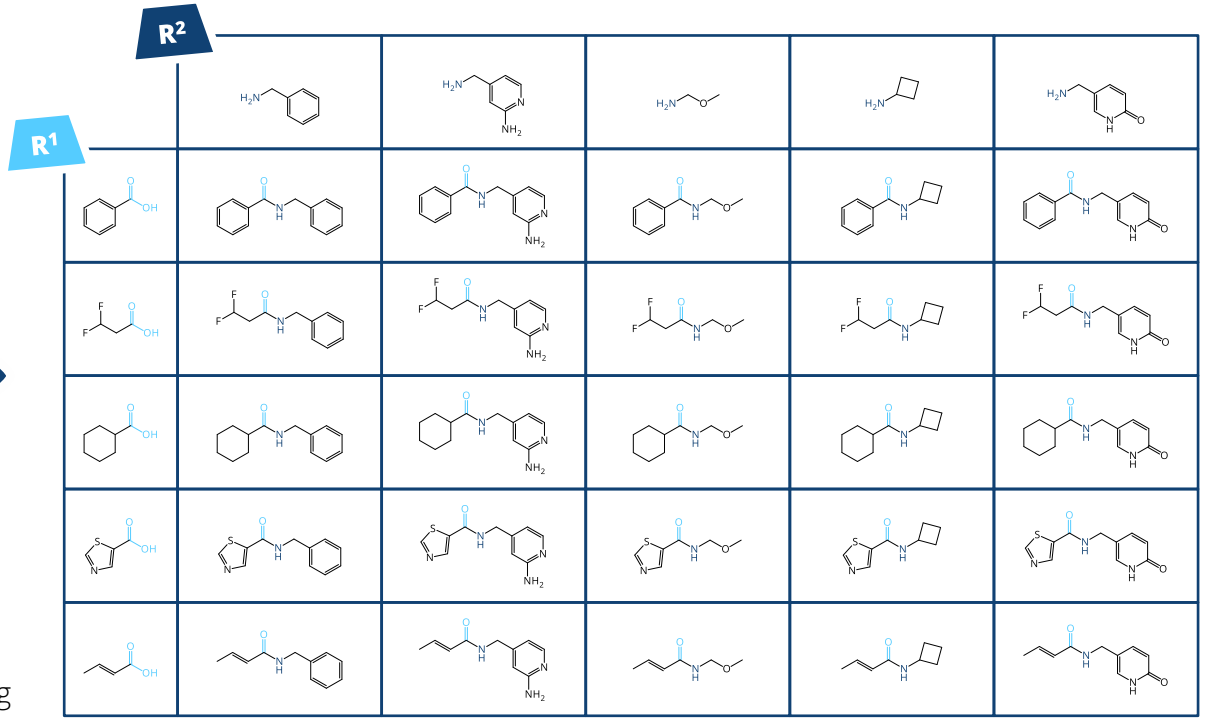
Reaction
e.g., amide coupling



- Define chemical reaction.
- Add and combine building blocks that match the definition.

- 5 × 5 building blocks = 25 combinations
- Combinatorial explosion: The more building blocks, the larger the numbers

1,000 × 1,000 building blocks = 10^6 compounds

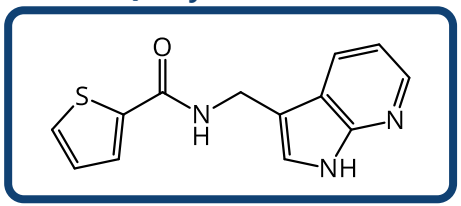


Accelerating the Screening

Utilizing Combinatorial Chemistry

Drug Discovery challenge: Find similar compounds

Query molecule



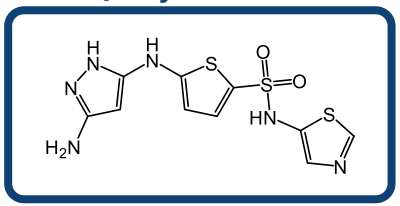
- Best matching results **always** emerge with highest scores.
- The **similarity of the entire molecule** is taken into account, not just that of its individual components.
- Less relevant compounds are displayed only when more results are requested. **Nothing is missed**.

On-the-Fly Generation of Results

Efficient Way to Handling Large Numbers

Drug Discovery challenge: Find similar compounds

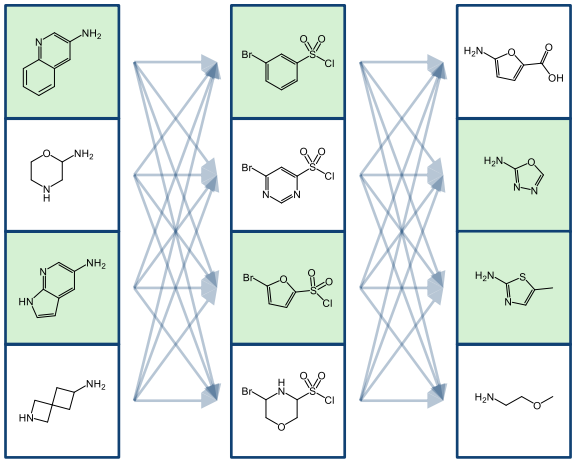
Query molecule



- Several search methods available to mine for similar compounds.
- Multi-component reactions are supported.

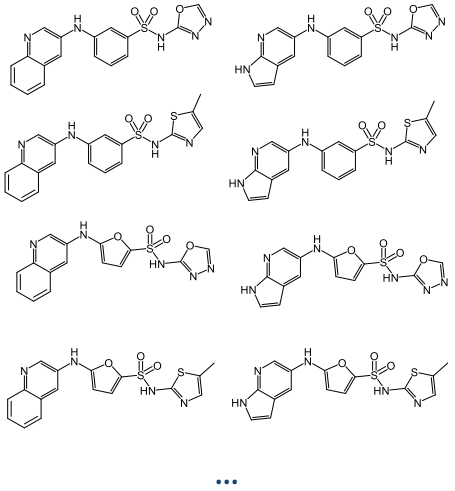
- A search run generates a requested number of the best scoring results out of a × b × combinations.

Combinatorial build up of molecules



... a + ... b + ... c

Results (enumerated during search)



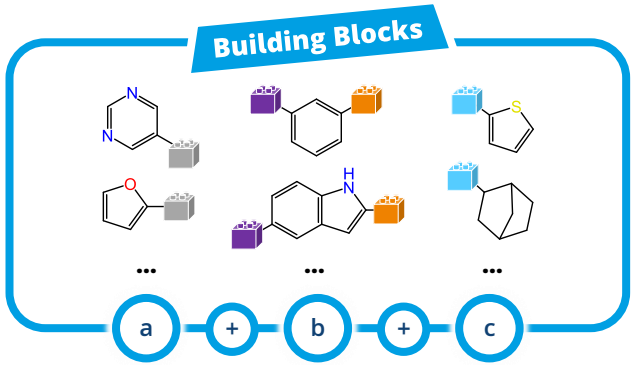
- Top scoring results matching the query are enumerated.

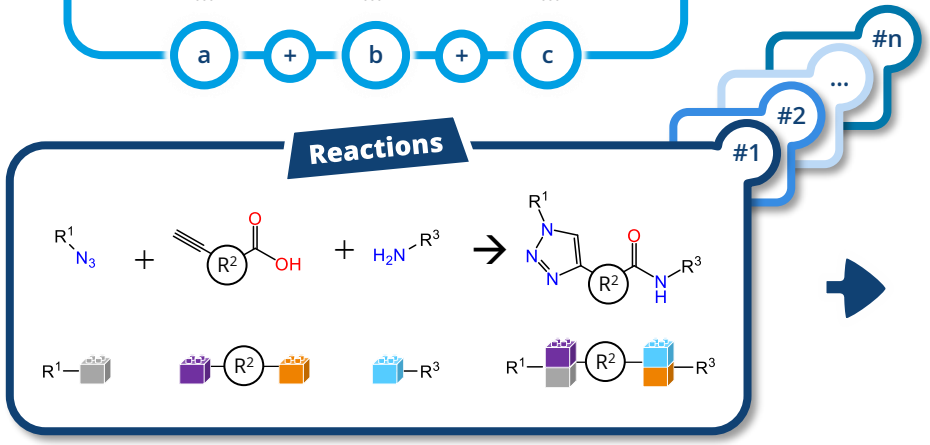
(Requested number of a × b × c combinations.)

- Efficient enumeration puts **focus on what is important** = faster computation
- While ALL possibilities are assessed, no computational power is wasted on low scoring candidates.

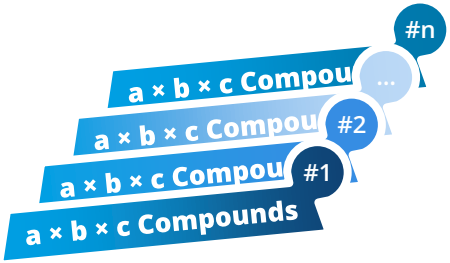
Building Blocks and Reactions

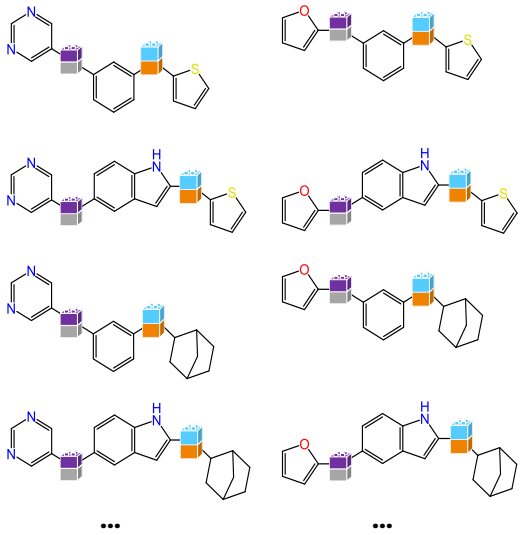
Ingredients to a Chemical Space





e.g., click chemistry and amide coupling



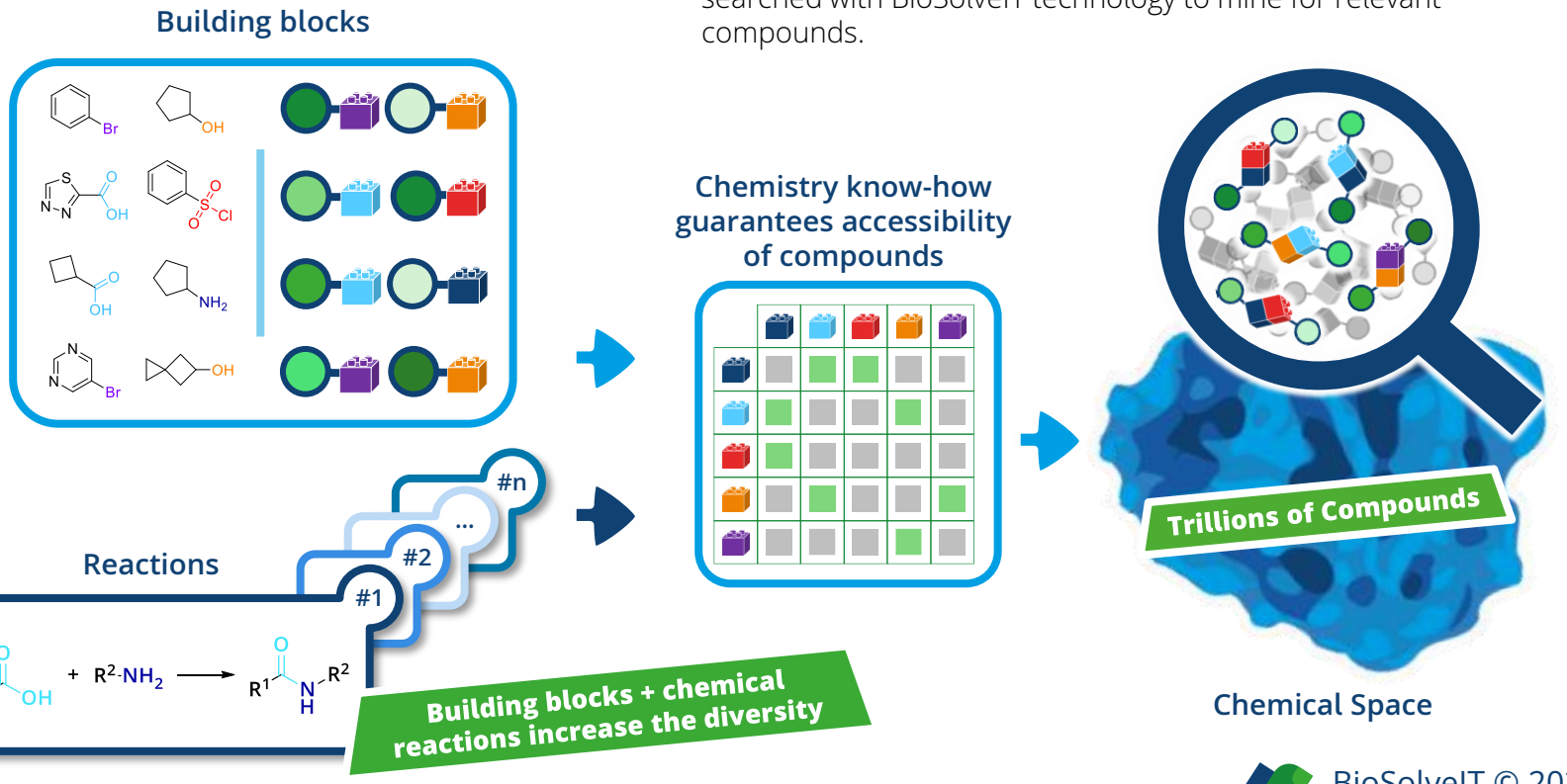


Combinatoric explosion of possible compounds!

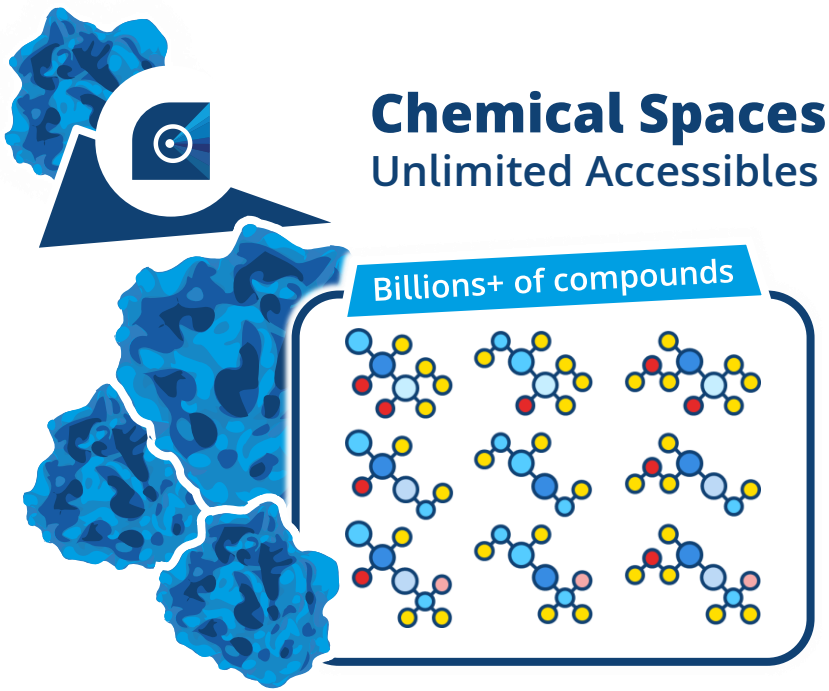
CoLibri

Building Chemical Spaces

CoLibri is a stand-alone command-line tool used to create ultra-vast Chemical Spaces using building blocks and reaction rules on how to combine them.
The generated combinatorial Chemical Spaces can then be searched with BioSolveIT technology to mine for relevant compounds.



Chemical Spaces
Unlimited Accessibles



Vast, tangible compound catalogs
We collaborate with compound suppliers to create unique Chemical Spaces reflecting their synthetic capabilities and potential.
Compounds of interest can be purchased or synthesized in-house.

Make-on-demand:
- 1.1×10^{11} AMBrosia (Ambinter)
- 7.0×10^{10} REAL Space (Enamine)
- 1.2×10^{10} GalaXi (WuXi AppTec)
- 1.2×10^{10} CHEMriya (OTAVA)
- 5.1×10^{9} Freedom Space (Chemspace)

Do-it-yourself*:
- 5.0×10^{12} eXplore (eMolecules)

* Vendors are delighted to collaborate with a CRO of your choosing.

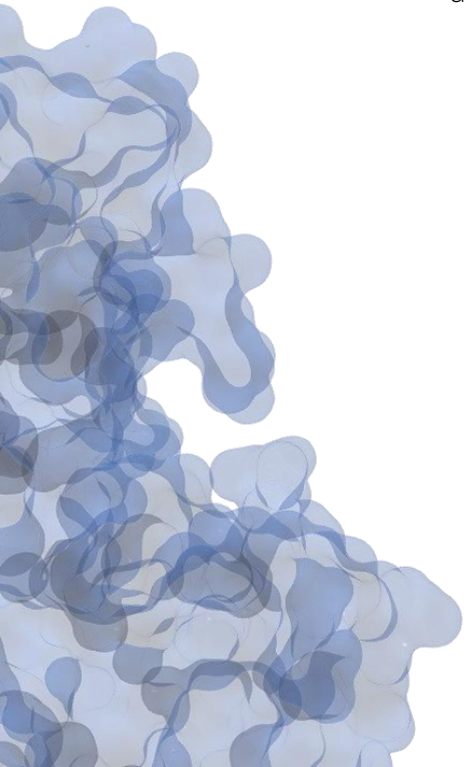
Ideation:
- 2.9×10^{14} Knowledge Space (BioSolveIT)

Exclusive to infiniSee xREAL
- 2.4×10^{12} xREAL Space (Enamine)

Why We Need Large Chemical Spaces

How to Profit from Massive Molecule Collections

Relying on brute force to manage billions of compounds is both inefficient and costly. Chemical Spaces offer a more advanced and cost-effective approach to find tangible and attractive molecules.



Bigger is better. Combinatorial Chemical Spaces are over 1000-times larger than enumerated libraries. The largest hunting grounds out there!

The retrieved results are **accessible** per design. Chemistry rules guarantee synthesizable and purchasable results.

Chemical Spaces can be screened on your **own hardware** without the need of a supercomputer or a server cluster.

Boost the efficiency of your internal resources with your **own Space**. Your building blocks and reactions create the perfect setting for finding top molecule candidates.

Access to our partners' Chemical Spaces unlocks **diverse** chemical options, increasing the chances of finding the ideal candidate.

Multiple methods have been developed to extract chemistry from Spaces, with settings that can be customized to meet your needs.

Empowering Drug Discovery

Tools to Mine From Chemical Spaces

Handling the immense task of browsing through billions of compounds calls for new methods and approaches to complete the process in a timely manner.

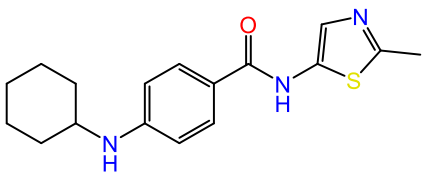
At BioSolveIT we develop novel technologies to extract the most relevant compounds for modern drug discovery challenges based on similarity to compounds of interest.

All search methods are **also available as command-line versions** that can easily be implemented into your workflows.

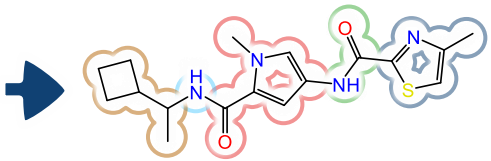
Possibilities of Chemical Space Exploration

Find Molecules Guided by the Needs

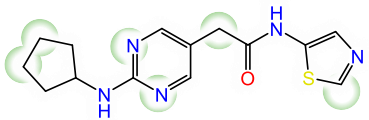
infiniSee features distinct modes, each aimed at identifying compounds according to the targeted drug discovery challenge.



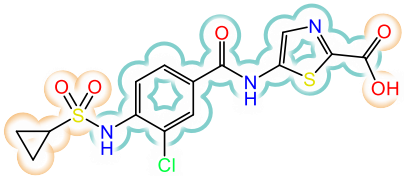
Query Molecule



Scaffold Hopper
Fuzzy matching of molecular features
(similar arrangement of functional groups)



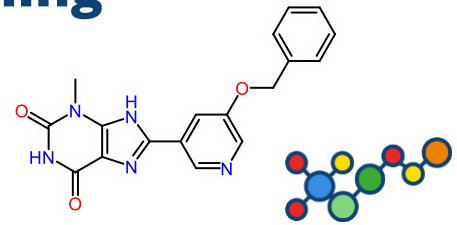
Analog Hunter
Close analogs with high fingerprint similarity
(variances like repeated patterns)



Motif Matcher
Substructure matching
(maximum overlap of a continuous arrangement)

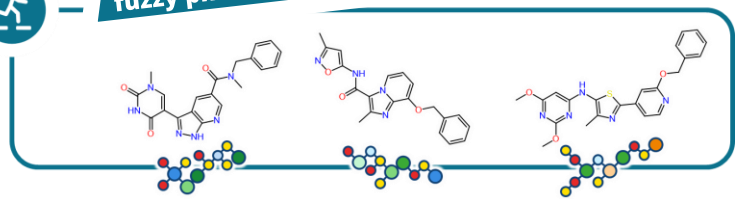
Trinity of Compound Mining

What Results to Expect



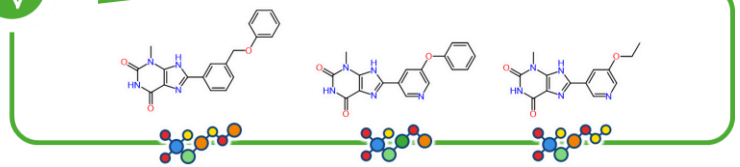
Query compound
Every Chemical Space exploration starts with a compound of interest.

fuzzy pharmacophore



scaffold hopping • novel IP

close analogs

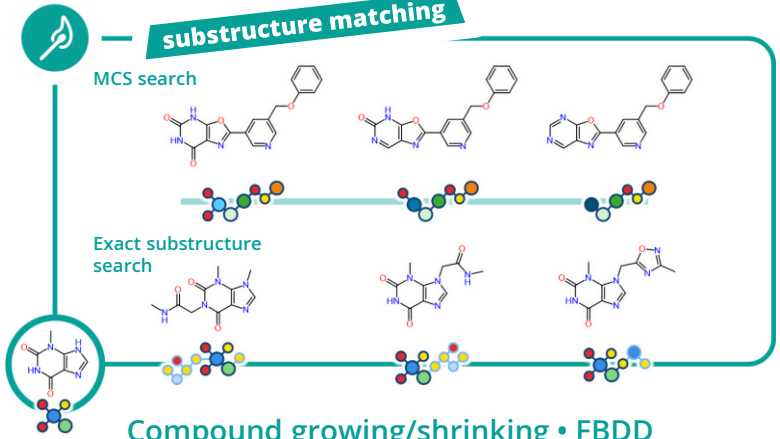


SAR analysis

substructure matching

MCS search

Exact substructure search



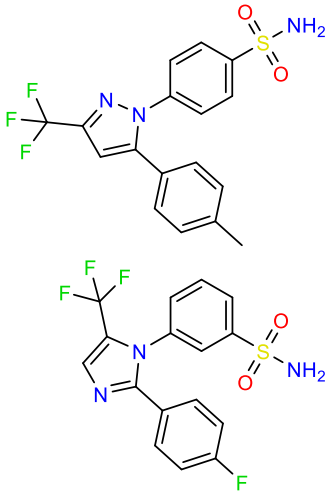
Compound growing/shrinking • FBDD

Scaffold Hopper

Fuzzy Pharmacophore Search

Search for distant neighbors of a query molecule with the Scaffold Hopper to spot non-obvious similarities. Retrieved compounds will surprise you with novel structures matching the flavor of your molecule.

The molecular feature descriptor captures a molecule's overall topology and its pharmacophore properties.

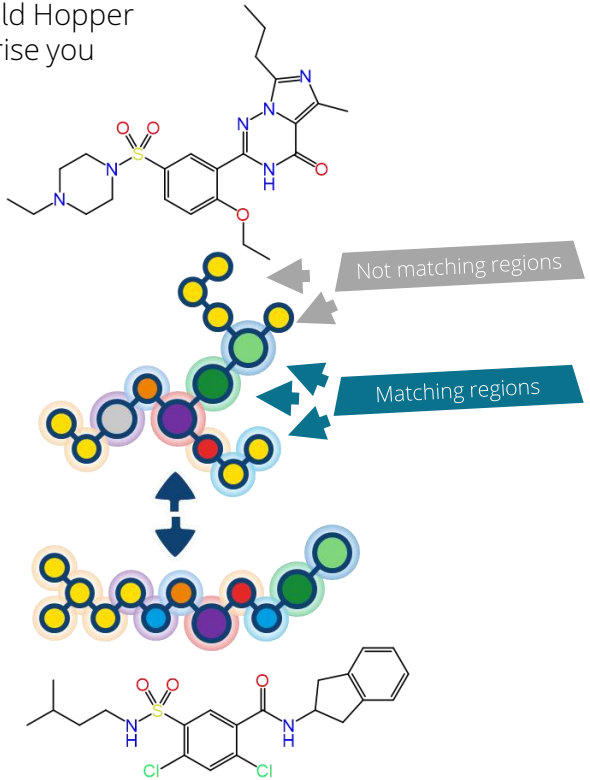


Tanimoto similarity: 0.328
FTrees similarity: 0.989

Compare apples and oranges

The underlying **FTrees algorithm** finds related and relevant molecules which would be missed by other methods. As an **orthogonal method** to fingerprint-based screenings, it overcomes their weaknesses to generate relevant results.

The similarity of two compounds is calculated based on the alignment of the molecular features.



Analog Hunter

Molecular Fingerprint Similarity

The Analog Hunter retrieves compounds of interest based on molecular fingerprint similarity from vast Chemical Spaces with the **SpaceLight algorithm**.

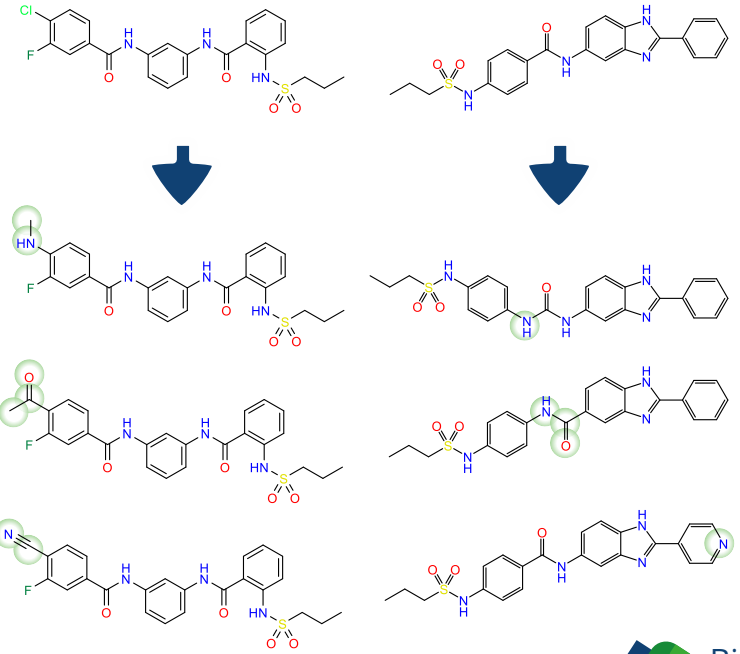
Proximal chemical space

Compounds retrieved with the Analog Hunter contain smaller modifications inside the molecule. Therefore, they are perfect candidates to explore structure-activity relationships around a molecule of interest.

Discover compounds for any scenario

While infiniSee applies the well-known ECFP4 fingerprint, novel CSFP (Connected Substructure Fingerprint) methods and their variants can be used in the command-line version of SpaceLight.

Query compounds



Fingerprints for Chemical Spaces

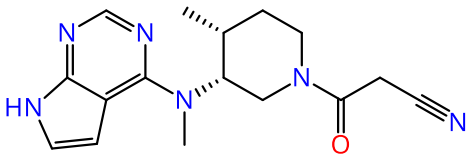
Connected Substructure Fingerprints

Nice to know

CSFPs are developed to mine relevant chemistry from combinatorial Chemical Spaces.

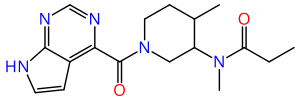
Similar ≠ Similar

The sense of similarity may change based on the aim of the project. Connected substructure fingerprints (CSFPs) address this research perspective.

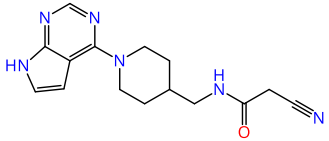


A SpaceLight search: Tofacitinib

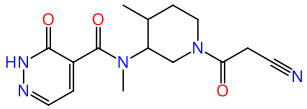
On the right side, results examples of CSFPs screens are presented with were retrieved with the command-line tool. Each of them ranks best for the respective fingerprint and may be of interest for different campaigns.



Closest neighbor



Focus on functional groups



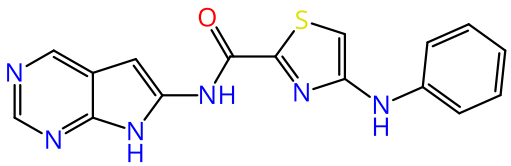
Similar arrangement of functionalities

fCSFP	**1**	**0.720**
iCSFP	4,093	0.726
tCSFP	25	0.743
fCSFP	328,523	0.472
iCSFP	**1**	**0.913**
tCSFP	581,923	0.523
fCSFP	4,666	0.620
iCSFP	81,174	0.656
tCSFP	**1**	**0.794**

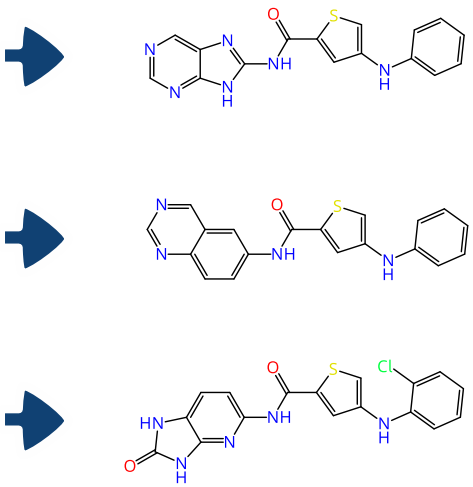
Motif Matcher

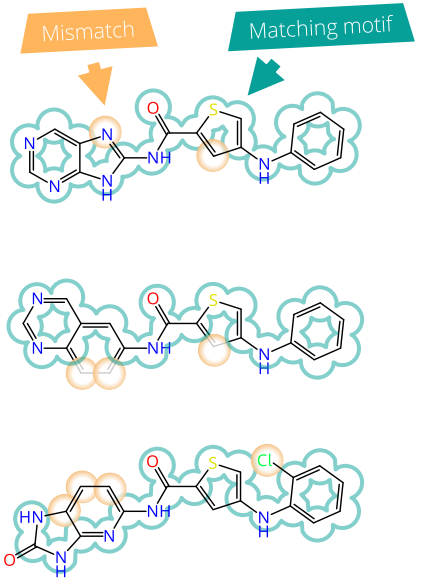
Substructure Search

Find compounds that contain the specific motifs you are seeking from extremely vast Chemical Spaces. The **SpaceMACS algorithm** advances FBDD and substructure-driven campaigns to the next level.



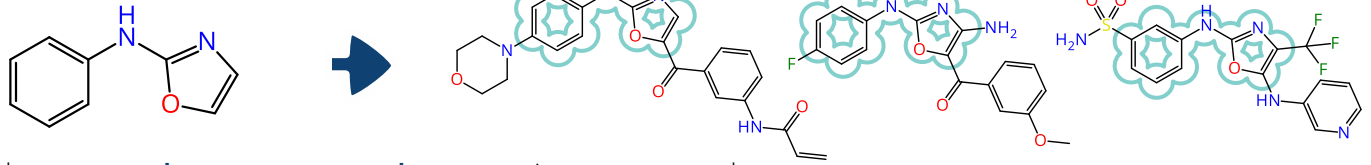
More than an exact substructure search
Using the **maximum common substructure (MCS) search** in the Motif Matcher provides you with molecules that have the highest similarity based on the longest identical sequence of heavy atoms of your query compound.





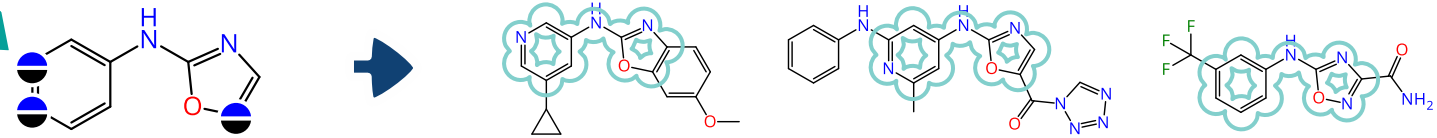
Versatile Claw Machine

More Than Just Straight Forward



Using the **exact substructure search**, you retrieve compounds containing a particular motif of interest.
You can also set the number of additional heavy atom to match your needs and screen for drug- and lead-like structures.

Nitrogen OR Carbon



The command-line version of SpaceMACS supports **SMARTS** definitions to direct your search and obtain molecules that are most relevant to your projects by defining desired chemical pattern and functionalities.

Augmented Fragment-Based Drug Discovery

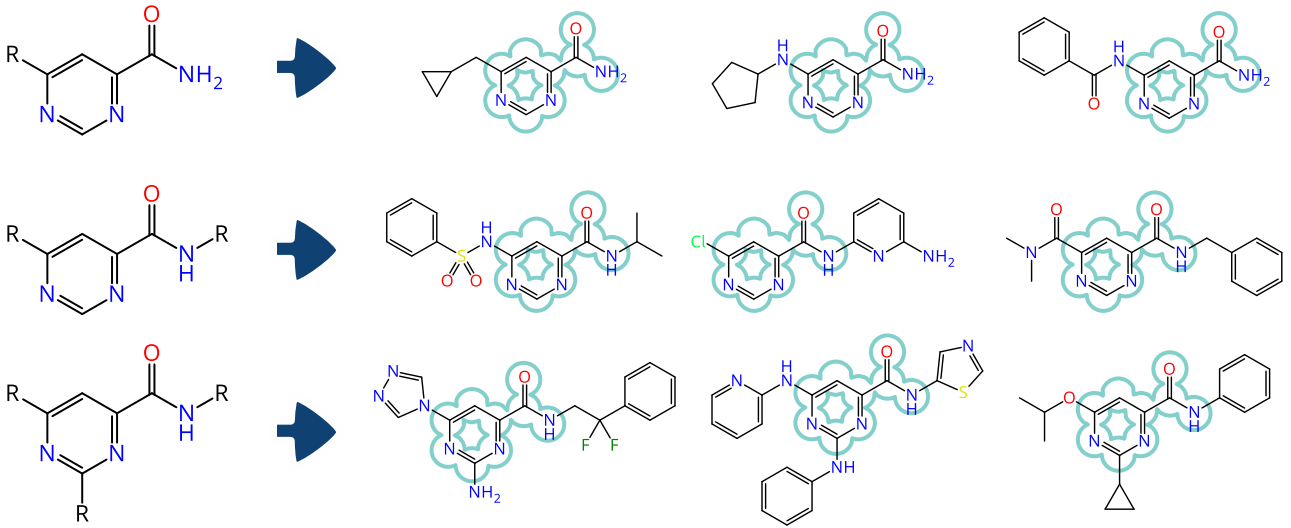
Extracting Millions of Extensions

Chemical Spaces are the largest catalogs for commercially available compounds. While common enumerated catalogs contain thousands of compounds for prominent substructures (e.g., hinge binders), Spaces go several orders of magnitude higher into millions and more!

R group search

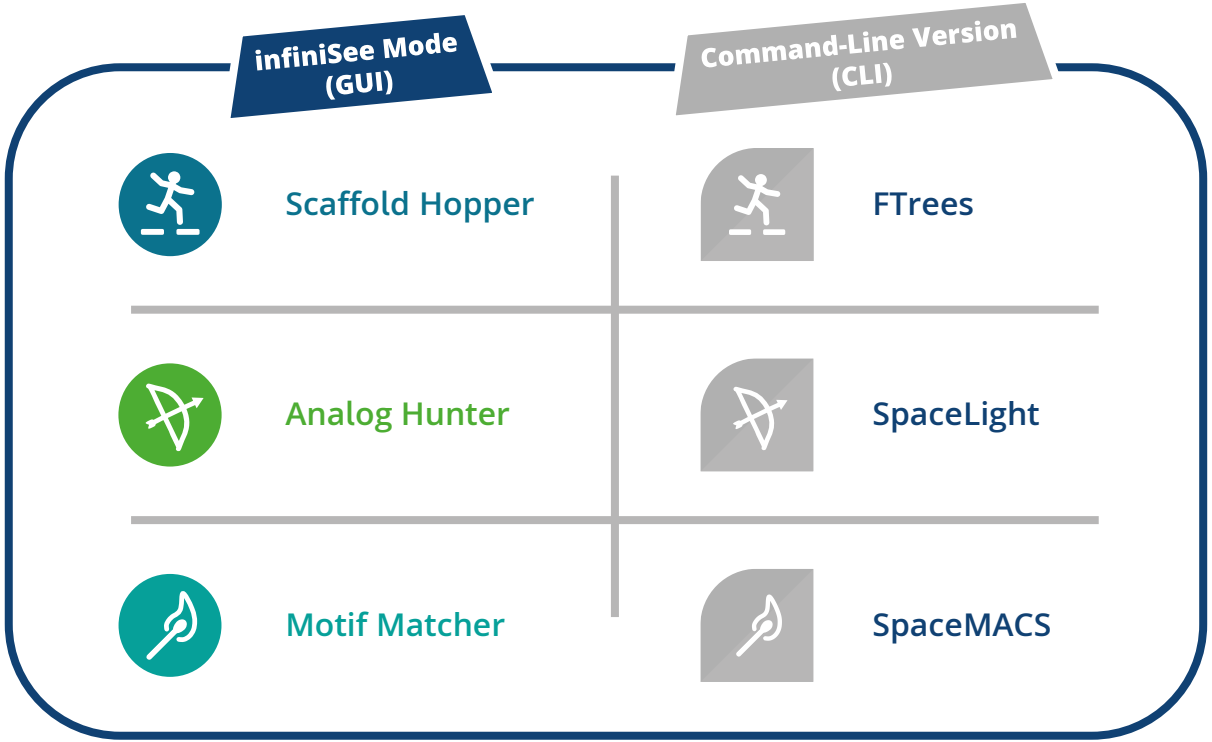
You can define the exact positions for allowed growing by introducing R groups into the molecule. The R group search will only retrieve compounds containing at least one additional heavy atom at the defined positions. The rest of the molecules remains untouched.

Use this to extract the most relevant compounds containing your fragment with extensions in the desired positions.



Overview of Our Tools

GUI and Command-Line Versions



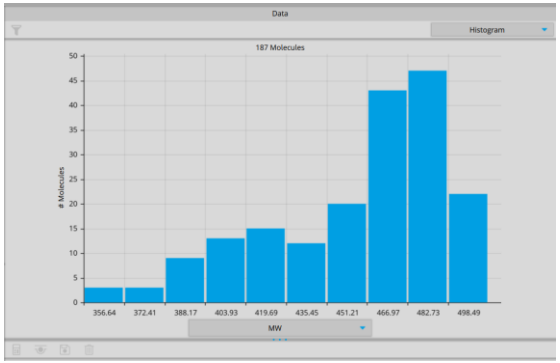
Analyzer Mode

Understand Your Data Set

infiniSee's Analyzer Mode is the aerial view panel for your compound pool your partner to develop a feeling for a molecule ensemble. Manage your data efficiently to focus on what is important.

Mine for the gems

Examine your infiniSee results or any compound source to refine the selection of compounds based on your specified parameters.



Interactive histograms

Split your results based on physicochemical properties into explorable segments.

Working with molecule subsets is an effective method to select chemically diverse candidates for follow-up. Compounds can be clustered based on **Bemis-Murcko scaffolds** and skeletons, vendors and reagents.