



THE EVOLUTION OF CHEMICAL COLLECTIONS AND DRUG-LIKE DIVERSITY: THE ROLE OF A COMPOUND AGGREGATOR

The story behind Molport's
5-million compound library

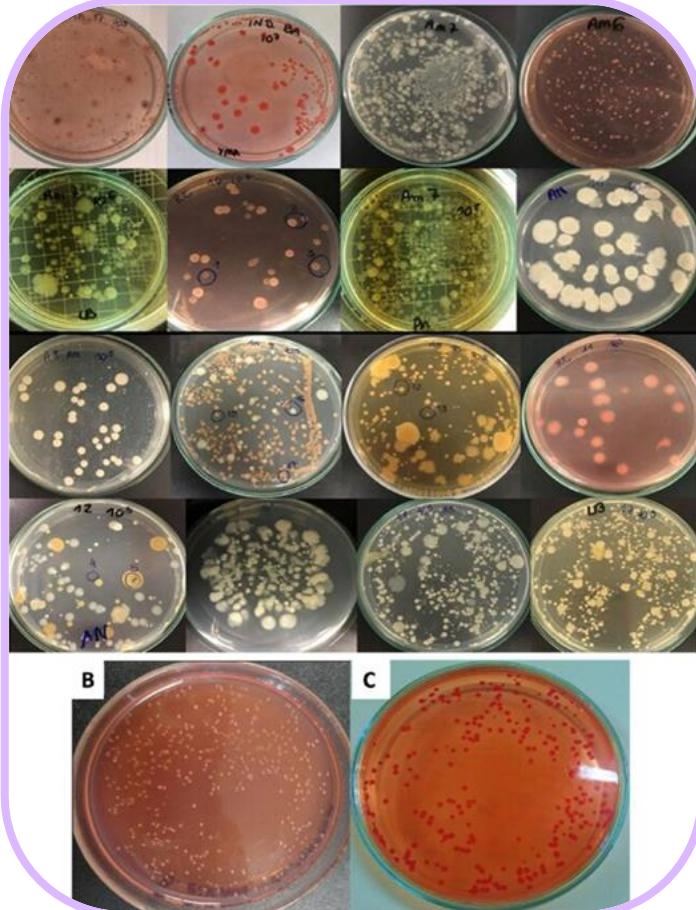
Andrea Altieri, PhD
Molport Chief Scientific Officer

September 11th 2025



Early Screening Approaches

Late 19th – early 20th century



End of 20th century



Gyula Takatsy



96-well plates

HTS and Compound Libraries



1990s: term of "High-Throughput Screening"

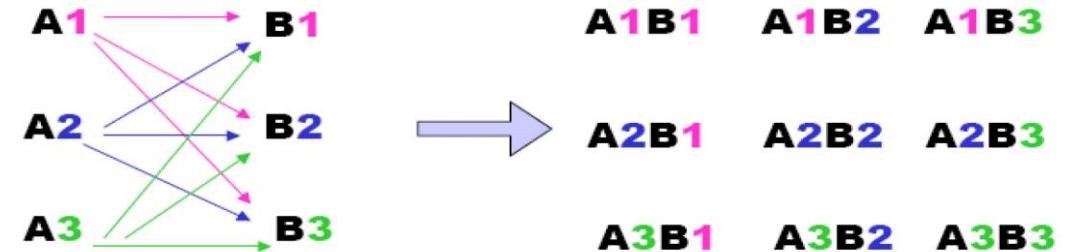
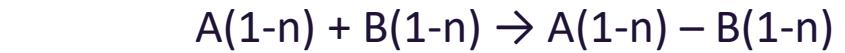
HTS facility: screening over **300,000** compounds per week (Pfizer).

The Need of Feeding HTS Platform

In-House/Historical
Compound Collections



Combinatorial Chemistry



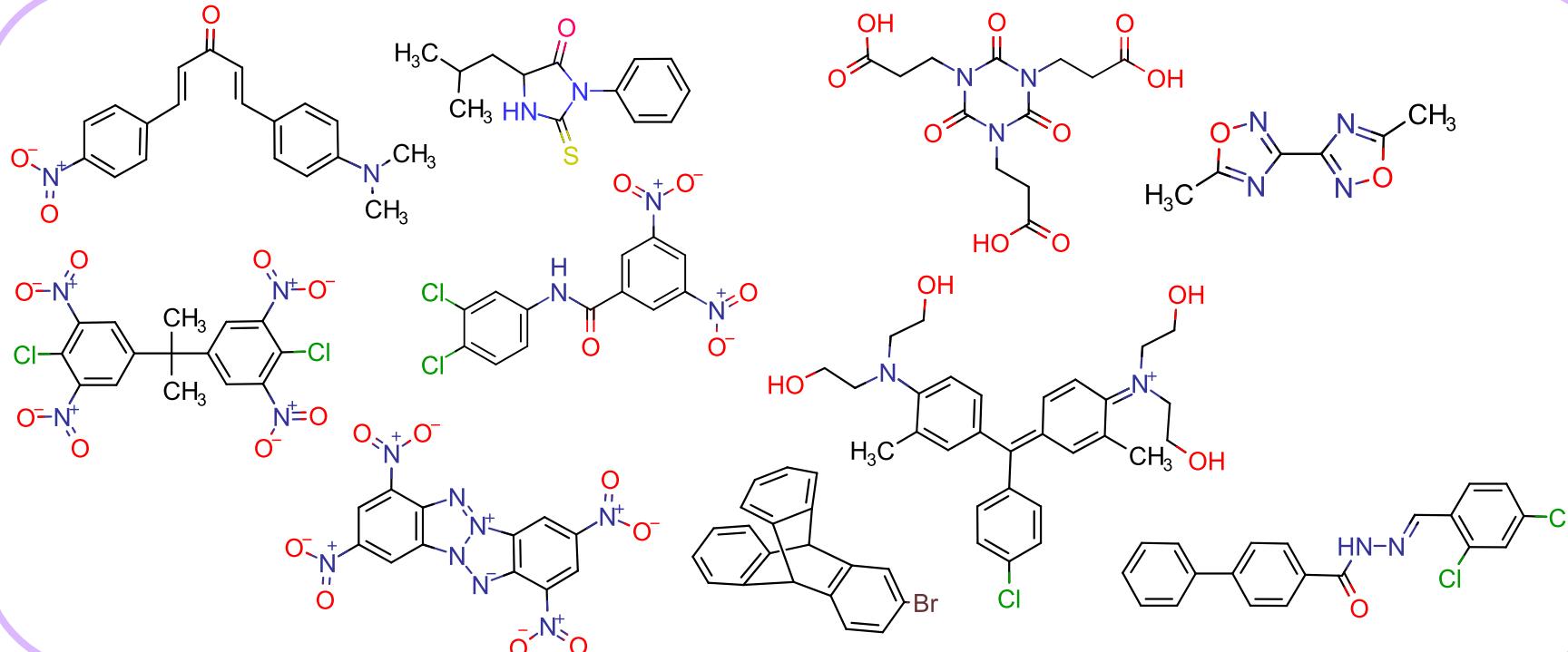
9 different compounds

Shaikh, S. M. et al. Res. Rev. J. Chem. 2017, 6 (2), 14–26.

Early Stage Compound Collections

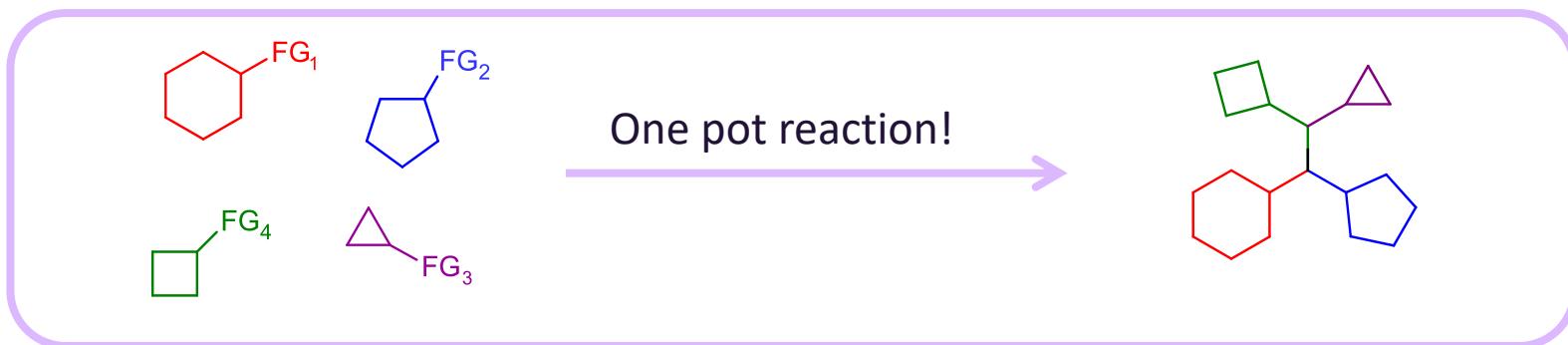
In the 1950s–70s, companies like Burroughs Wellcome, Pfizer, and Merck began assembling systematic collections of small molecules to test for biological activity.

What did the early compound collection look like (1994)?

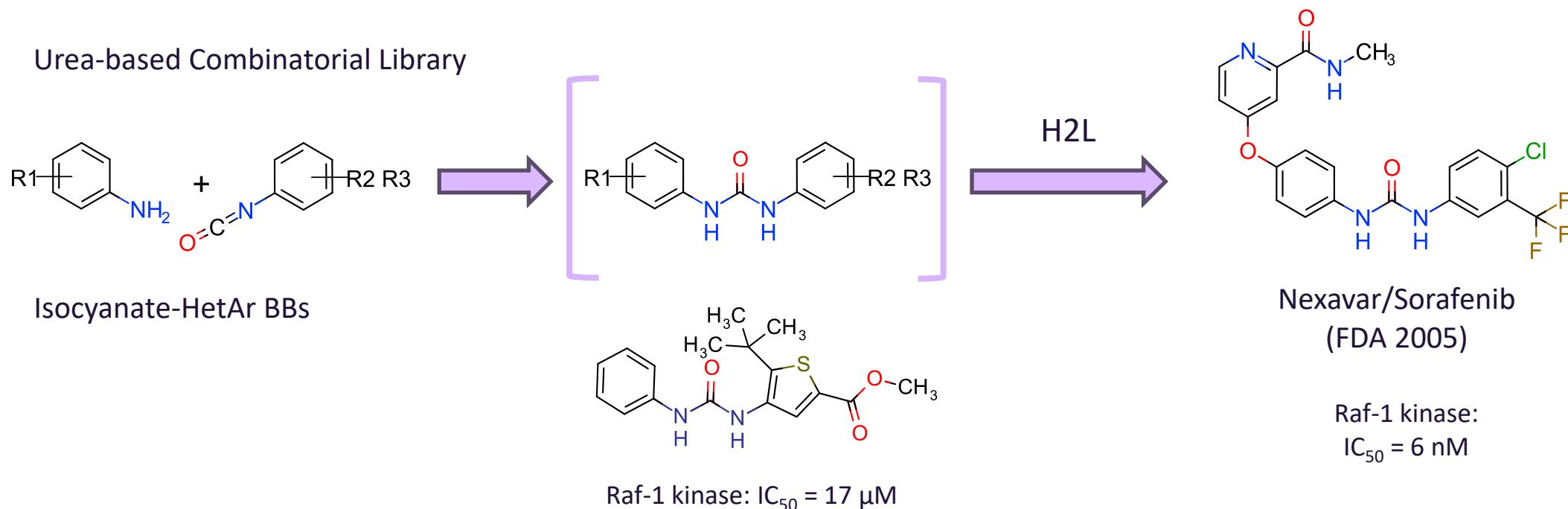


- 10,000 to 2-3 million compounds
- Highly diverse
- Old compounds
- Compound integrity?
- Novelty?
- Chemistry reproducibility?
- Drug-likeness?

Combinatorial Chemistry and Multi Component Reactions (MCRs)



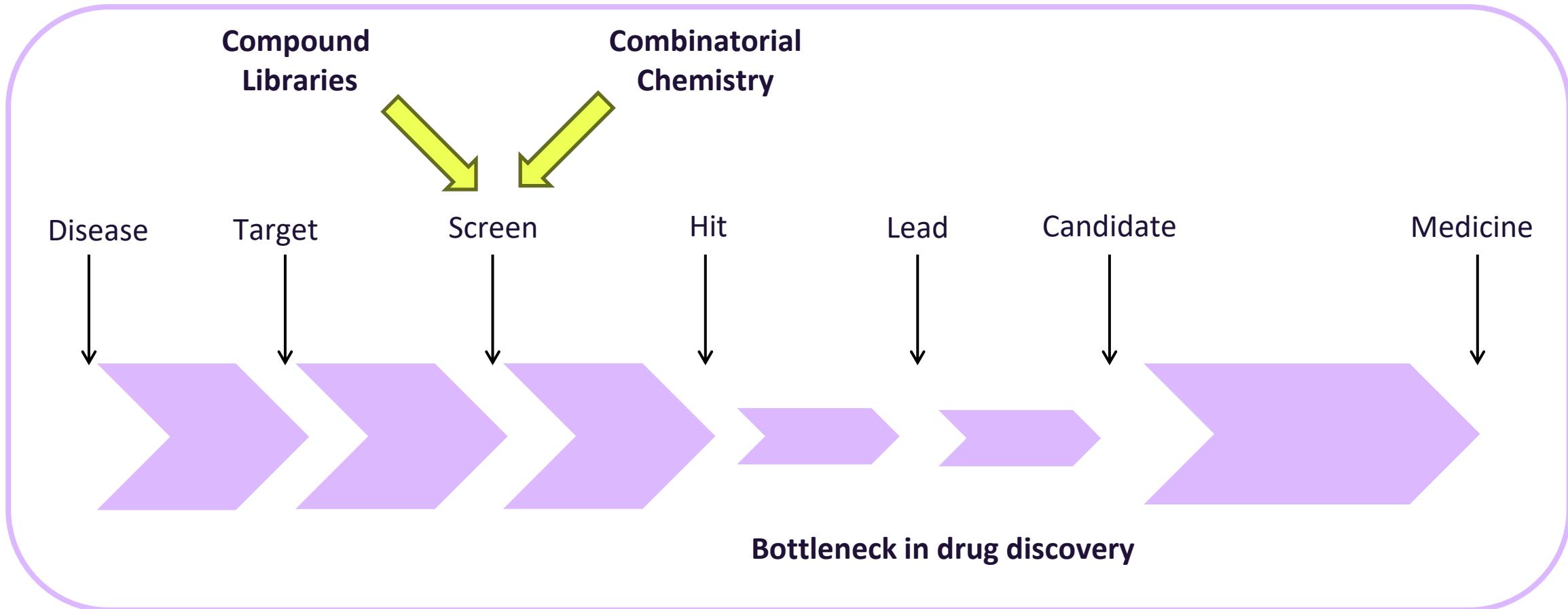
Urea-based Combinatorial Library



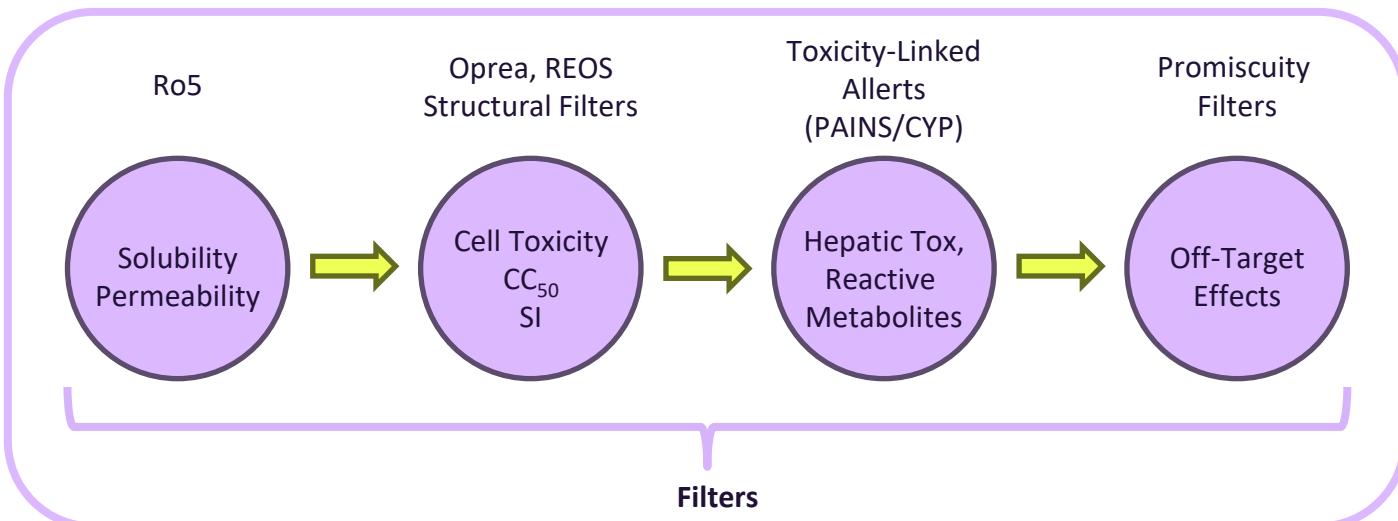
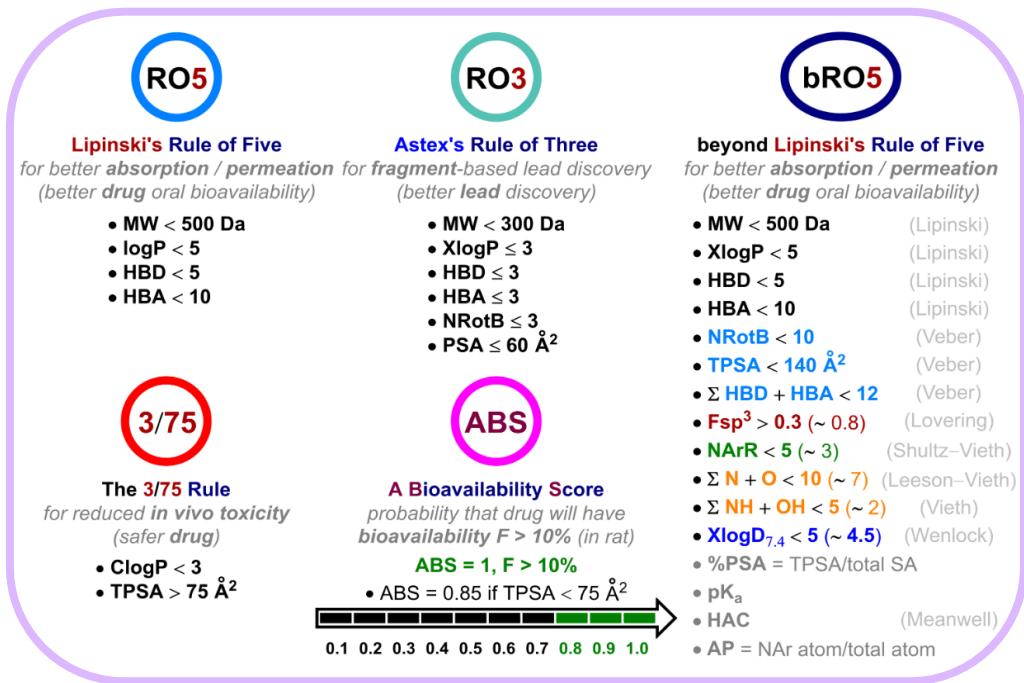
Bayer and Onyx
Pharmaceuticals
kinase inhibitor
(kidney & liver cancer)

How to Improve the Hit Rates in Early Drug Discovery?

What makes a good **Hit** and an ideal Screening Compound?



Lipinski's Rule of 5 and Drug-Likeness Concept



Filter the Compound Collections /Selected Screening Set for Hit Compound
Tox /Unwanted Properties

Sen, D. J. et al. World J. Adv. Healthc. Res. 2021, 5 (3), 206–211.

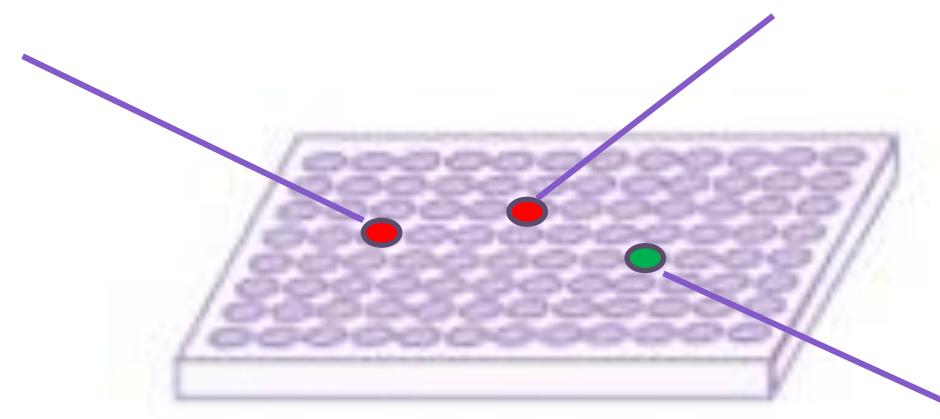
Hits vs Good Quality Hits: Pragmatic Aspects

The **singleton** problem!

No analogues
for SAR models
No synthetic protocols

The **stock-level** problem!

No enough material for hit confirmation

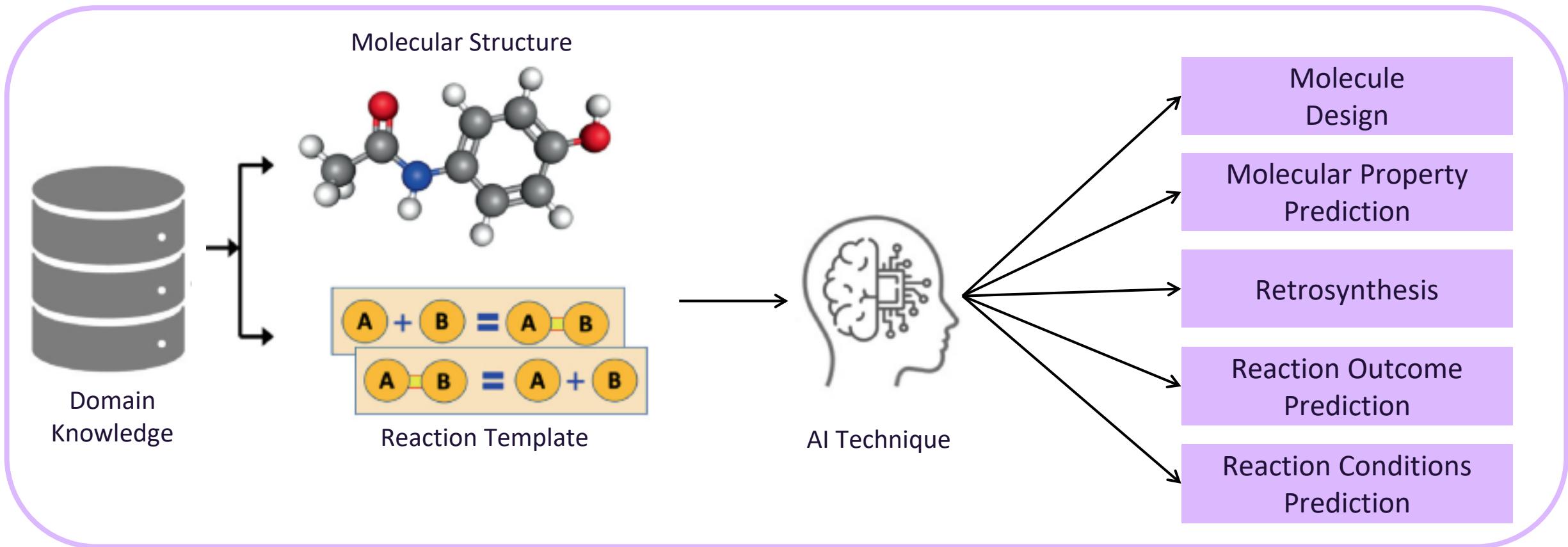


Verified Compounds!

“Cured” Libraries increase your chances of finding real hits

AI-Powered Drug Design: A New Era in Combinatorial Chemistry

Modern Combinatorial Chemistry and AI & Reaction's Portfolio



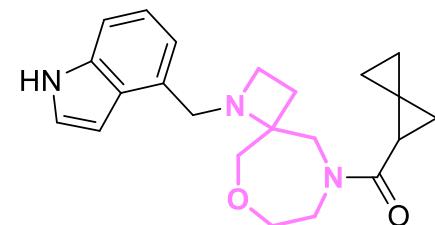
Molport Nova Deck Reagent Toolbox

1159 cured mono Boc protected di-amine BBs

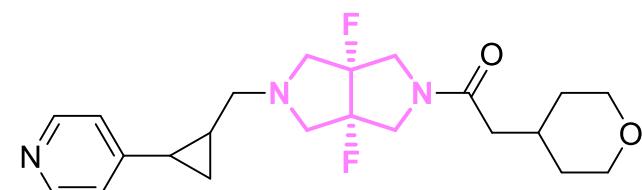
3420 cured RCOOH BBs

Predicted Yields											
1		2		3		4		5		6	
94	91	41	80	56	93	100	98	74	84	56	83
76	95	97	83	82	58	46	52	38	97	97	23
37	85	19	90	21	20	83	72	64	49	87	82
14	75	94	34	31	83	73	12	84	100	79	28
89	20	85	83	69	36	17	56	84	76	74	87
74	100	84	27	92	98	95	60	83	98	66	62
92	46	95	14	94	89	18	64	54	100	63	80
99	95	100	66	82	84	56	61	18	63	93	87
27	81	29	92	88	42	25	45	93	70	82	80
97	91	93	88	87	11	92	41	98	68	28	90
81	43	88	98	47	90	80	95	14	16	53	35
48	97	75	69	46	68	22	37	85	70	83	95
94	26	27	66	42	81	99	43	99	49	99	28
63	99	59	59	60	100	95	78	85	92	39	71
87	86	36	84	84	99	26	96	92	97	90	87
67	90	46	91	92	12	52	85	86	98	85	95
33	90	15	29	16	79	94	82	92	16	96	87
92	33	81	97	97	48	32	82	36	83	99	74
23	90	72	27	90	75	99	20	100	51	89	20
30	82	98	93	81	91	99	49	81	100	98	95
94	82	96	93	90	90	88	91	80	65	88	31
98	15	88	94	80	54	88	99	24	74	96	85
83	82	23	16	97	71	89	29	92	88	88	73
98	18	82	81	50	80	93	87	89	56	79	97
94	91	91	71	90	92	92	73	96	32	100	82
96	88	57	44	59	82	36	74	42	37	83	99
99	82	21	72	57	95	55	93	98	80	39	91
84	61	89	74	91	86	93	93	83	90	74	41
68	90	14	99	93	92	64	100	75	96	99	62
65	18	82	63	100	89	12	88	48	95	74	32
68	93	94	82	95	89	86	22	69	53	47	100
96	27	43	78	97	86	21	89	62	42	90	45
48	23	88	56	56	33	56	59	90	76	56	100
17	25	91	94	25	30	81	77	91	93	92	91

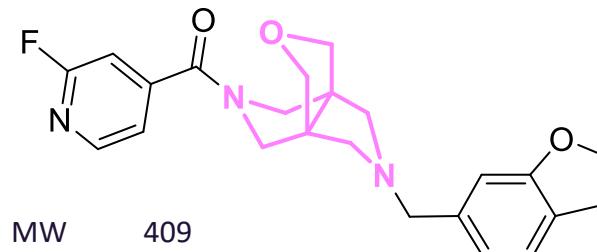
Examples of final virtual compounds:
(0 entrance in 2D SSS PubChem)



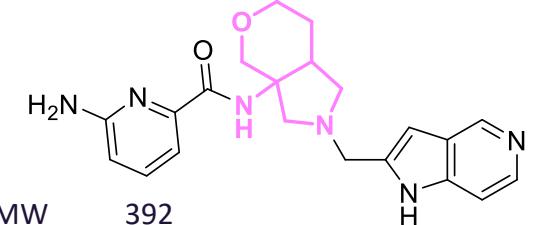
MW 365
cLogP 1.9
HBA 3
HBD 1
TPSA 46



MW 405
cLogP 1.5
HBA 4
HBD 0
TPSA 46

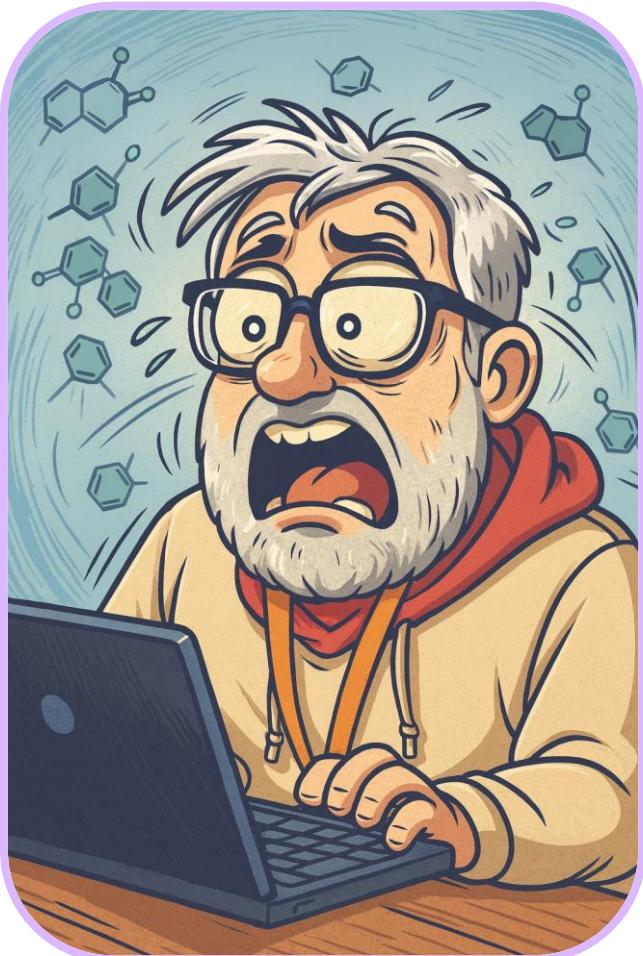


MW 409
cLogP 1.4
HBA 5
HBD 0
TPSA 55



MW 392
cLogP 0.72
HBA 6
HBD 3
TPSA 109

The Harsh Reality After In-Silico Screening



**Weeks/Months of computing...
One moment of disappointment!**

Molecule Was Not Found

Requested molecule was removed or was not found in the chemical search, [SMILES search](#) or use text search to find

- **Top priority compounds often OUT-OF-STOCK**
unavailable when needed the most
- **Beautiful molecules, but prohibitively expensive**
beyond project budget
- **Promising hits with no future**
zero or limited synthetic vectors, no room for optimization

Industry Issues: Orders from Different Vendors



- Multiple suppliers
- Complex search
- Unreliable data
- Made to order vs Stock compounds
- Several POs
- Several deliveries
- Mess in documentation
- Customs clearance

Advantages of Aggregated Libraries

- ✓ Efficiency in compound sourcing and procurement.
- ✓ Uniformity in data curation, quality control, and availability tracking.
- ✓ Enabling cross-supplier comparisons and selections within one platform.



Molport: Global Leader in Compound Sourcing



- ✓ 18+ Years of Expertise
- ✓ 5.5M Unique In-Stock Compounds
- ✓ 80+ Verified and Trusted Suppliers
- ✓ Publicly Available Database
- ✓ Advanced Search Tools
- ✓ High Quality Laboratory Services

The screenshot shows the Molport website homepage. At the top, there's a search bar with the placeholder "Search by Molport ID" and a magnifying glass icon. To the right of the search bar are buttons for "Ship to:" (with "US" selected), "Log In", and "Reg". Below the search bar is a navigation menu with links for "Find Chemicals", "Services", "Database", "Resources", and "About Us". A prominent banner in the center features the text "18+ Years of Expertise" and "Compound Sourcing and Formatting Made Easy". Below the banner is a search bar with the placeholder "Search by Name, ID, SMILES, CAS, InChI, MolFormula...". Further down, there are buttons for "List Search" and "Structure Search". To the right of the main content area, there's a sidebar with a purple header that says "5.5M In-stock Screening Compounds" and "83 Verified Suppliers". A "Download D" button is also visible.

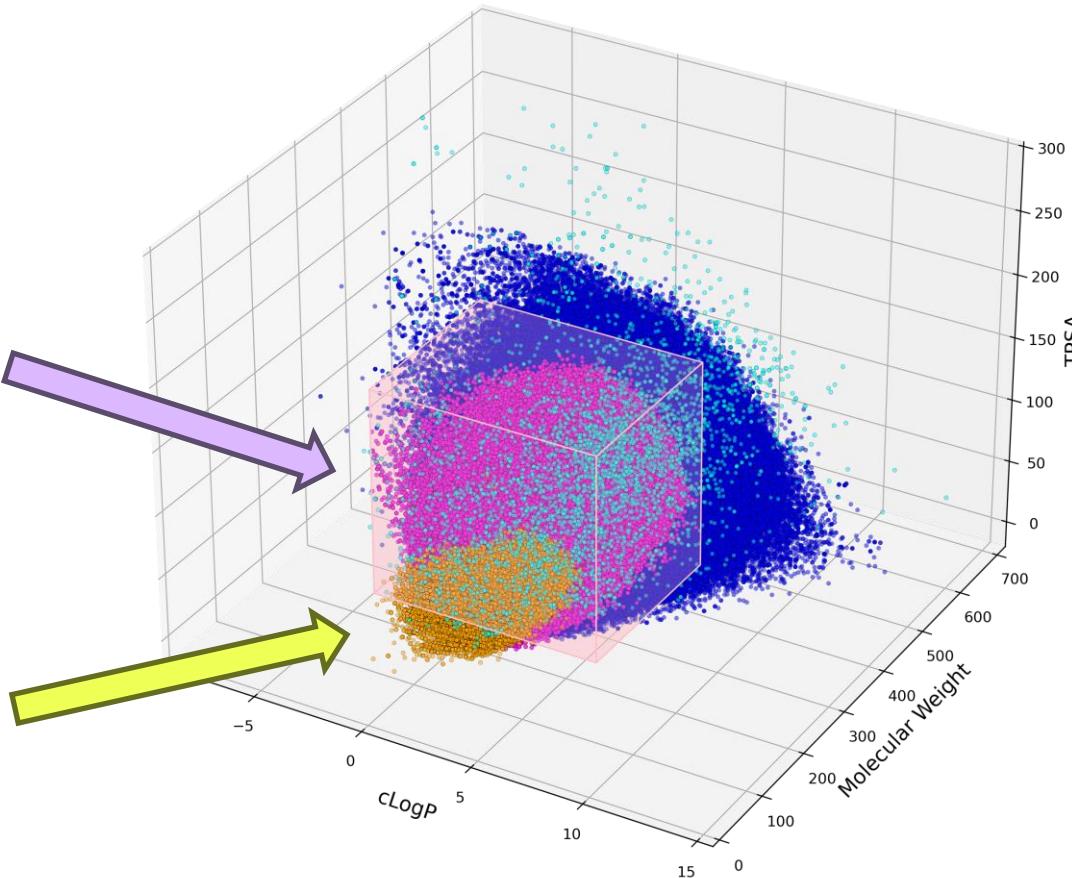


5.1M Molport Chemical Space Profile



2.1M drug-like compounds

150k fragments



MW, cLogP, and TPSA Distribution

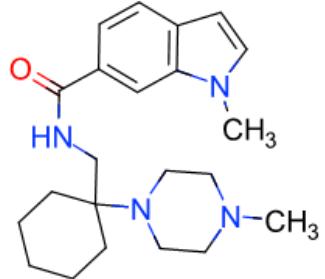
Full db (blue), Ro3-fragments (orange), Ro5-drug-like (magenta), and FDA-approved drugs and drug candidates (cyan)

Physicochemical filters
Ro3/Ro5

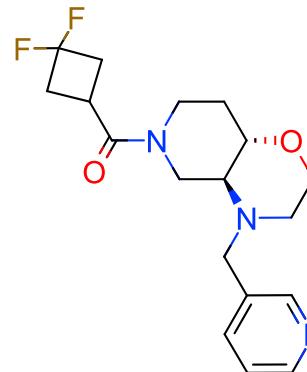
OPREA structural filters
PAINS filters

Built for Drug Discovery: Molport's Curated Screening Sets

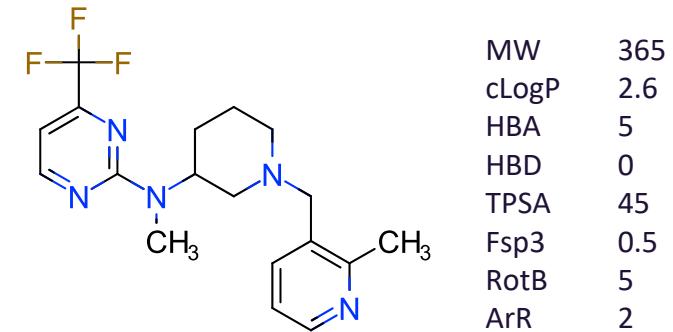
Lead-like, toxicity-filtered, and optimized for real results: compound representatives



MW	368
cLogP	2.5
HBA	5
HBD	1
TPSA	41
Fsp3	0.8
RotB	4
ArR	2



MW	351
cLogP	1.1
HBA	5
HBD	0
TPSA	46
Fsp3	0.8
RotB	3
ArR	1

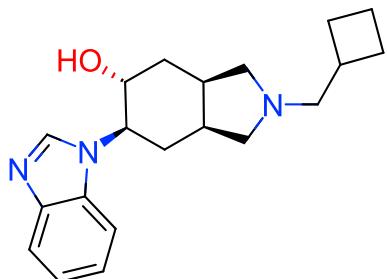


MW	365
cLogP	2.6
HBA	5
HBD	0
TPSA	45
Fsp3	0.5
RotB	5
ArR	2

Molport-030-044-170

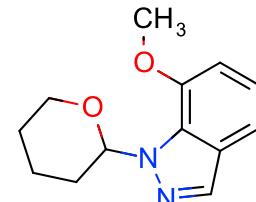
Molport-047-698-376

Molport-051-879-838



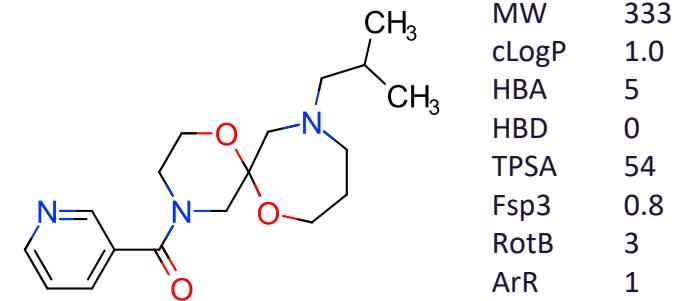
MW	325
cLogP	2.9
HBA	3
HBD	1
TPSA	30
Fsp3	0.6
RotB	3
ArR	2

Molport-052-032-756



MW	232
cLogP	2.1
HBA	5
HBD	0
TPSA	36
Fsp3	0.8
RotB	3
ArR	2

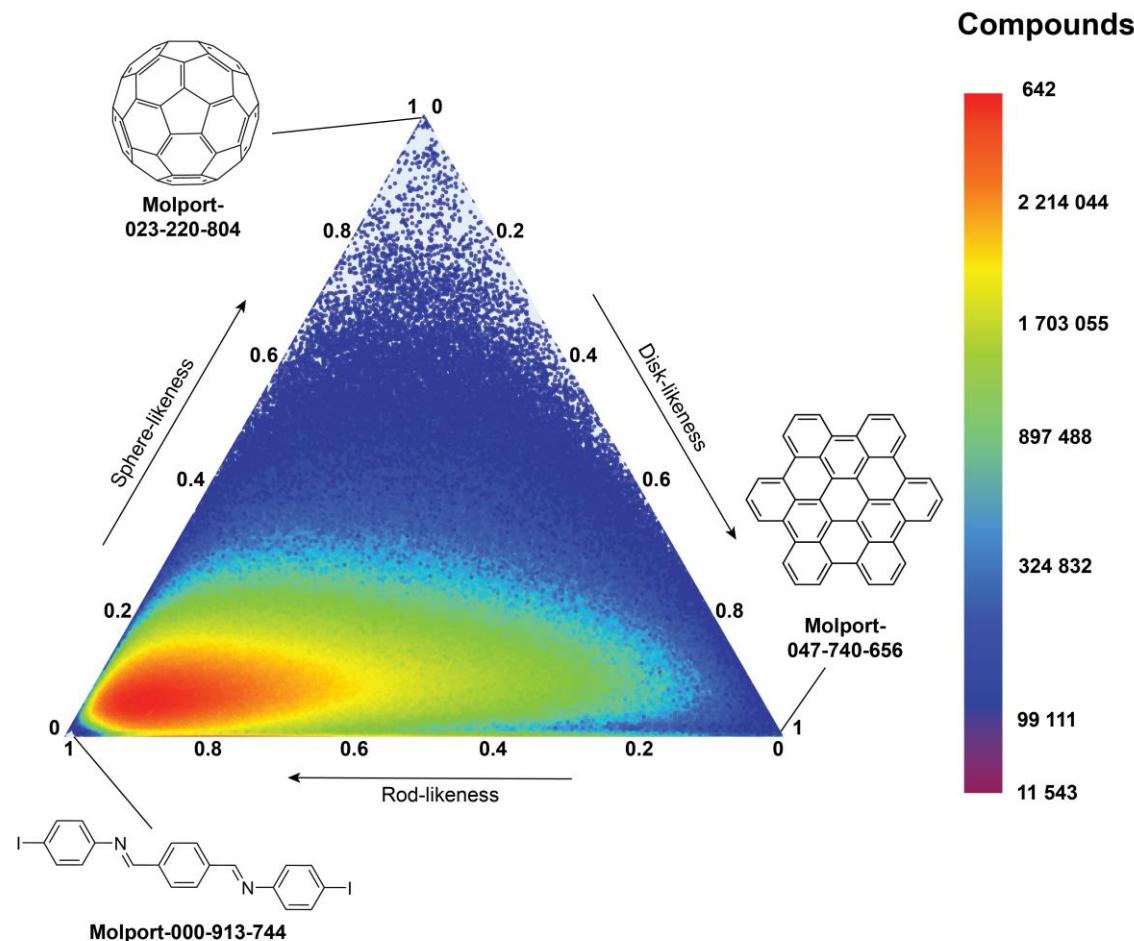
Molport-051-993-042



MW	333
cLogP	1.0
HBA	5
HBD	0
TPSA	54
Fsp3	0.8
RotB	3
ArR	1

Molport-051-933-188

3D-Shape Analysis of 5.1M Molport database



Generate 3D conformers

Minimize conformers by MMFF
force field

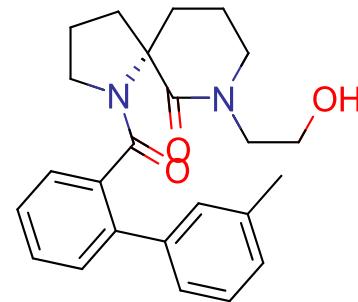
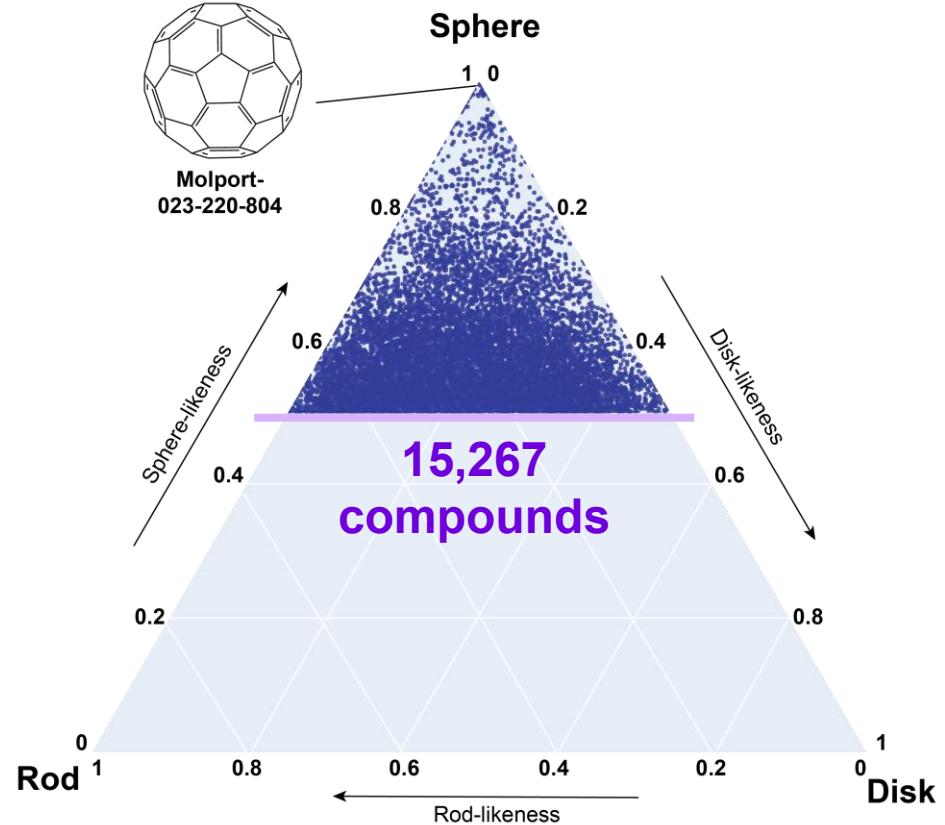
Compute sphere-likeness
Rod-likeness, and
disk-likeness

RDKit 2024.09.3

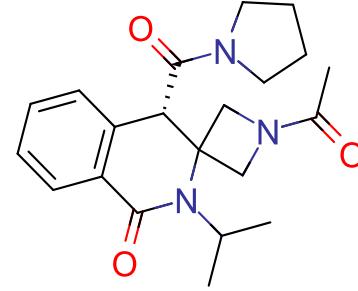
Heat Map of Normalized Principal Moments of Inertia (PMI)

The most occupied regions colored in red, less occupied – in purple

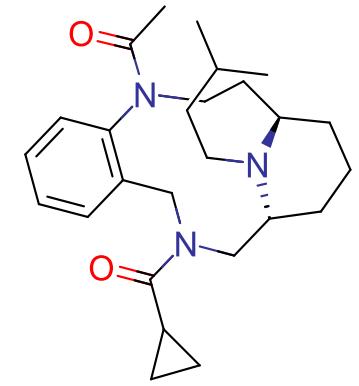
3D-Supershape Enhanced Library



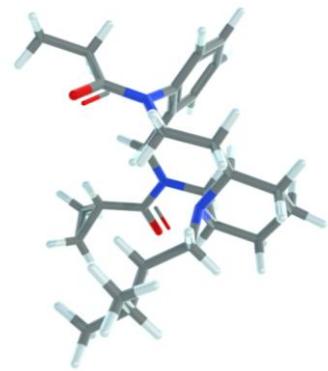
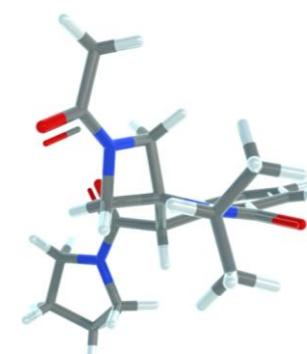
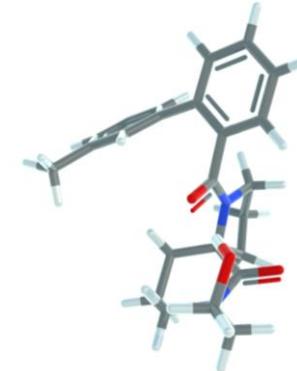
Molport-051-474-927



Molport-046-628-845



Molport-047-061-797



Structural Diversity

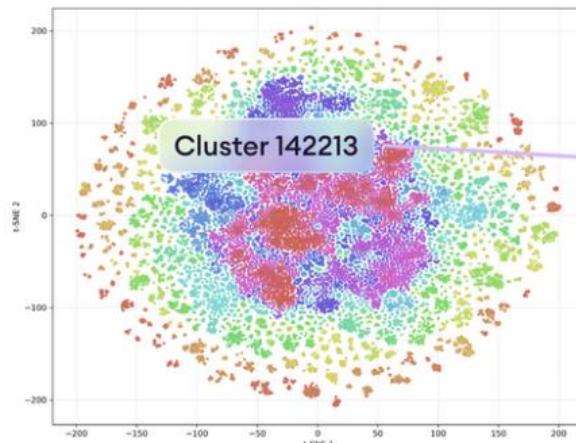
The 5.1M Molport compound repertory
 drug-like properties calculated
 Ro5 modified filter applied
 2.1M compounds

OPREA structural filters
 Stock-level filters

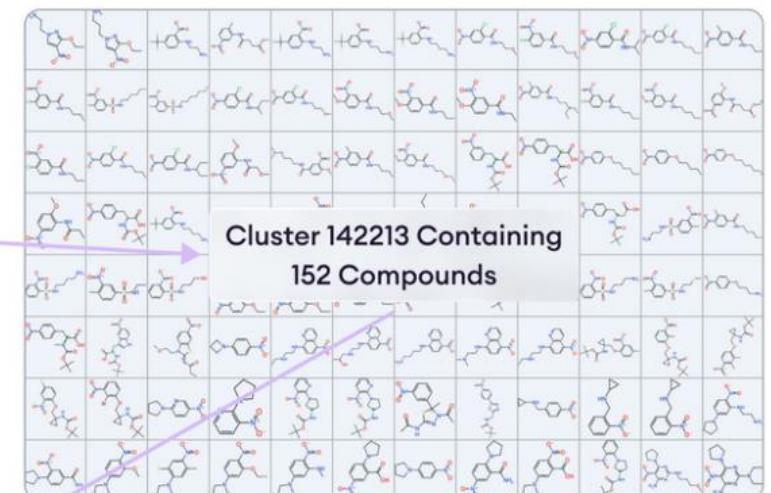
Tanimoto similarity of 0.7
 Morgan fingerprints
 (radius = 2, 2048 bits, chiral)

Singleton filtering
 Cut off < 10 representatives

Clustered Molport Drug-Like Chemical Space,
 Visualized with t-SNE and HDBSCAN



Total: 16,522 clusters



Ro5 and Drug-Like Conform



RDKit 2024.09.3

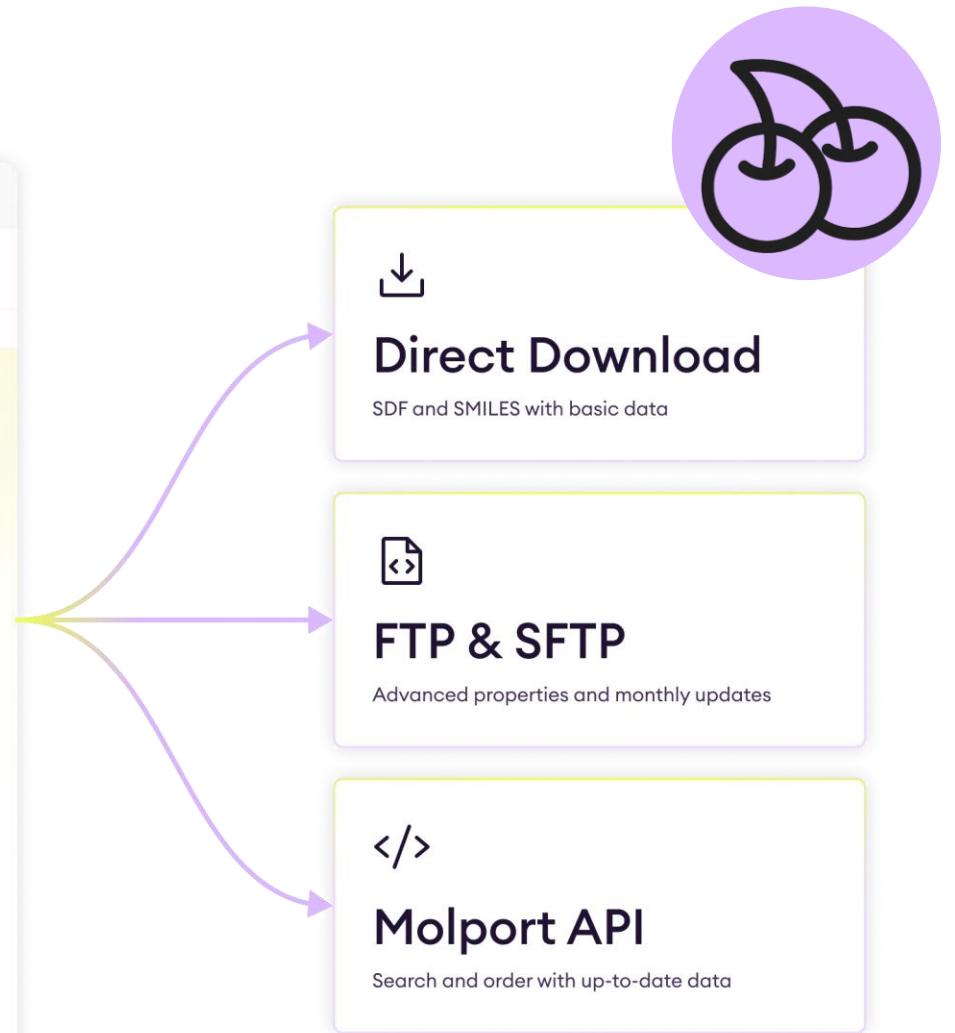
Database Access

Publicly Available Molport Database

The screenshot shows the molport.com homepage. At the top, there's a navigation bar with links for "Find Chemicals", "Services", "Database", "Resources", and "About Us". Below the navigation is a banner with the text "18+ Years of Expertise" and a large heading "Compound Sourcing and Formatting Made Easy". There are two search bars: one for "Search by Molport ID" and another for "Search by Name, ID, SMILES, CAS, InChI, MolFormula...". Below these are buttons for "List Search" and "Structure Search". On the right side of the homepage, there's a section titled "Compound Libraries" featuring icons and details for five libraries: Amino Acid Library (42,216 compounds), Covalent Library (419,822 compounds), GPCR Library (415,426 compounds), Kinase Library (105,639 compounds), and DEL Reagent Library (119,913 compounds). To the right of the homepage, a detailed view of the "Molport In-Stock Database" is shown. This view includes a summary table with the following data:

In-stock Screening Compounds	In-stock Building Blocks
5.1M	617k
83 Verified Suppliers	15 Compound Libraries

A prominent button labeled "Download Database" with a download icon is centered below the summary. A note at the bottom states: "The full database files were last updated on June 6, 2025." A "View All Libraries" link is located at the bottom right of this panel.



Target-Driven Sets

Filtered from 5.1M Database:



Natural Products

426,787 compounds



Kinase Library

105,639 compounds



PPI Library

107,266 compounds



GPCR Library

415,426 compounds



Carbohydrate Library

28,535 compounds



Known Drugs & Clinical Candidates

28,535 compounds

Technology-Driven Sets

Filtered from 5.1M Database:



Drug-Like Library

1,767,717 compounds



Fragment Library

107,266 compounds



Covalent Library

419,822 compounds



Amino Acid Library

42,691 compounds



DEL Reagent Library

119,913 compounds

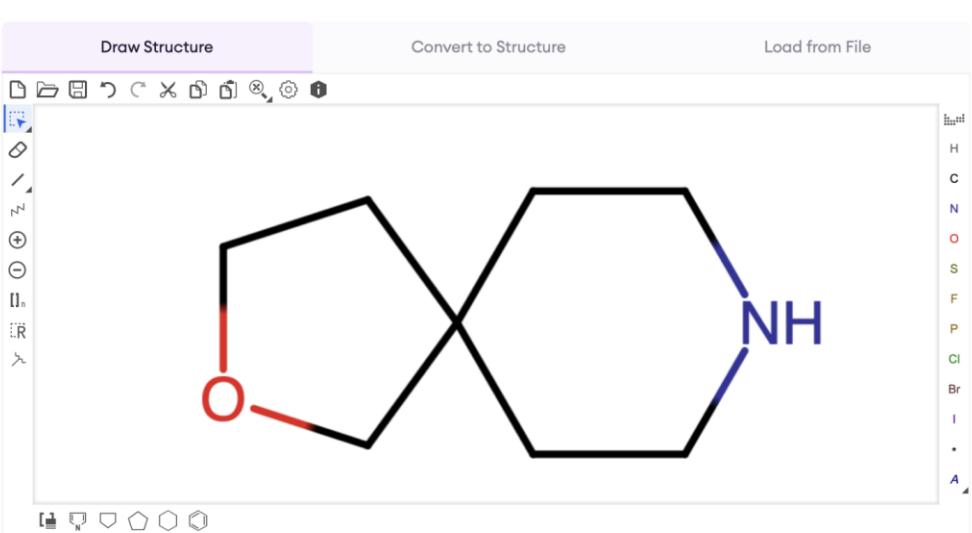


Diversity Library

307,739 compounds

Structure and Similarity Search

Draw Structure Convert to Structure Load from File



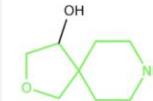
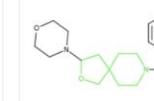
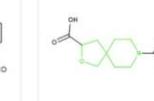
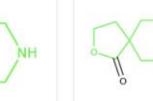
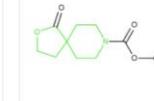
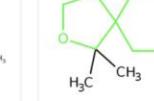
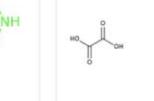
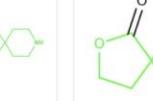
Maximum Search Results: 10000

Draw and Search

Search results

2397 results were found for chemical search

0 Selected

In stock	Molport-004-765-402	In stock	Molport-006-841-782	In stock	Molport-009-679-570	In stock	Molport-011-133-347	In stock	Molport-023-198-733
	View Pricing		View Pricing		View Pricing		View Pricing		View Pricing
	View Pricing		View Pricing		View Pricing		View Pricing		View Pricing

Review and Select

Structure and Similarity Search

2-oxa-8-azaspiro[4.5]decane

In stock

Molport ID: Molport-011-133-347
Molecular Formula: C₈H₁₅NO
Molecular Weight: 141.214
SMILES: C1CC2(CO)CCNCC2
CAS Number: 176-69-2; 479195-19-2

InChI: InChI=1/C8H15NO/c1-4-9-5-2-8(1)3-6-10-7-8/h9H,i-7H2
InChI Key: OXXXNISRXFPBK-UHFFFAOYSA-N
InChI Main Layer: C8H15NO/c1-4-9-5-2-8(1)3-6-10-7-8/h9H,i-7H2

Functional Groups: All Amines

[Edit Structure](#) [Find Similar](#)

Marketplace Offers
31 offers from 6 suppliers

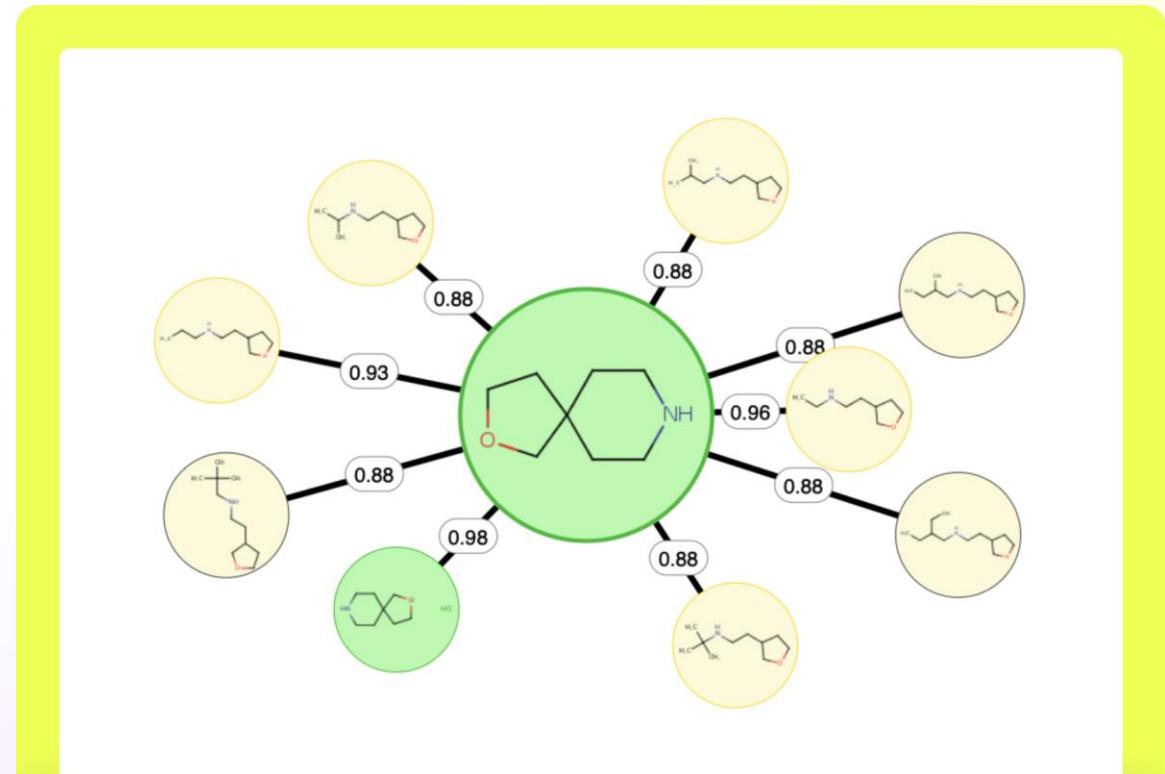
Pack Size: Best (165.00 \$ • Delivery by 3 Jul 2025), Fastest (224.00 \$ • Delivery by 25 Jun 2025), Cheapest (165.00 \$ • Delivery by 3 Jul 2025)

Compound Price: Min 98, Max 3094

A2B Chem LLC offers:
100 mg, Purity 95%, Warehouse in United States, Delivery by 3 Jul 2025, 130.00 \$ + 35.00 \$ Direct Shipping to US
250 mg, Purity 95%, Warehouse in United States, Delivery by 3 Jul 2025, 186.00 \$ + 35.00 \$ Direct Shipping to US

[Add to cart](#)

Explore Marketplace Offers



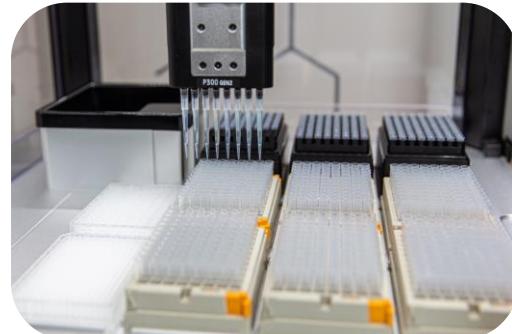
Compound Sourcing & Formatting Made Easy

Order



Formatting

Custom Format



Package

Custom Format

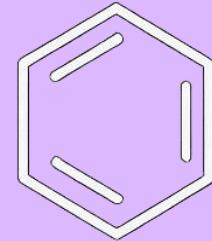


Molport In-house Formatting
Department

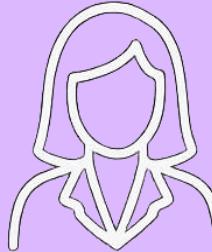


Why Collaborate with Molport?

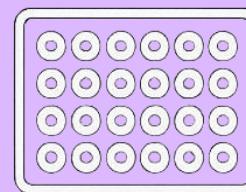
molport



Access to pre-selected vendors and their compound space



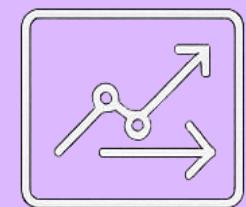
MedChem expertise supporting custom screening projects



Flexibility in compound formatting, plating, and logistics



Cost and time savings (Aggregation Value)



Integration into existing end-user workflows & informatics systems

Acknowledgments

Molport Team

Diana Zelencova-Gopejenko	Ieva Kluga
Andrea Altieri	Ieva Kovalenko
Martin Razuks-Ebels	Rolands Cinis
Janis Oslejs	Rolands Capars
Kristine Skulme	Elza Dzene
Artis Erglis	Helena Daiga Litvina
Kristaps Boss	Enia Tina Susta
Normunds Kinne	Dita Duka
Andrievs Auzins	Inese Sulca
Raitis Brodezonoks	Kristine Rauska-Zeltina
Inga Salmina	Kirils Bidzans
Uldis Ledainis	

andrea.altieri@molport.com
info@molport.com

