Navigating Synthon Space: Property-Driven Molecular Optimization for Pharmacokinetics

Rafał A. Bachorz

Simulations Plus

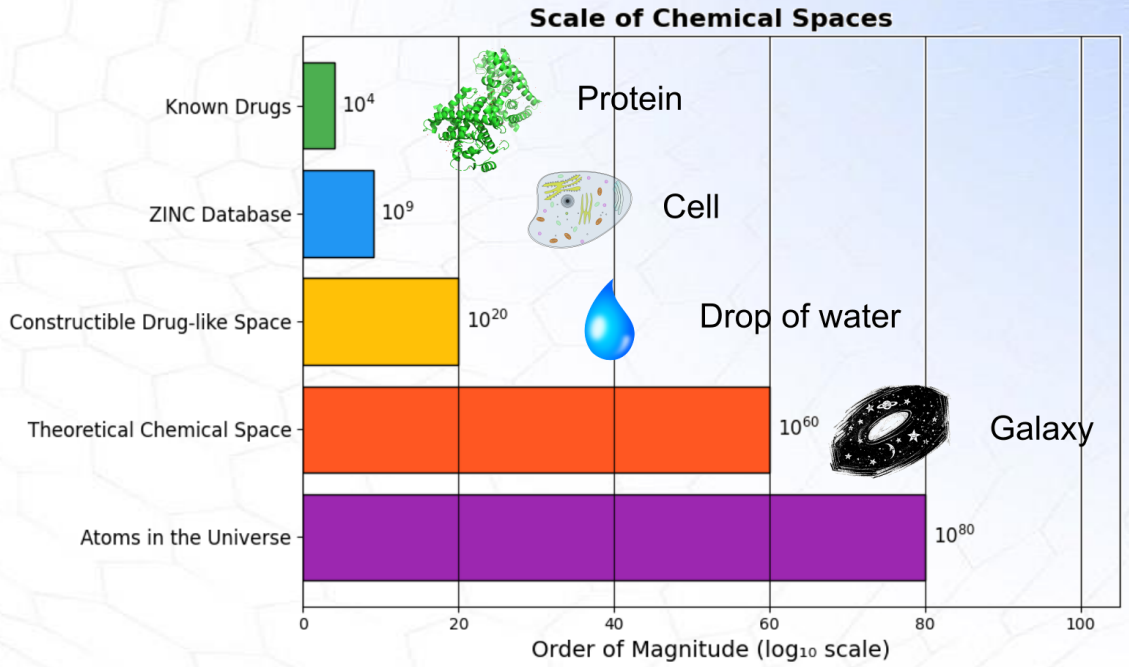
Agenda

- The Complexity of Chemical Space
- Synthons as a Smarter Representation
- PK Profiling with ADMET Predictor
- Multicriteria Decision Analysis with VIKOR
- SynthonForgePlus – multiobjective generative chemistry framework
- Workflow and Implementation
- Use Case:
  - RORγ Agonist Optimization (unconstrained) with 3 objectives
  - RORγ Agonist Optimization (targeted) with 4 objectives
- Summary & Outlook

The Complexity of Chemical Space

The Chemical Space is huge

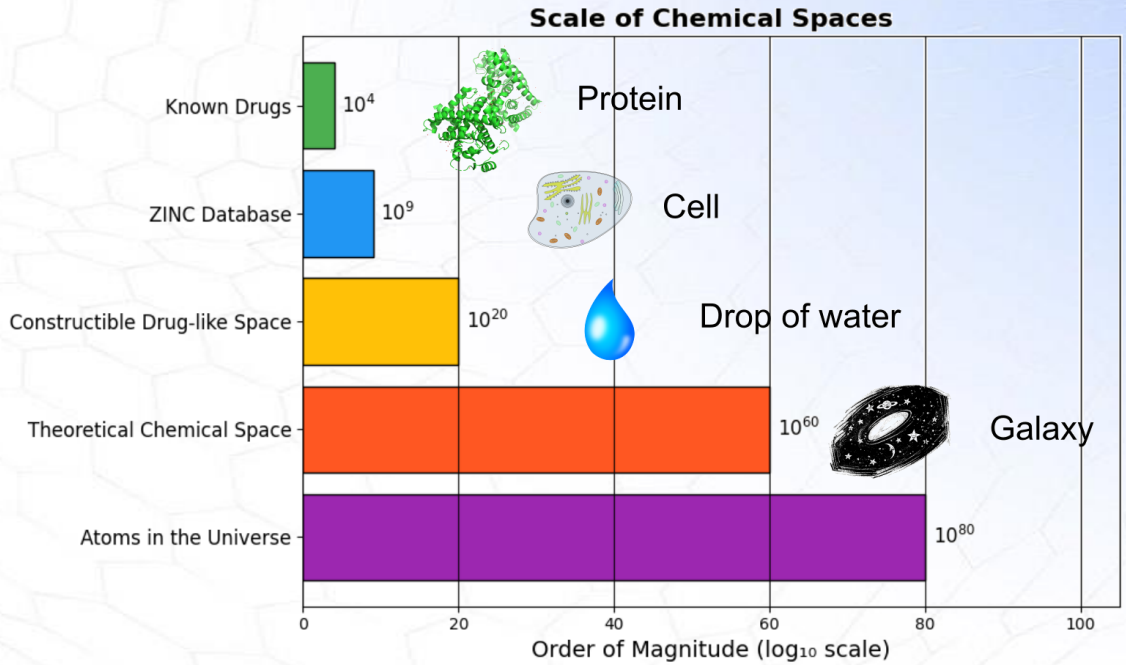
- The sizes:
  - Known drugs (10^4)
  - ZINC database (10^9)
  - Constructible drug-like space (10^{20})
  - Total theoretical space (10^{60})
  - Total number of atoms in universe (10^{80})
- Exhaustive enumeration is computationally infeasible and chemically redundant.
- Even large commercial databases (ZINC, REAL) represent only a tiny fraction.
- We need **goal-directed** exploration to extract value from this vastness.
- Smart, navigable, and property-driven strategies.



Brute-force enumeration is like searching for needles in a cosmic haystack.

Beyond enumeration: Constructing the Search Space

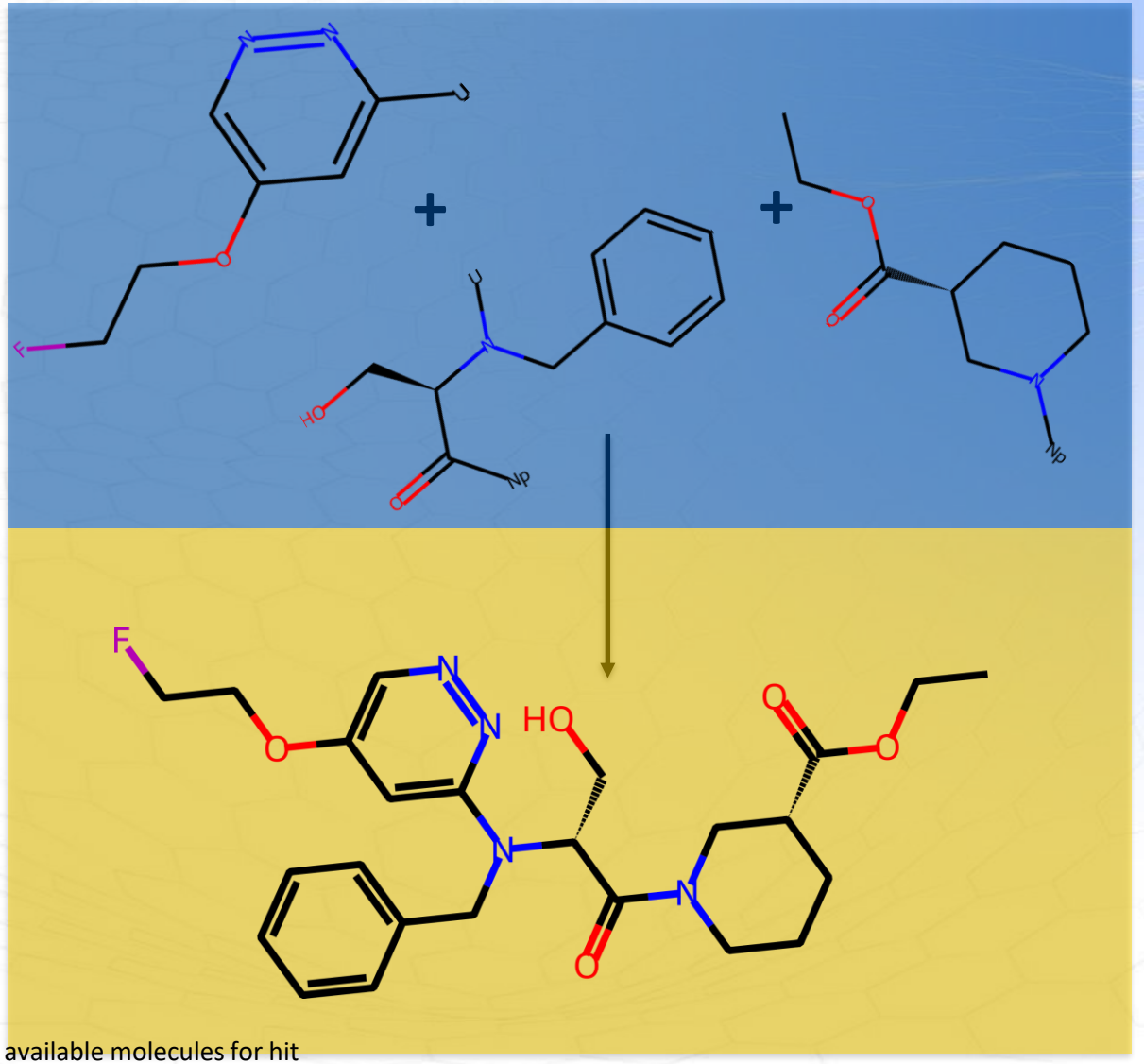
- Traditional enumerated libraries are:
  - **Static**: fixed compounds, no adaptation to project needs.
  - **Redundant**: large portions are chemically or pharmacologically similar.
  - **Costly**: in silico screening of billions still misses the mark.
- Constructive strategies (like synthon-based design) provide:
  - **Scalable expansion** of space on-demand.
  - **Relevance-filtered** by synthesis feasibility and property objectives.
  - **Efficiency**: dramatically smaller yet more diverse effective libraries.



Don't search the haystack—build the needle.

Synthons as a Smarter Representation

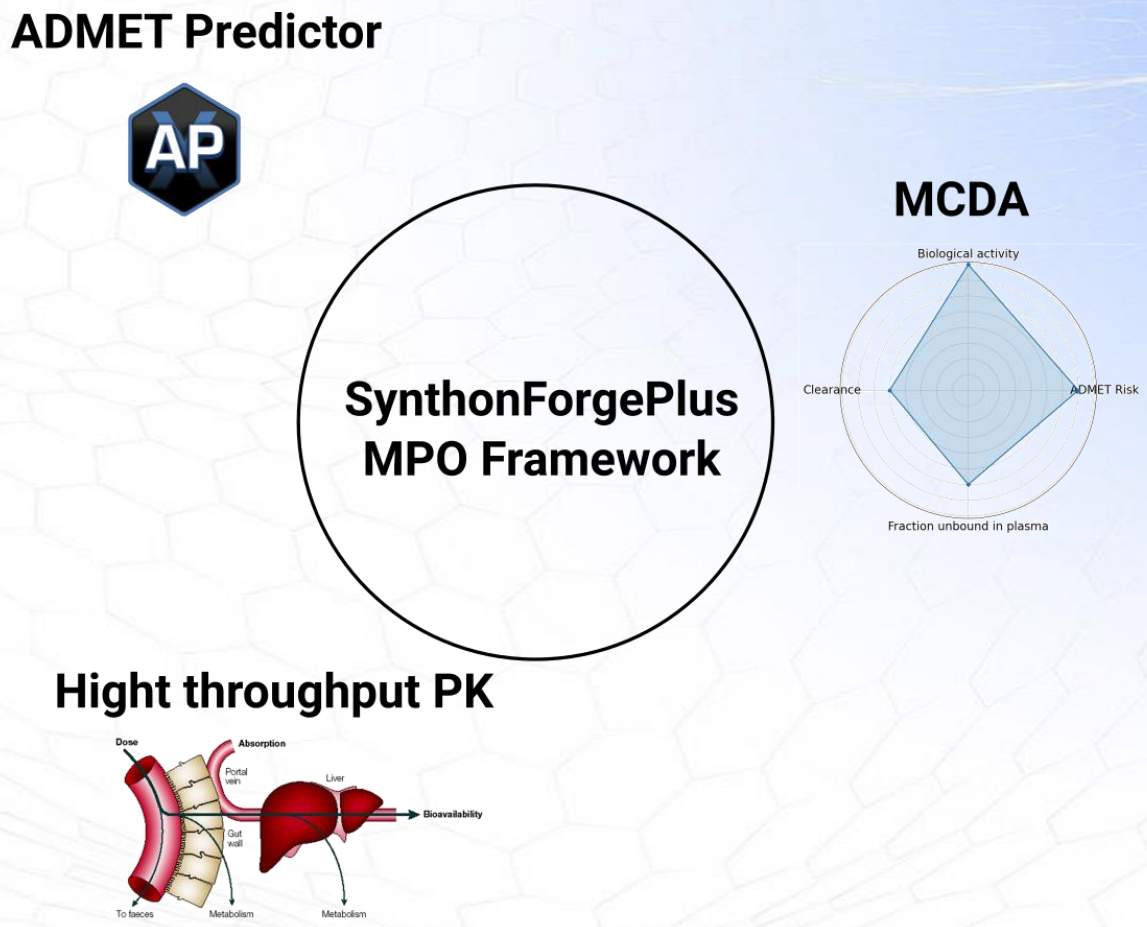
- Synthons = reactive building blocks enabling constructive exploration.
- Synthetic accessible (80-90% success rate).
- Flexible and modular; far more efficient than enumerated space.
- Based on **Freedom Space 3.0** from **ChemSpace**[1]: curated, synthesis-relevant fragment library.
- Other fragment libraries also possible (e.g.: e-molecules, WuXi).



[1] Protopopov MV, Tararina VV, Bonachera F, et al. The freedom space – a new set of commercially available molecules for hit discovery. *Molecular Informatics*. 2024;43(12):e202400114. doi:10.1002/minf.202400114

SynthonForgePlus

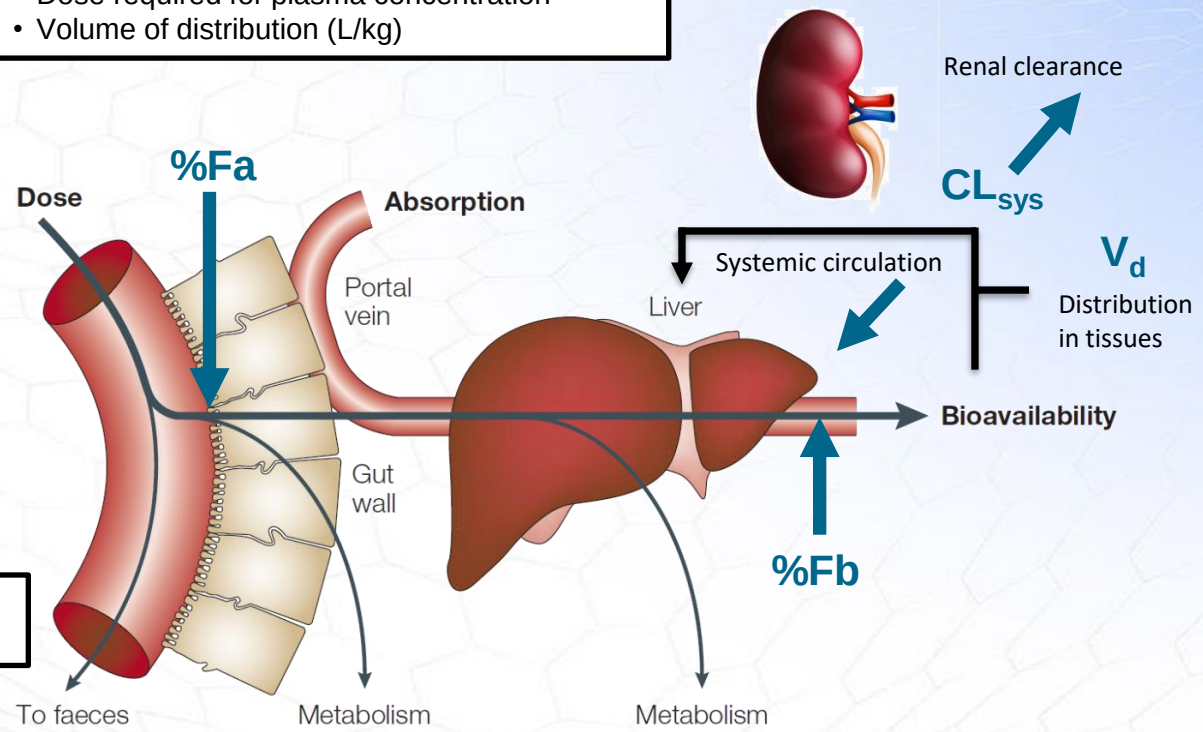
- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Combines:
  - Synthon-based construction/orchestration
  - ADMET Predictor® (via pyADMETPredictor)
  - MCDA (Multicriteria Decision Analysis)
  - Integration with HTPK for rapid PK modeling
- Enables dynamic, goal-oriented exploration of fragment chemical space.



HTPK profiling

- SynthonForgePlus integrates ADMET Predictor to assess key PK endpoints.
- Enables early-stage selection with predictive modeling:
  - Clearance
  - Fraction bioavailable
  - Cp(t) profiles
  - C_{min}, C_{max}
  - AUC
  - $T_{1/2}$
  - etc.

Predicts:
- Human, rat, mouse, dog, and monkey species
- IR tablet, IR solution, and IV bolus
- Percent absorbed (aka HIA)
- Oral bioavailability (aka F%)
- Dose required for plasma concentration
- Volume of distribution (L/kg)



ACAT™ model is used to simulate absorption

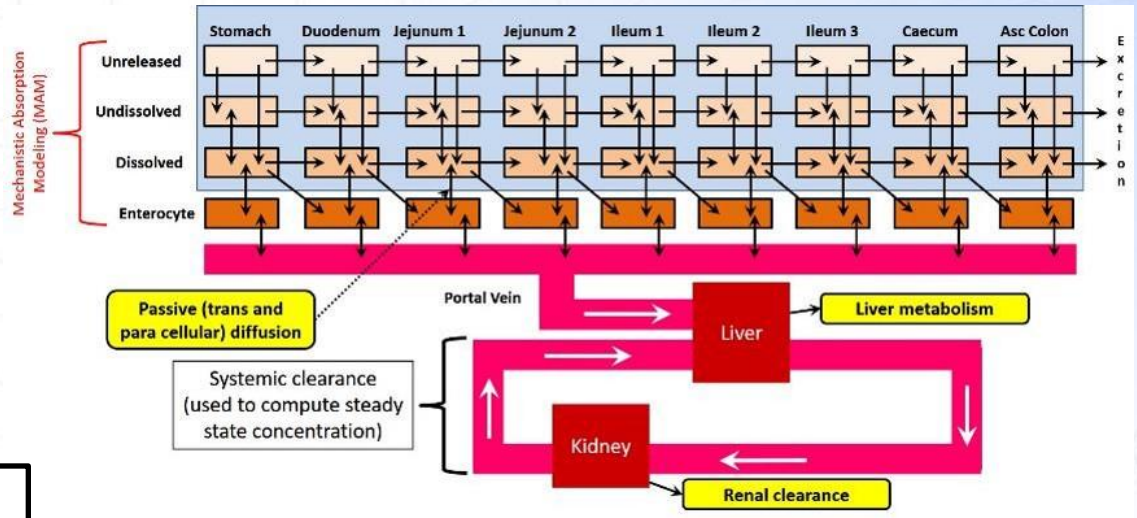
Gut clearance and active transport (efflux/influx) are **not** considered

Modified from van de Waterbeemd, H, and Gifford, E. Nat. Rev. Drug Disc. 2003, 2:192-204

HTPK profiling

- SynthonForgePlus integrates ADMET Predictor to assess key PK endpoints.
- Enables early-stage selection with predictive modeling:
  - Clearance
  - Fraction bioavailable
  - Cp(t) profiles
  - C_{min}, C_{max}
  - AUC
  - $T_{1/2}$
  - etc.

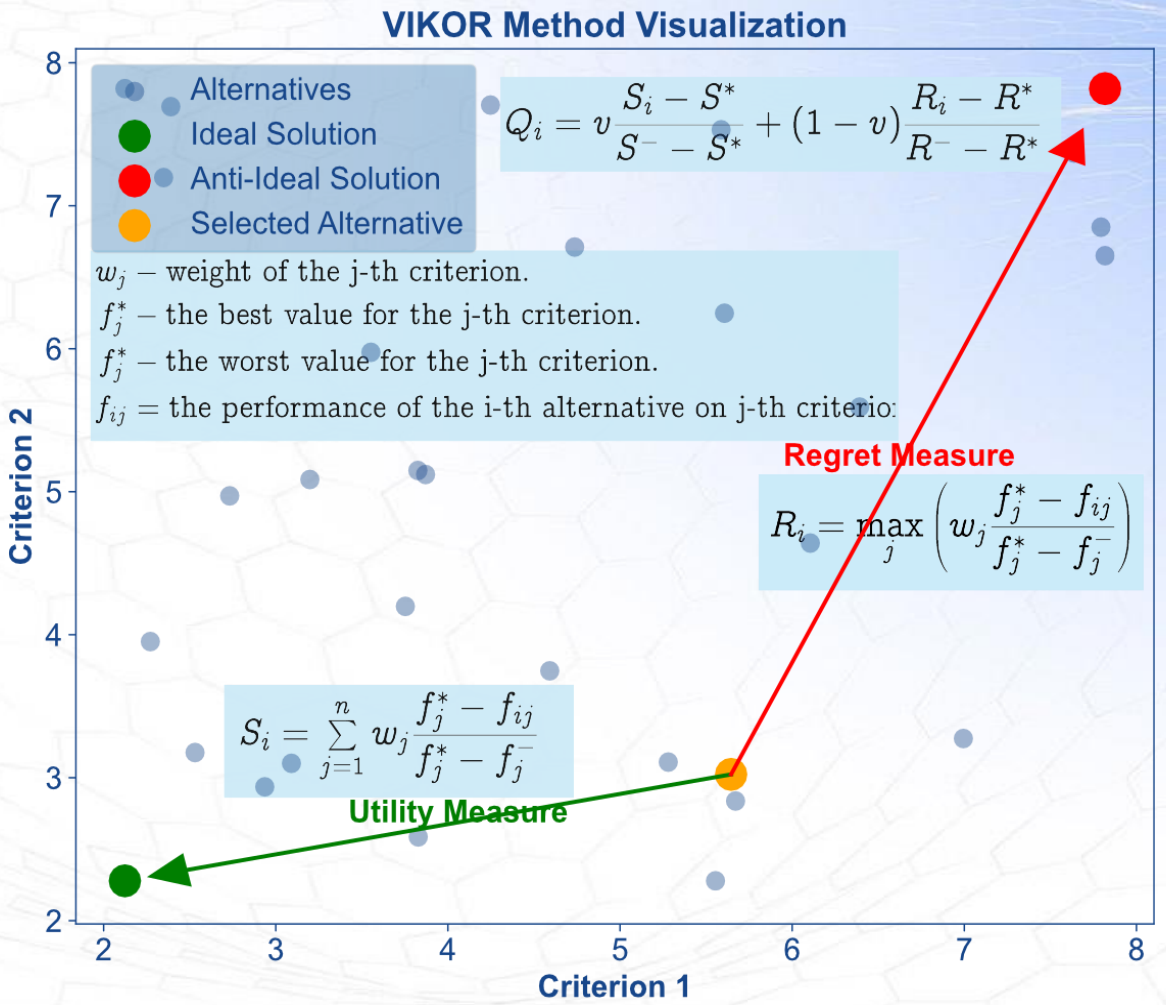
Predicts:
- Human, rat, mouse, dog, and monkey species
- IR tablet, IR solution, and IV bolus
- Percent absorbed (aka HIA)
- Oral bioavailability (aka F%)
- Dose required for plasma concentration
- Volume of distribution (L/kg)



ACAT™ model is used to simulate absorption

MCDA: VIKOR method

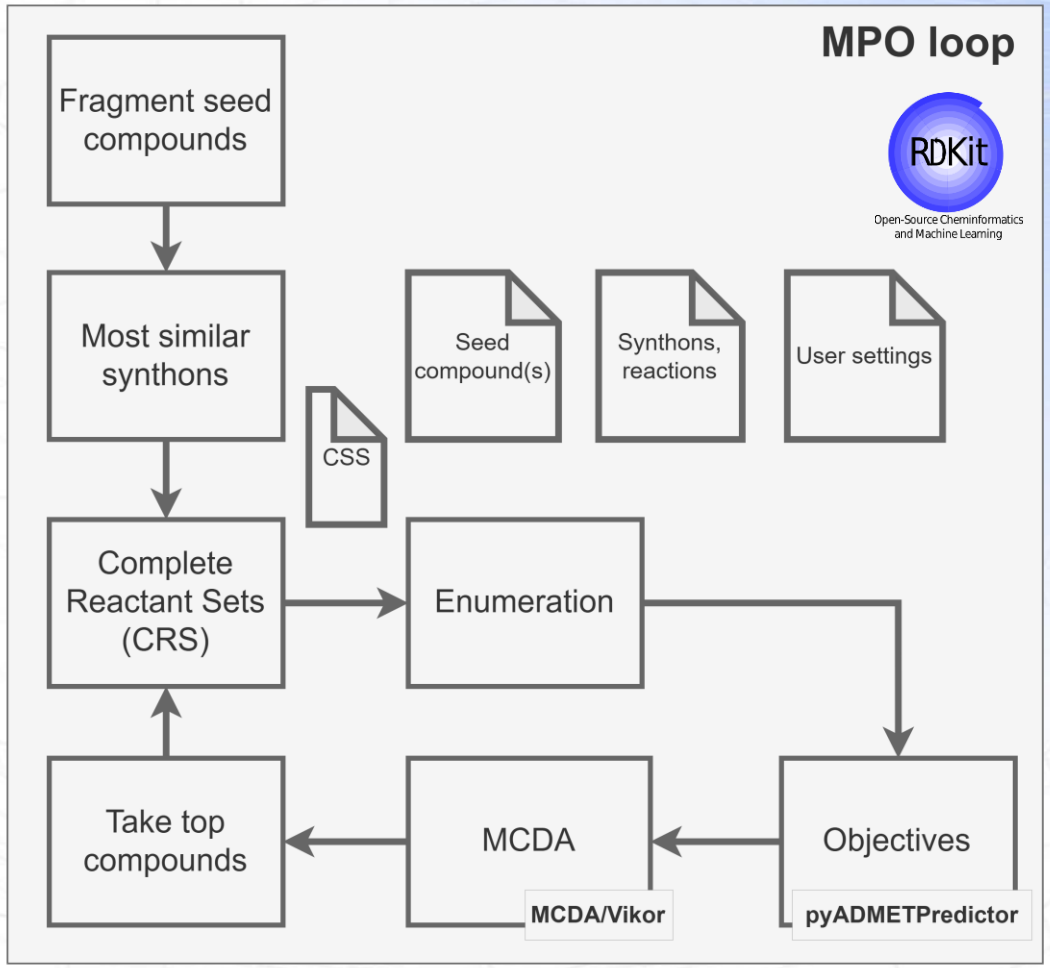
- VIKOR method (VlseKriterijumska Optimizacija I Kompromisno Resenje).
- Multi-criteria decision analysis (MCDA) technique.
- Ranking from among alternatives in the presence of conflicting criteria.
- Regret and Utility measures.
- Each alternative ranked individually.
- Incorporation of user preferences.
- Extensively used in **SynthonForgePlus**.



SynthonForgePlus

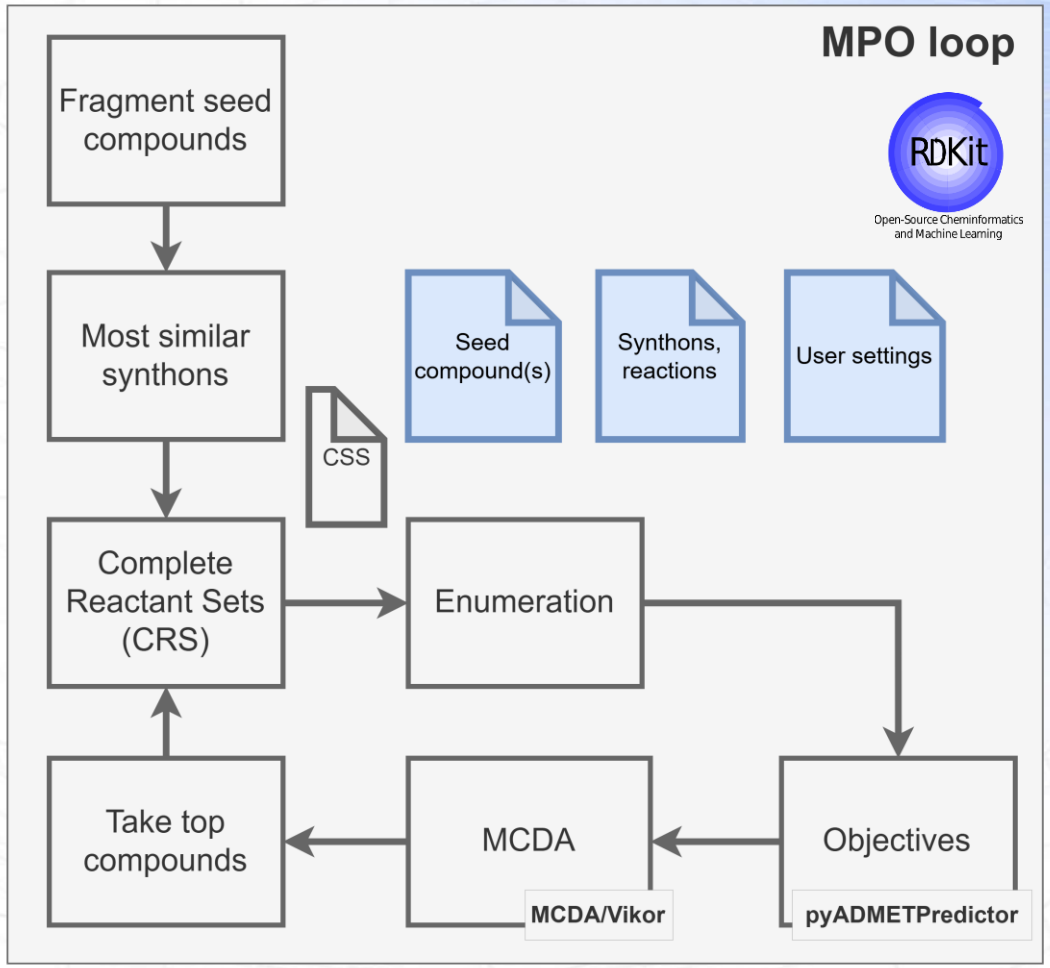
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- General diagram of Multiparameter Optimization (MPO) Loop.
- We will go step-by-step.



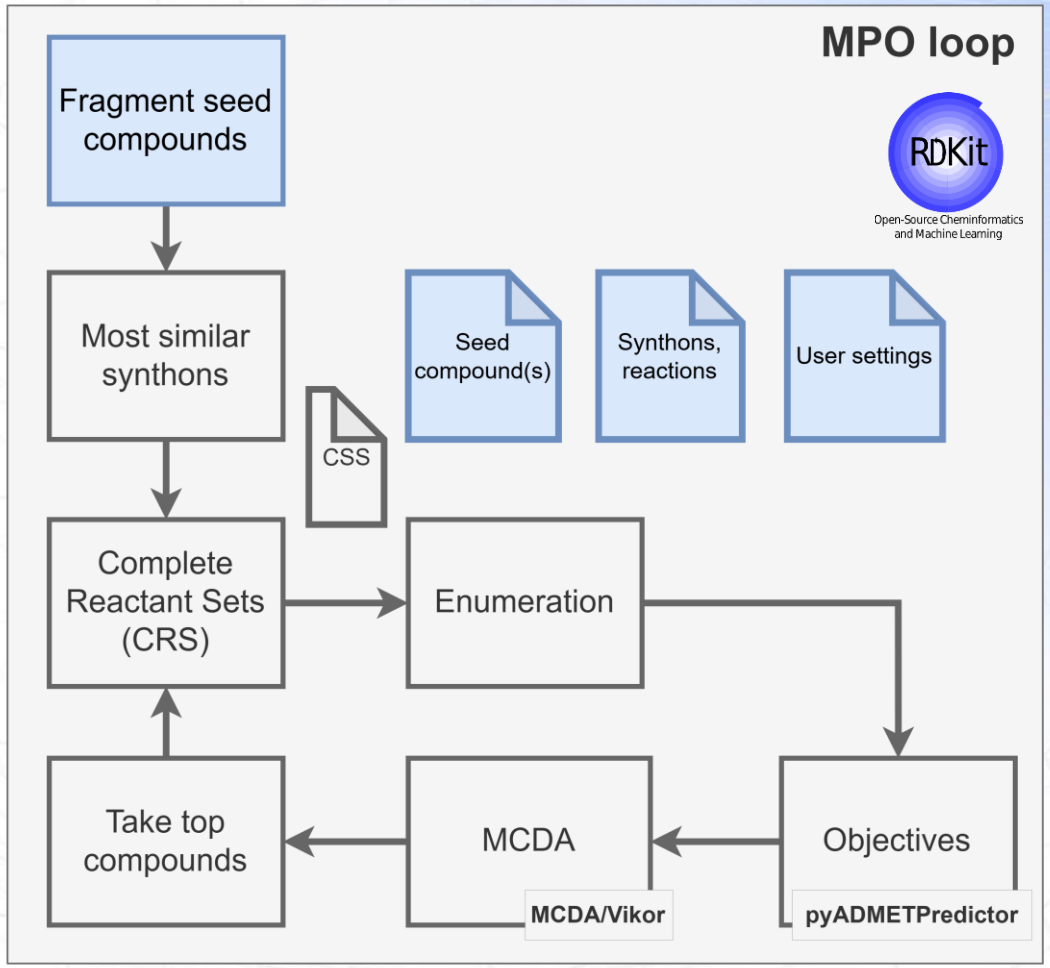
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- User inputs:
  - Seed compounds
  - Definition of synthon space (here Freedom Space 3.0 from Chemspace)
  - Other settings



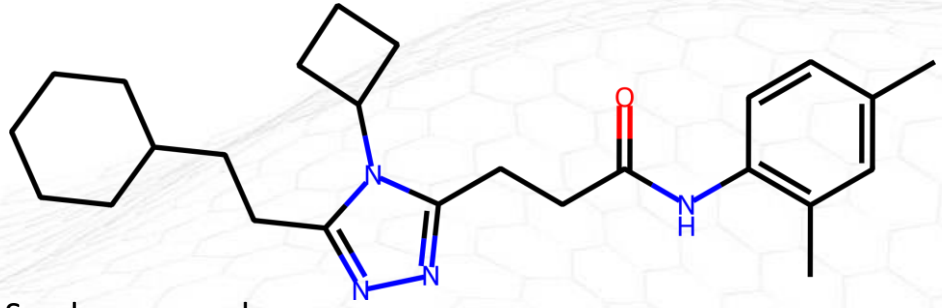
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Compound fragmentation.
- BRICS (**B**reaking of **R**etrosynthetically **I**nteresting **C**hemical **S**ubstructures)[1] or exhaustive approach.
- The latter: fragmenting the compound into many possible pieces.
- Neglecting small, not-relevant fragments.



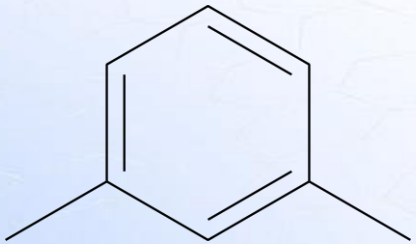
[1] Degen J, Wegscheid-Gerlach C, Zaliani A, Rarey M. On the Art of Compiling and Using "Drug-Like" Chemical Fragment Spaces. *ChemMedChem*. 2008;3(10):1503-1507. doi:10.1002/cmdc.200800178

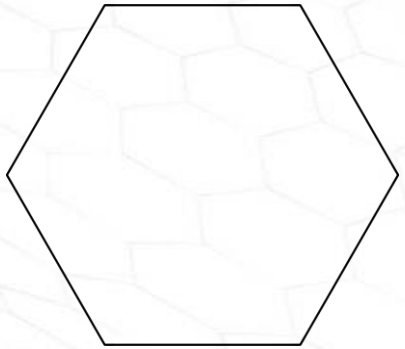
SynthonForgePlus

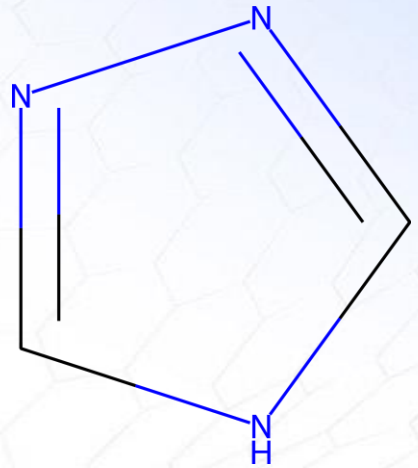


Seed compound

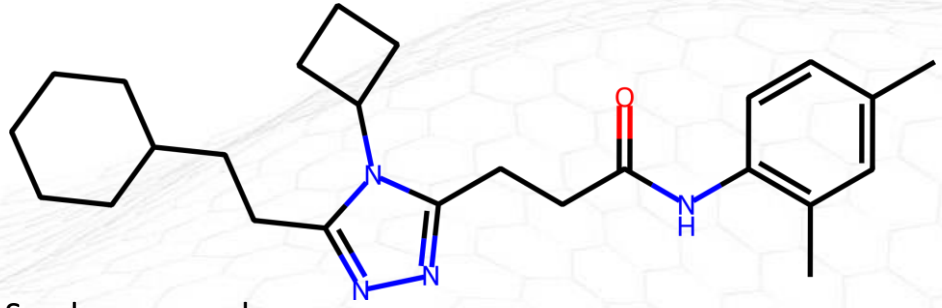
BRICS fragmentation





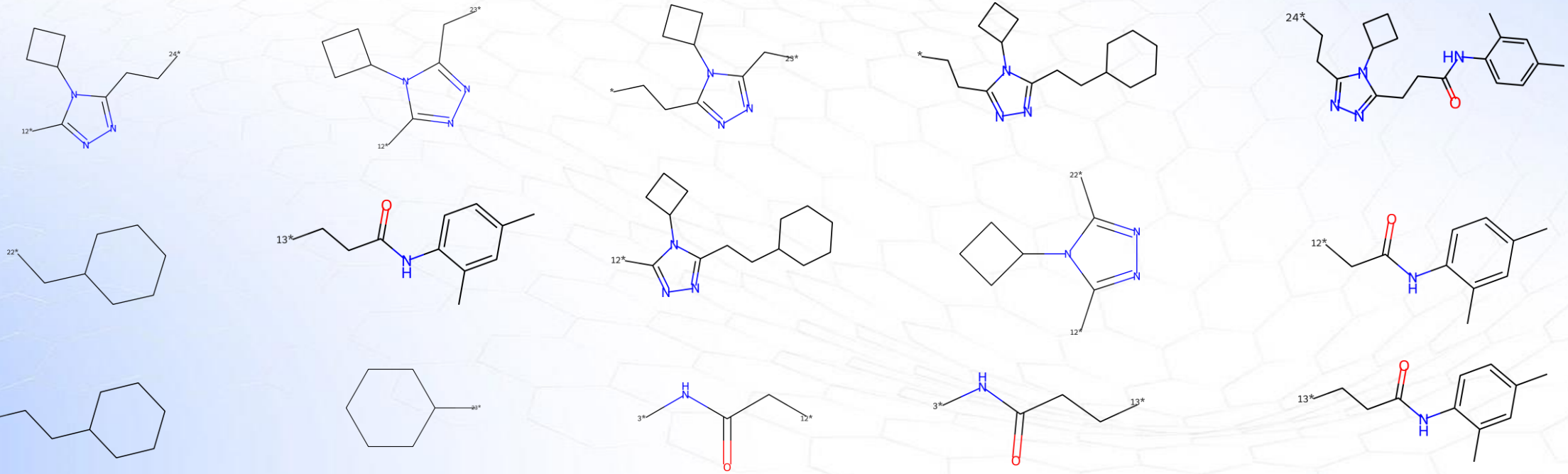


SynthonForgePlus



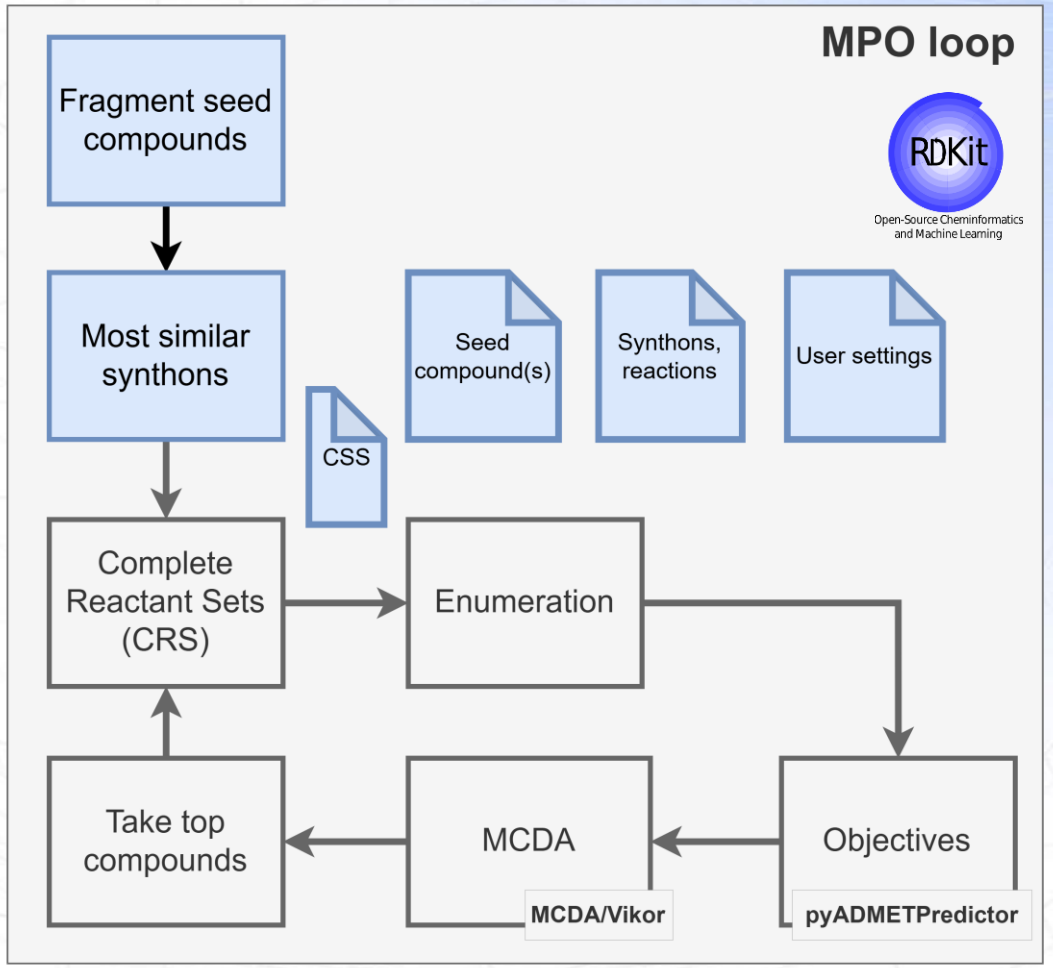
Seed compound

Exhaustive fragmentation
(selected fragments)



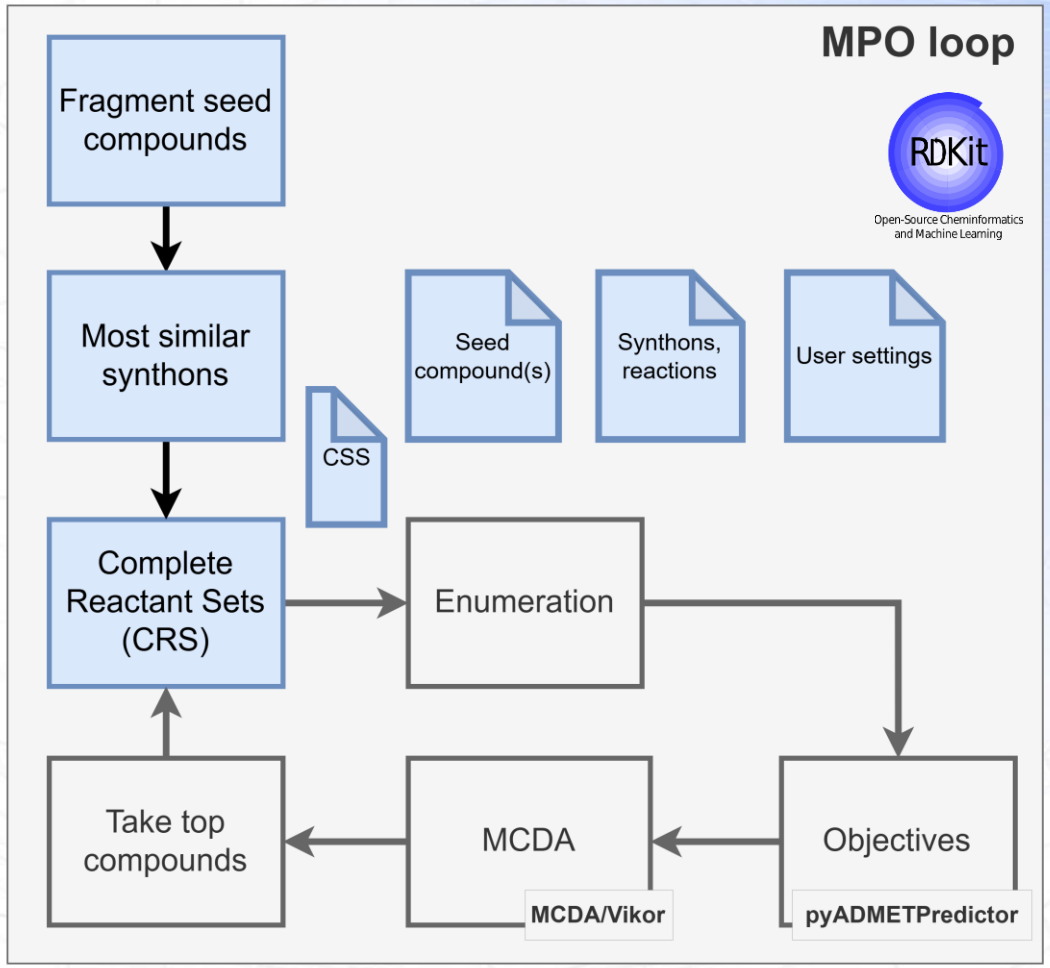
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Finding most similar synthons.
- Morgan fingerprints (FPs) with FPSim2.
- Form **Current Synthons Set** (CSS).
- CSS: key artifact in the optimization loop.



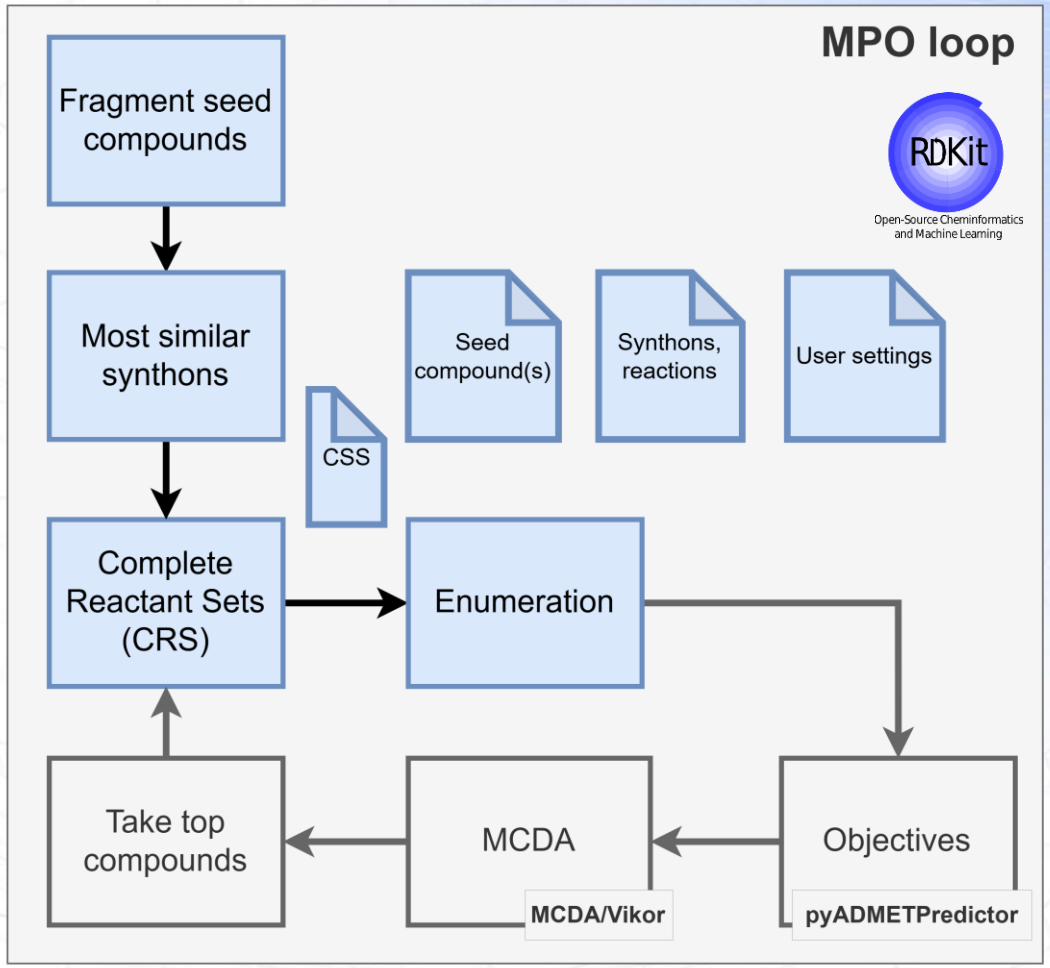
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- CSS is the object which is iteratively refined.
- Random selection of reactions and complementary synthons.
- Form **Complete Reactant Sets**.



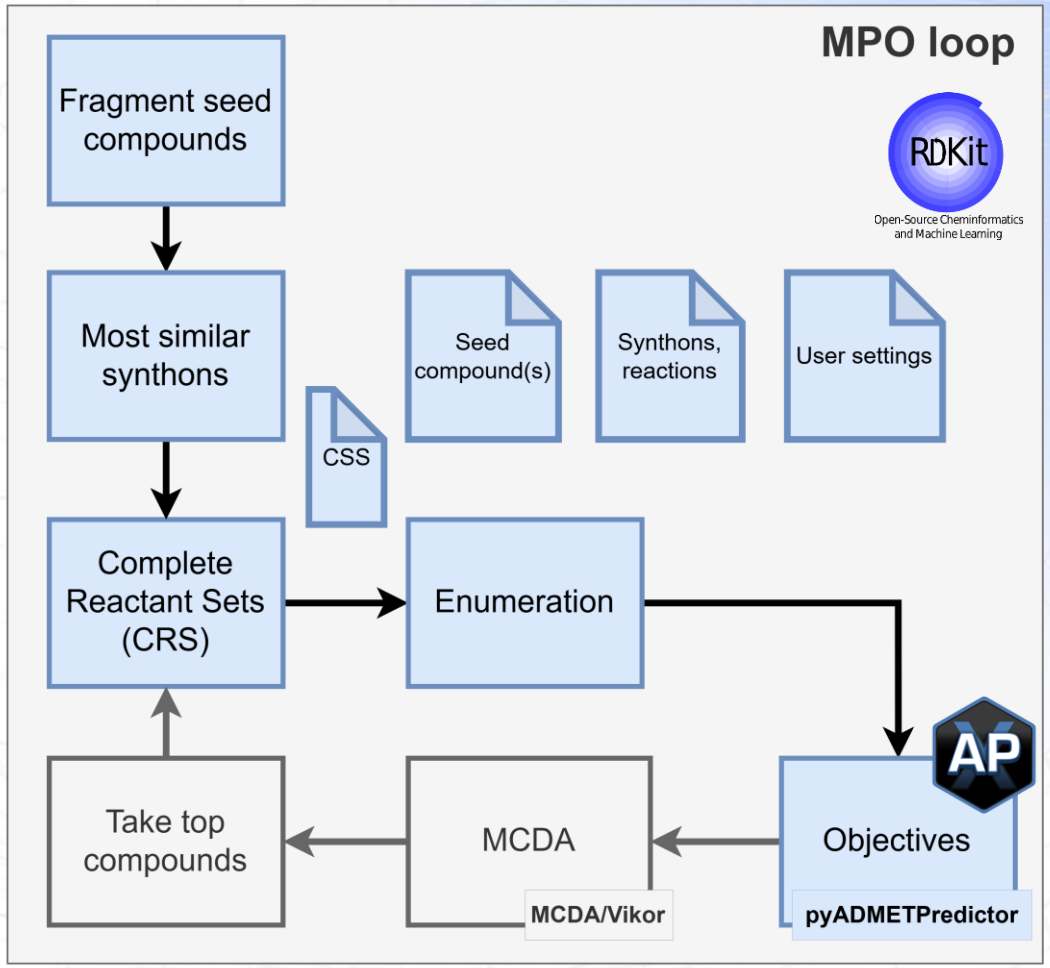
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Enumeration.
- Turning Complete Reactant Sets (CRS) into compounds.
- Current Candidates.
- Needed only for quality estimation.



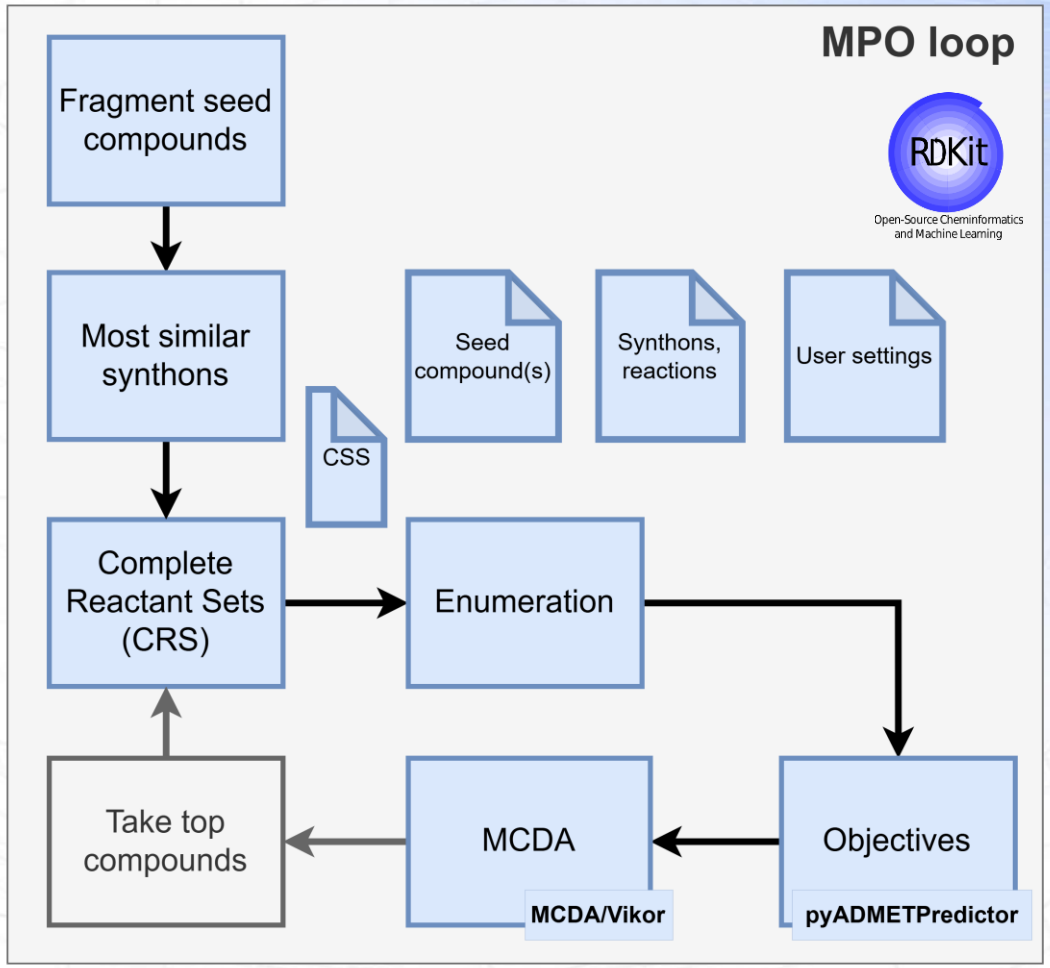
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Application of ADMET Predictor and HTPK to get the properties.
- Other objectives also possible (e.g. similarity to a reference compound).



SynthonForgePlus

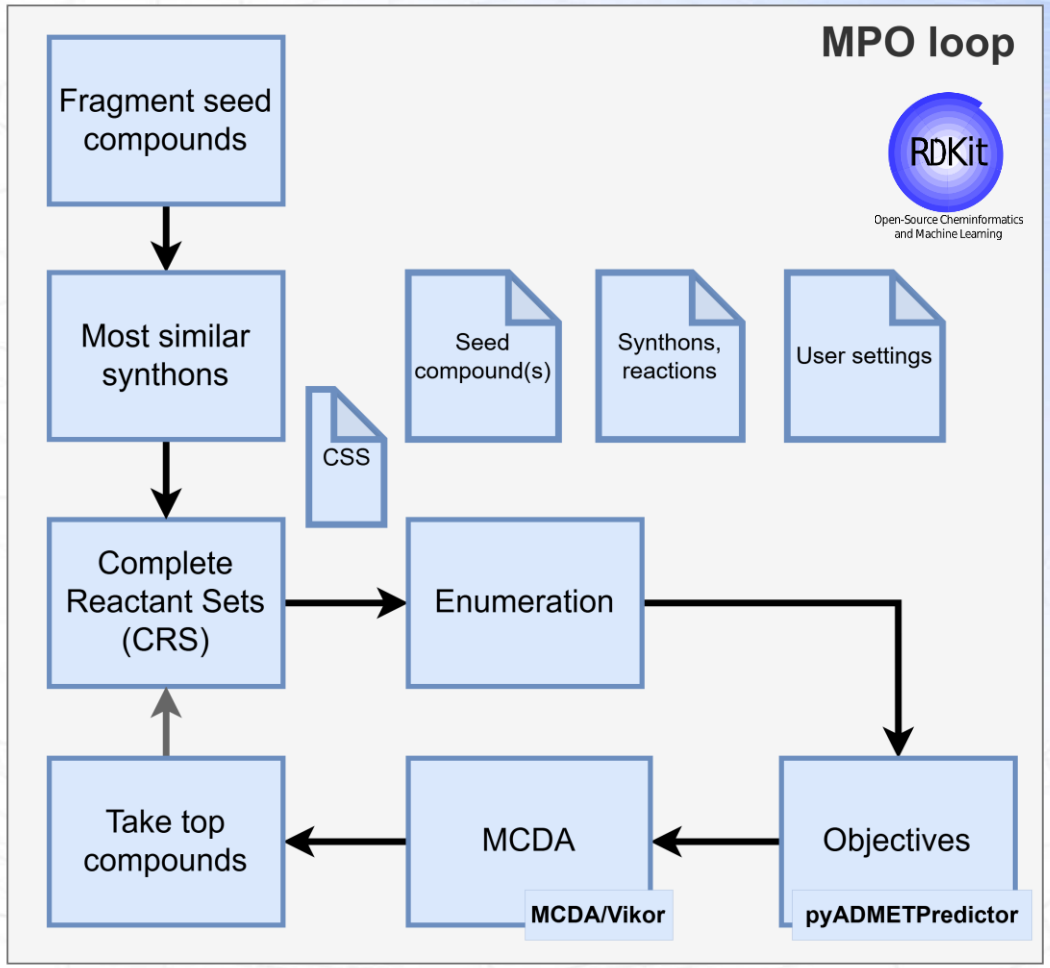
- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- MCDA: Multicriteria Decision Analysis/Aiding.
- Application of MCDA Vikor method in drug discovery[1].
- Global utility function incorporating all objectives (with user weights).
- Ranking, best compounds on top.



[1] Bachorz RA, Lawless MS, Miller DW, Jones JO. Multi-Criteria Decision Analysis in Drug Discovery. *Applied Biosciences*. 2025;4(1):2. doi:10.3390/applbiosci4010002

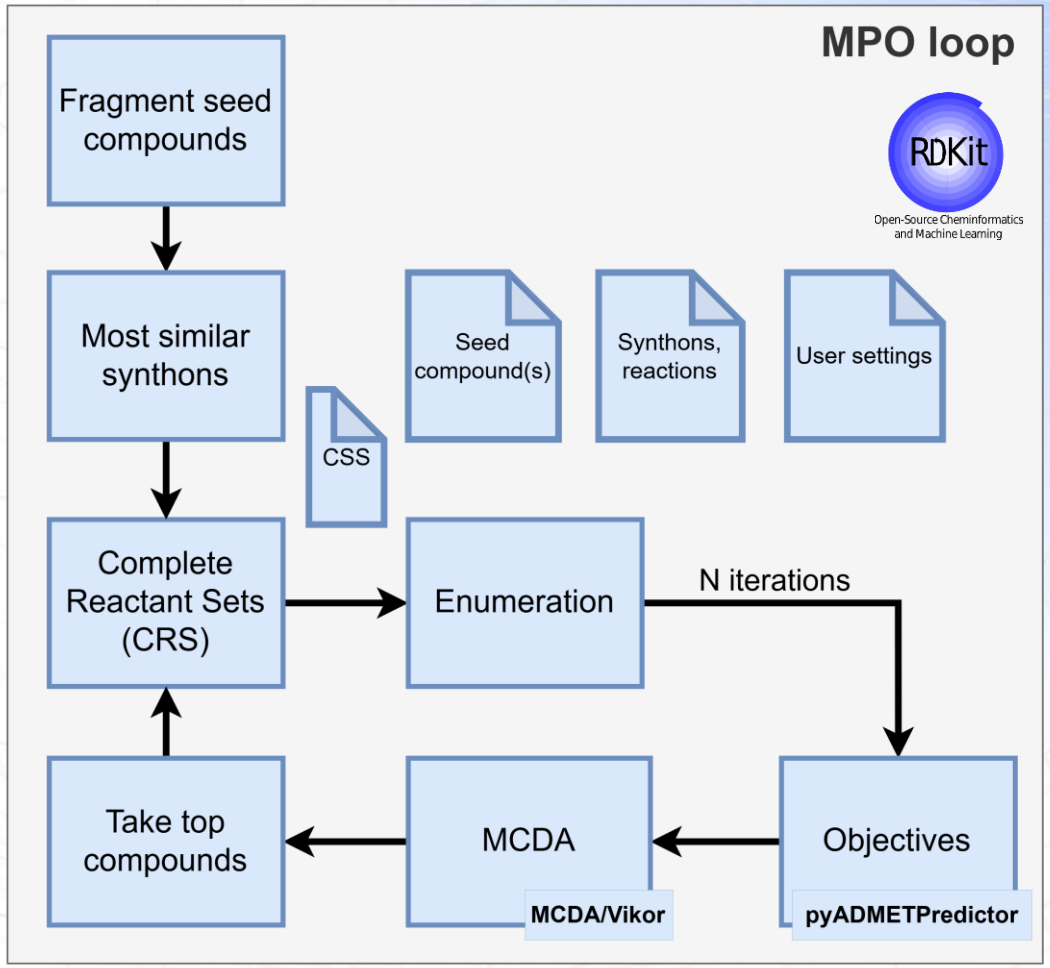
SynthonForgePlus

- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Pruning phase.
- The best compounds with highest MCDA ranks.
- The best compounds turned back into synthons.
- Form new CSS and are inherited to next iterations.



SynthonForgePlus

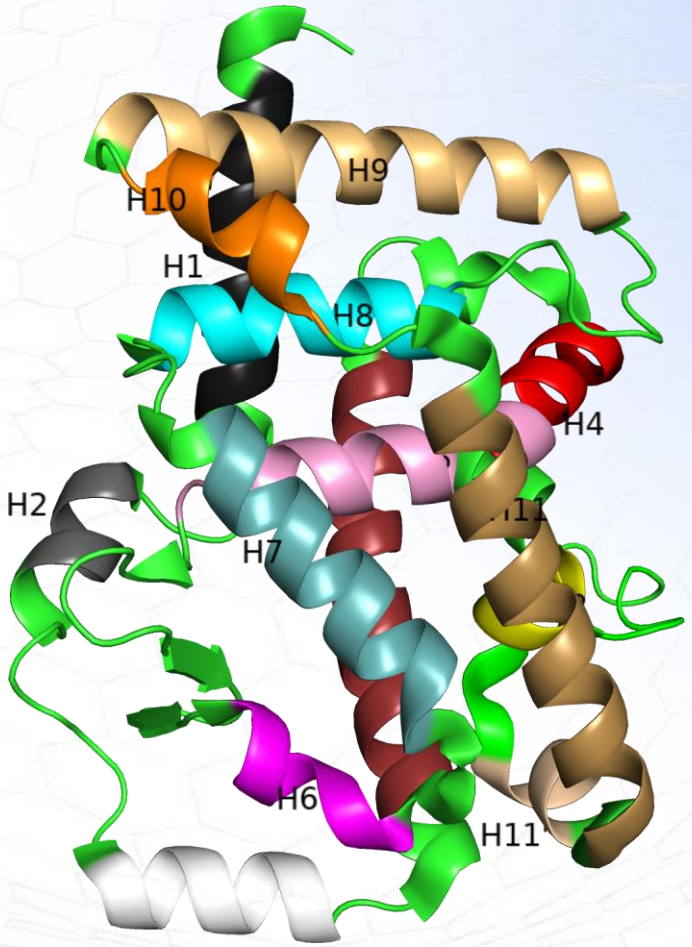
- **SynthonForgePlus**: A framework for property-driven and multi-objective molecular design.
- Closing the MPO loop.
- The optimization process continues till convergence criteria are met:
  - Predefined number of iterations.
  - CSS does not change significantly.
  - Calculation time.
- Results provided as MCDA-sorted compounds.



Use cases

Use case: RORγ agonist

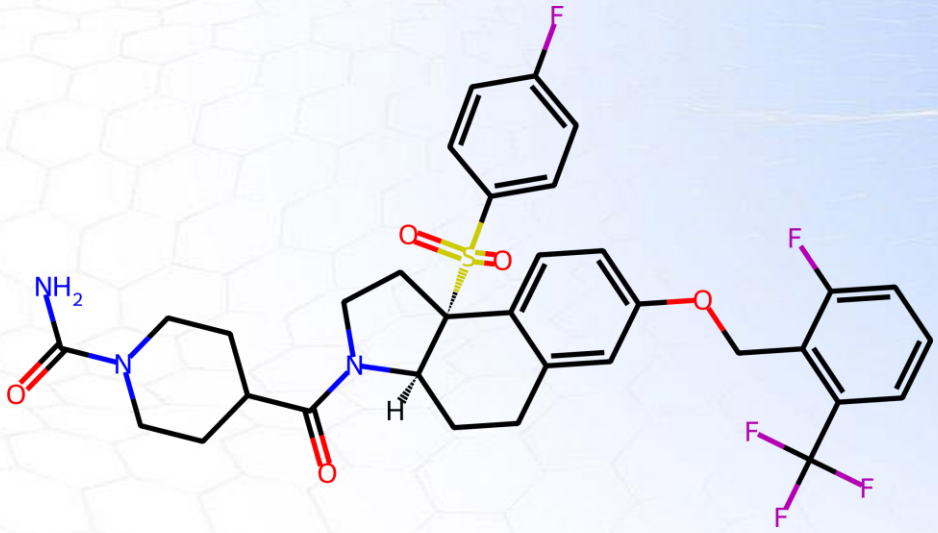
- **Retinoic acid receptor**-related orphan receptor gamma (RORγ) is a nuclear receptor that regulates:
  - Differentiation and function of Th17 cells.
  - Expression of IL-17 cytokines, central to pathogenesis of psoriasis, rheumatoid arthritis, and Crohn's disease.
- RORγ agonists:
  - Cancer immunotherapy.
  - Autoimmune modulation.
- Recent experiences with the design of inverse agonists[1].



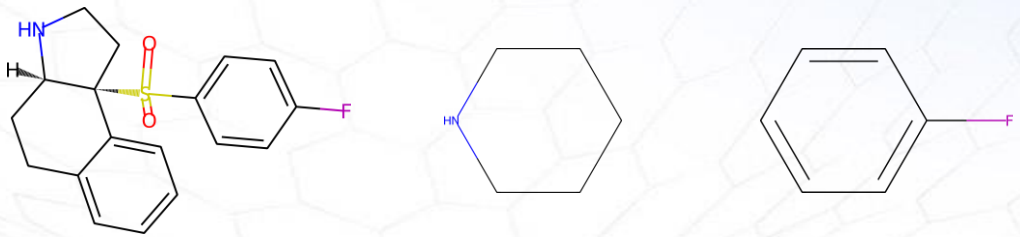
[1] Bachorz RA, Paswwińska J, MS, Miller DW, Sałkowska A, Karaś K, Karwaciak I, Jones JO, and Ratajewski M. Identification of a novel indolizine RORγT inverse agonist using the AI-driven Drug Design platform. Accepted in *ACS Med. Chem. Lett.*

Use case: RORγ agonist

- **Setup:**
  - Seed: well-known strong agonist of RORγ (pEC50: 8.2).
  - Fragmenting procedure: BRICS
  - Objectives:
    - ADMET_Risk (0.25), direction: min
    - %Fb (0.25), direction: max
    - pEC50 (0.50), direction: max
  - HTPK setup:
    - Species: human
    - Route: IR Tablet
    - Dose: 1 mg/kg BW

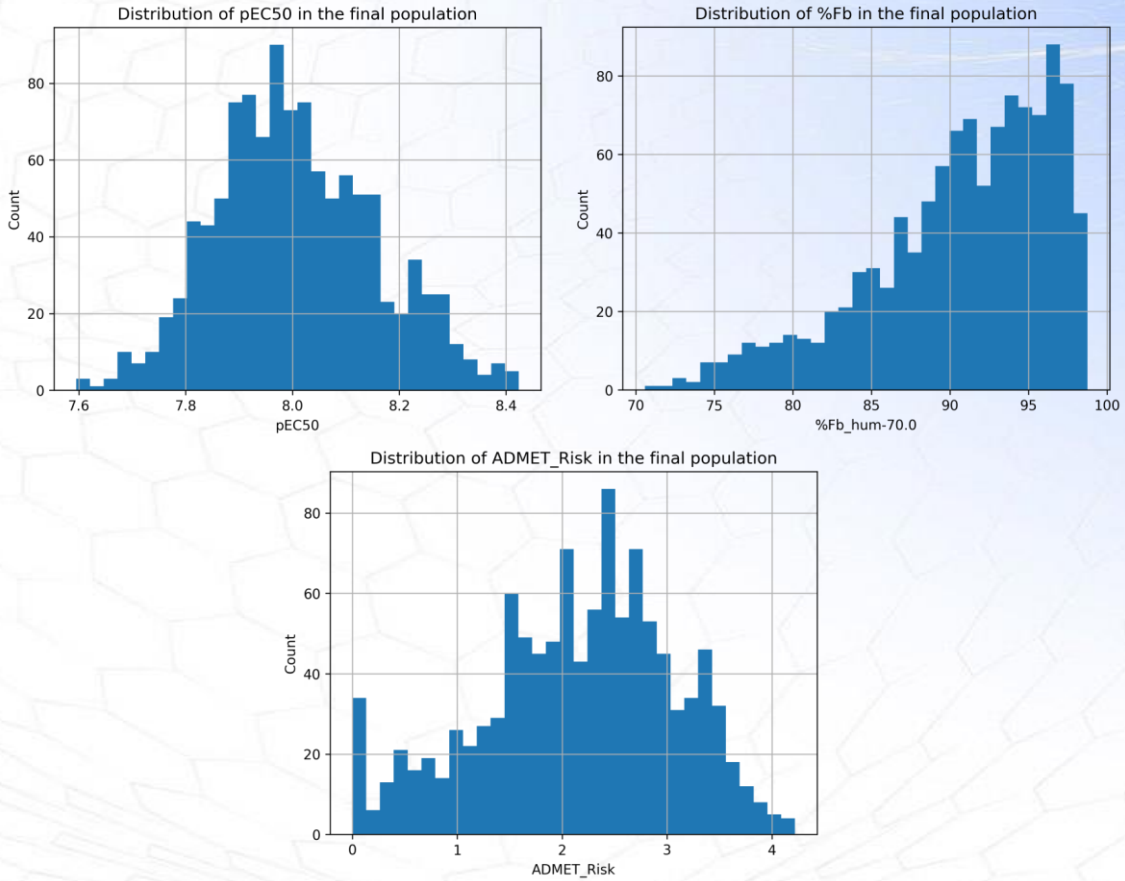


After BRICS fragmentation:



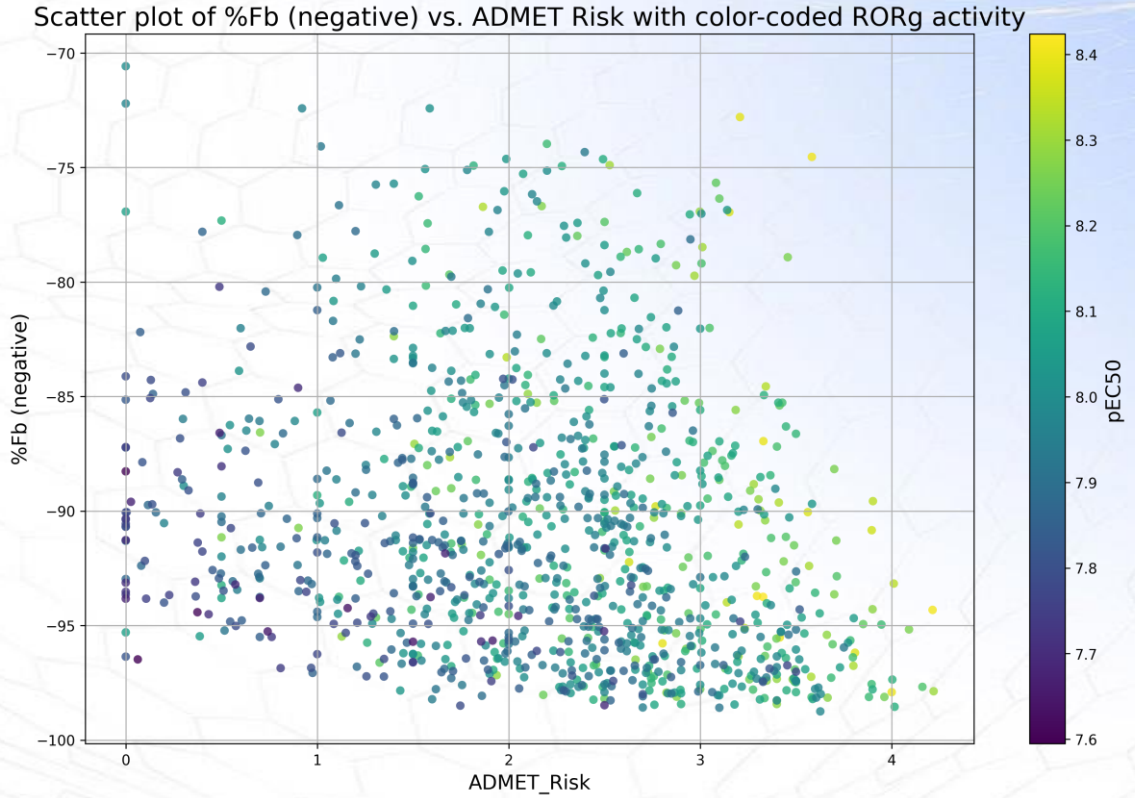
Use case: RORγ agonist

- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Pareto front (10 compounds).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



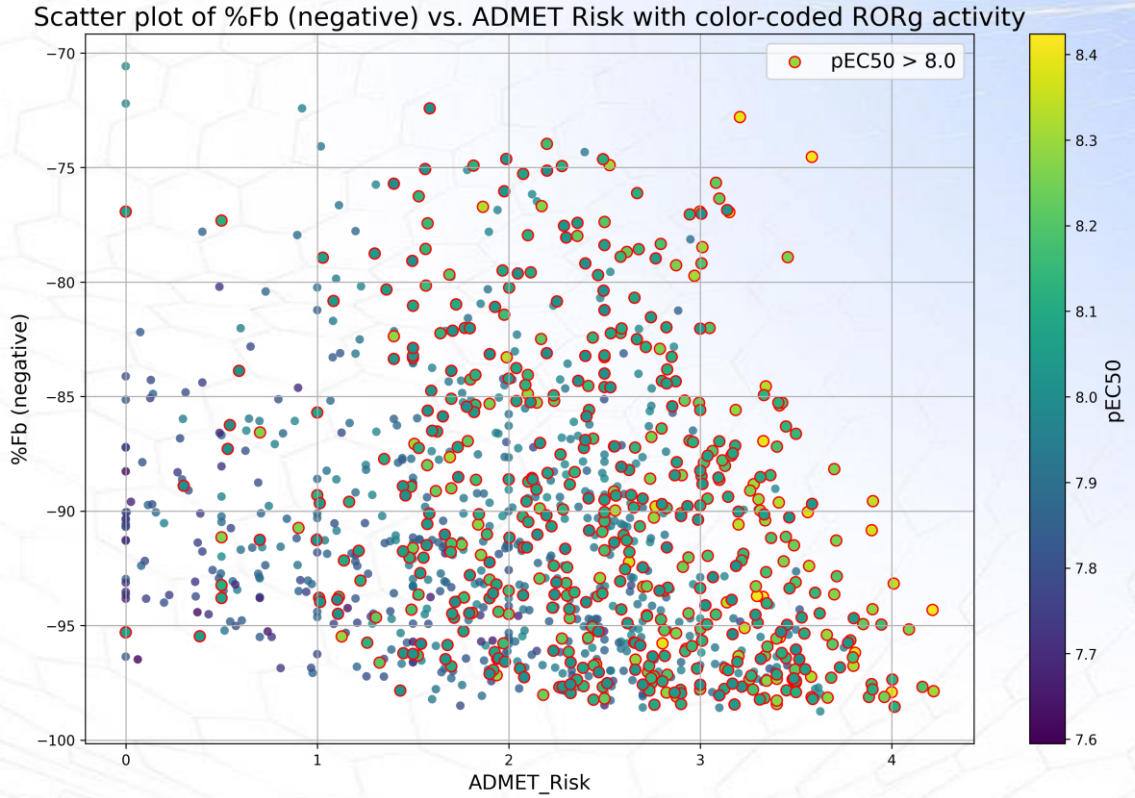
Use case: RORγ agonist

- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Top MCDA compounds.
  - Pareto front (10 compounds).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



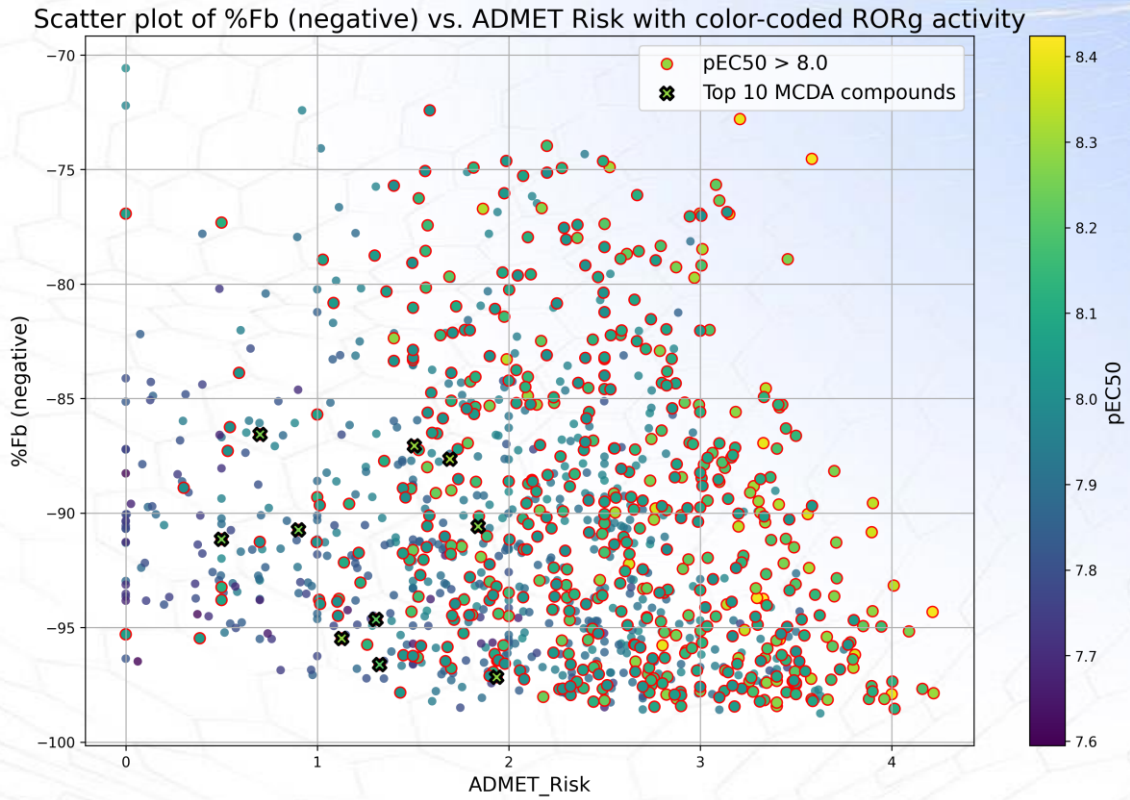
Use case: RORγ agonist

- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Top MCDA compounds.
  - Pareto front (10 compounds).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



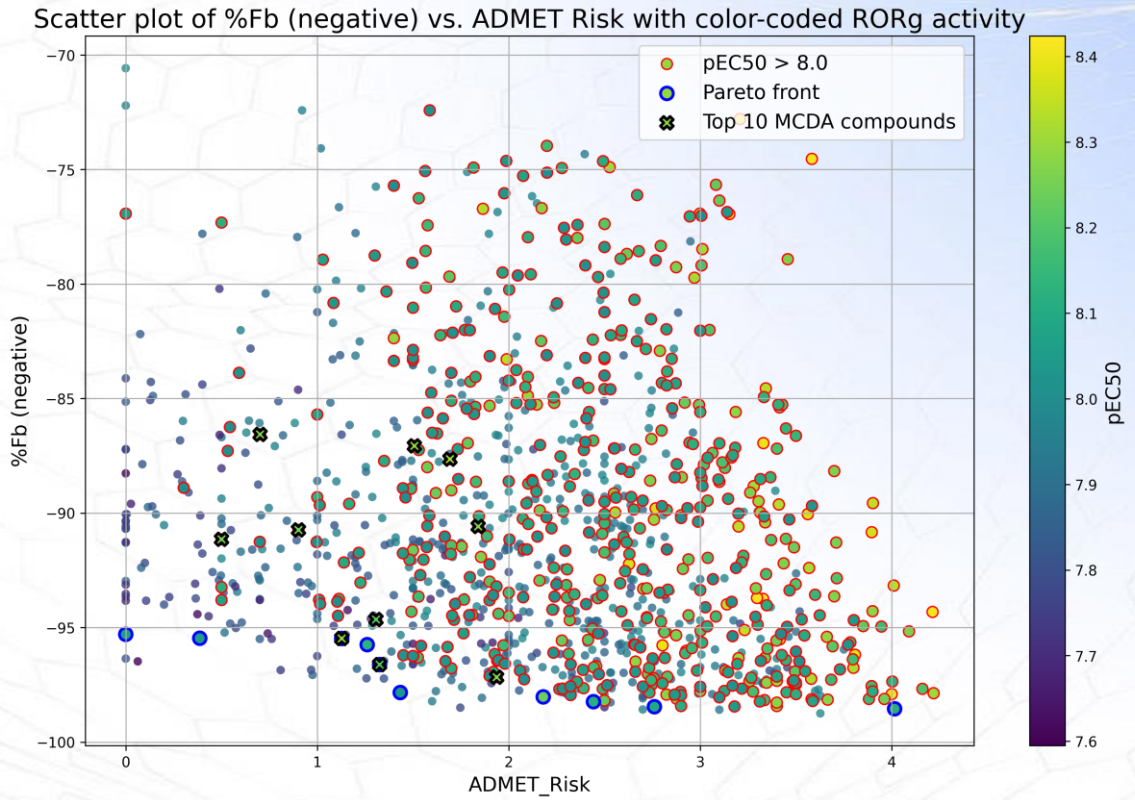
Use case: RORγ agonist

- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Top MCDA compounds.
  - Pareto front (10 compounds).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



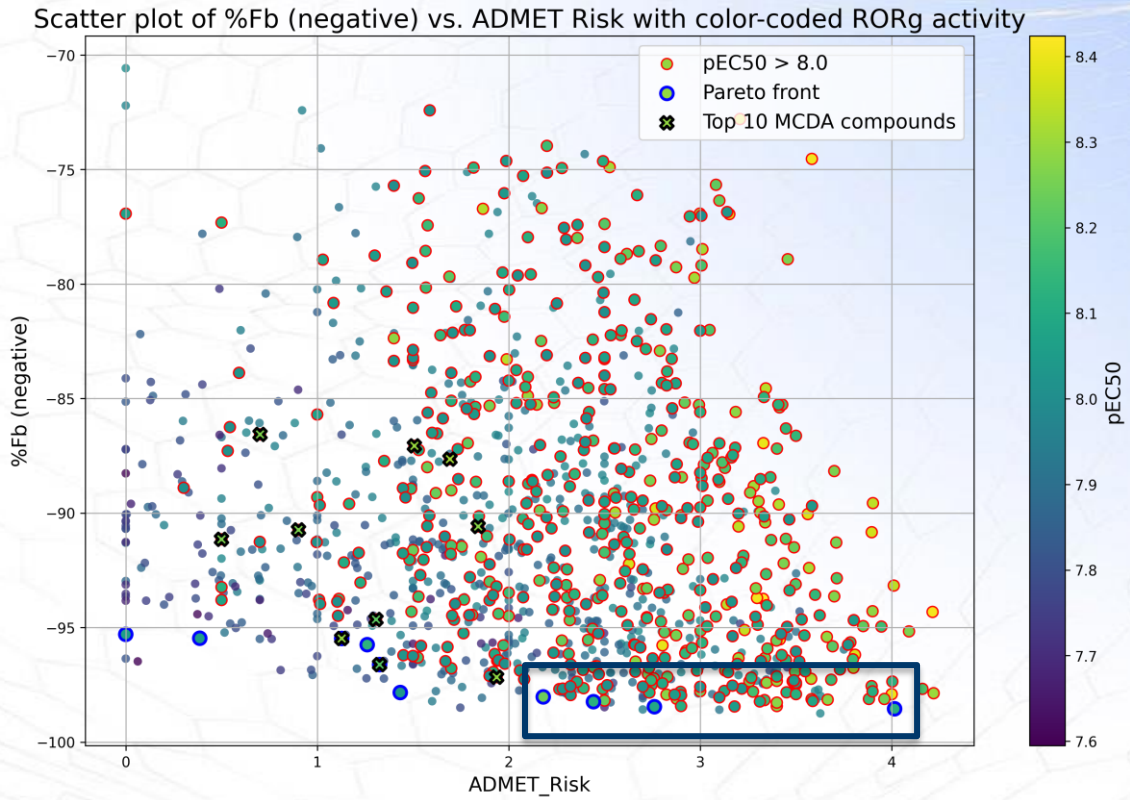
Use case: RORγ agonist

- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Top 10 MCDA compounds.
  - Pareto front compounds (10).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



Use case: RORγ agonist

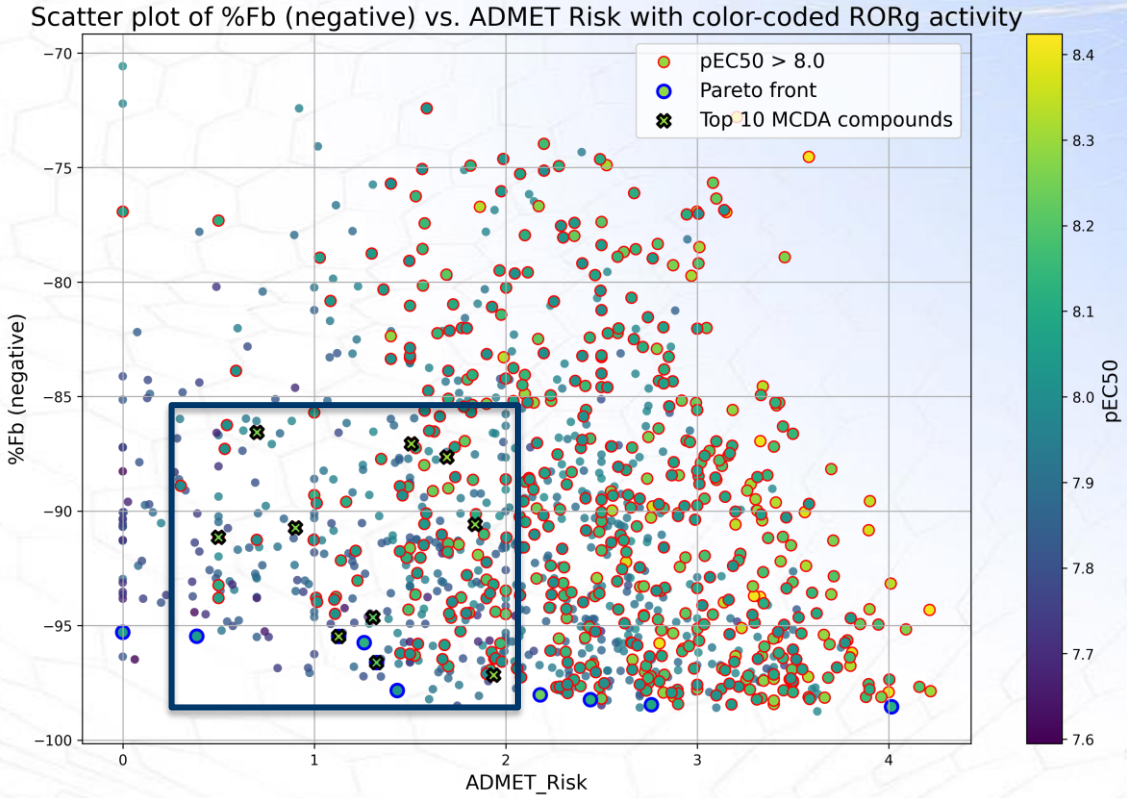
- **Results:**
  - The count of resulting population: 1098.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Top 10 MCDA compounds.
  - Pareto front compounds (10).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



Compounds on Pareto-front not always relevant, here due to high ADMET_Risk.

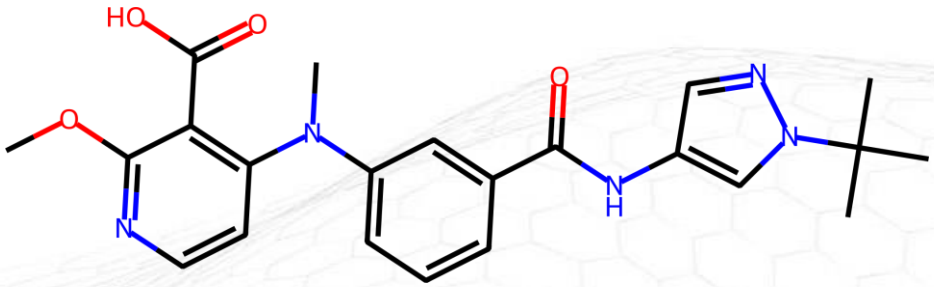
Use case: RORγ agonist

MCDA rank	ADMET_Risk	%Fb	pEC50
0	1.1	95.5	8.29
1	0.5	91.2	8.26
2	0.9	90.7	8.26
3	1.9	97.2	8.29
4	0.7	86.6	8.27
5	1.3	96.6	8.21
6	1.3	94.7	8.22
7	1.7	87.7	8.31
8	1.5	87.1	8.30
9	1.8	90.6	8.28

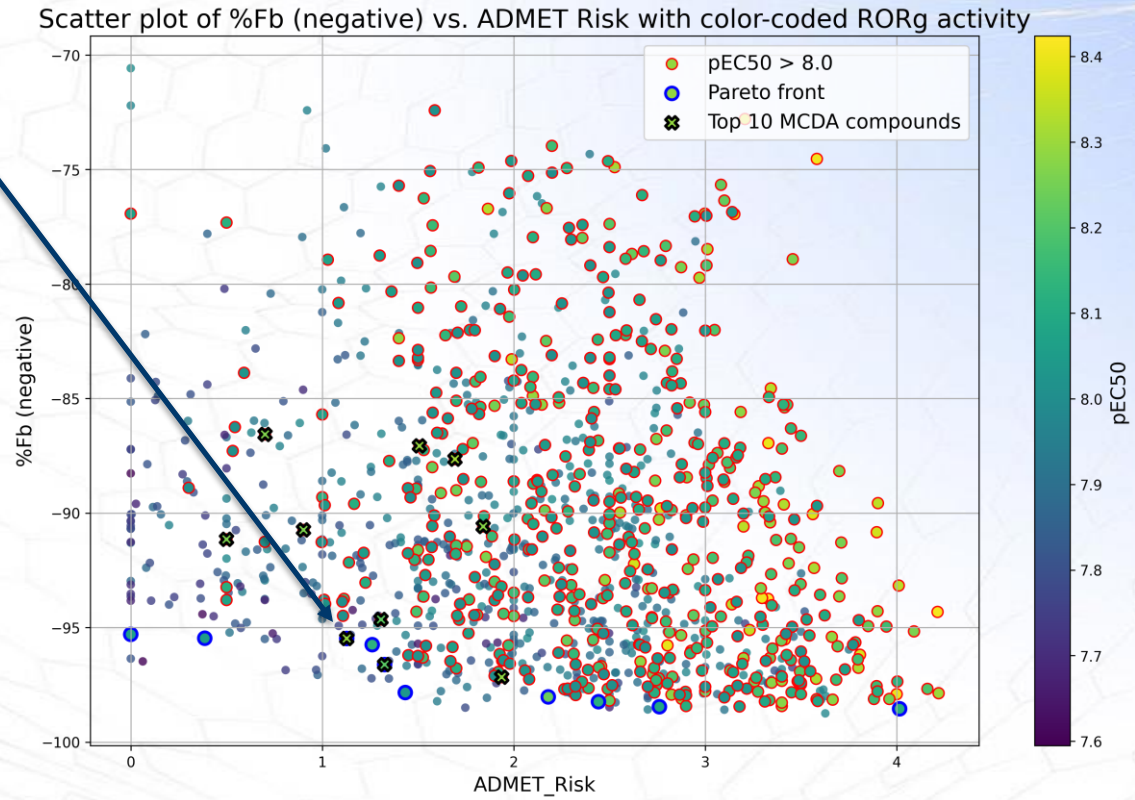


Therefore, we look at the top 10 MCDA compounds first.

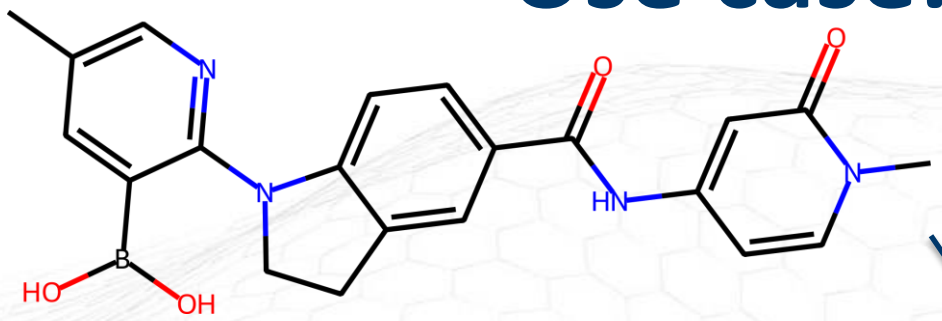
Use case: RORγ agonist



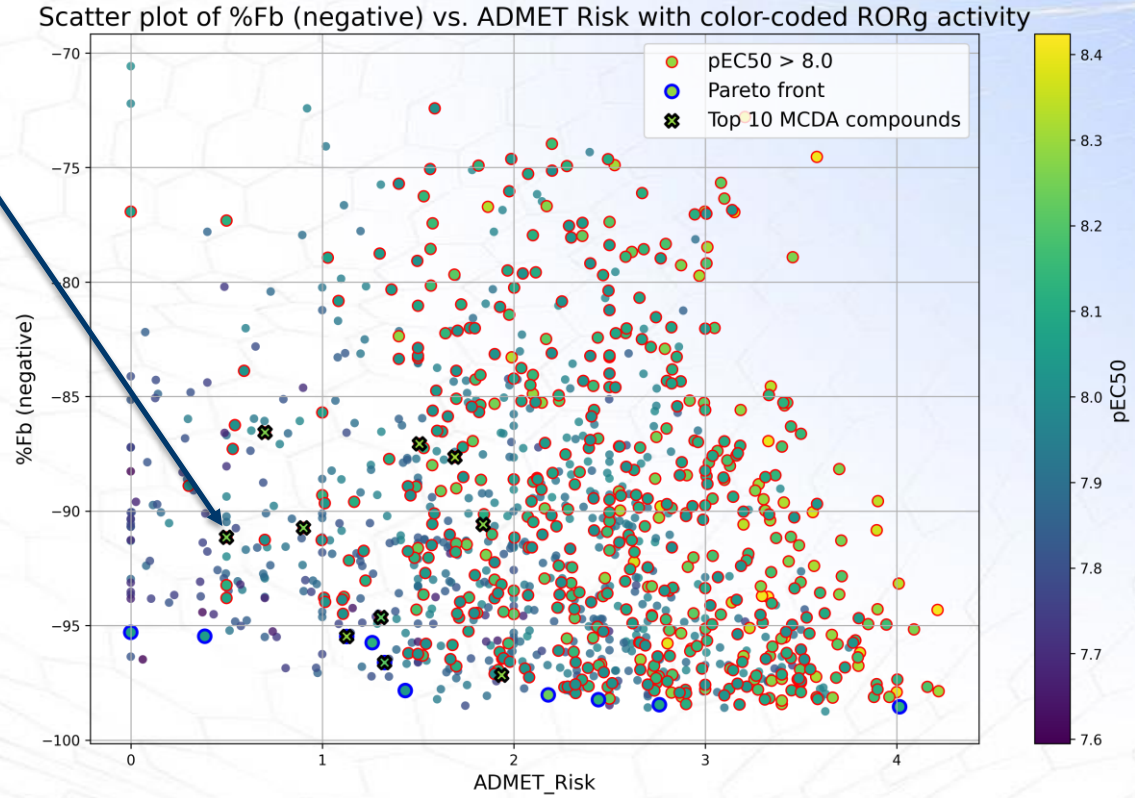
MCDA rank	ADMET_Risk	%Fb	pEC50
0	**1.1**	**95.5**	**8.29**
1	0.5	91.2	8.26
2	0.9	90.7	8.26
3	1.9	97.2	8.29
4	0.7	86.6	8.27
5	1.3	96.6	8.21
6	1.3	94.7	8.22
7	1.7	87.7	8.31
8	1.5	87.1	8.30
9	1.8	90.6	8.28



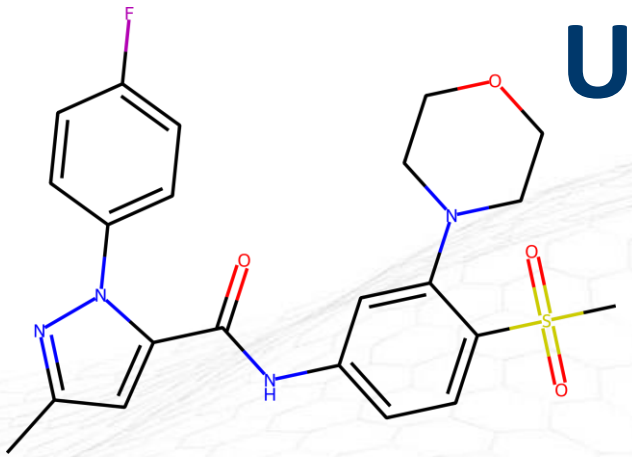
Use case: RORγ agonist



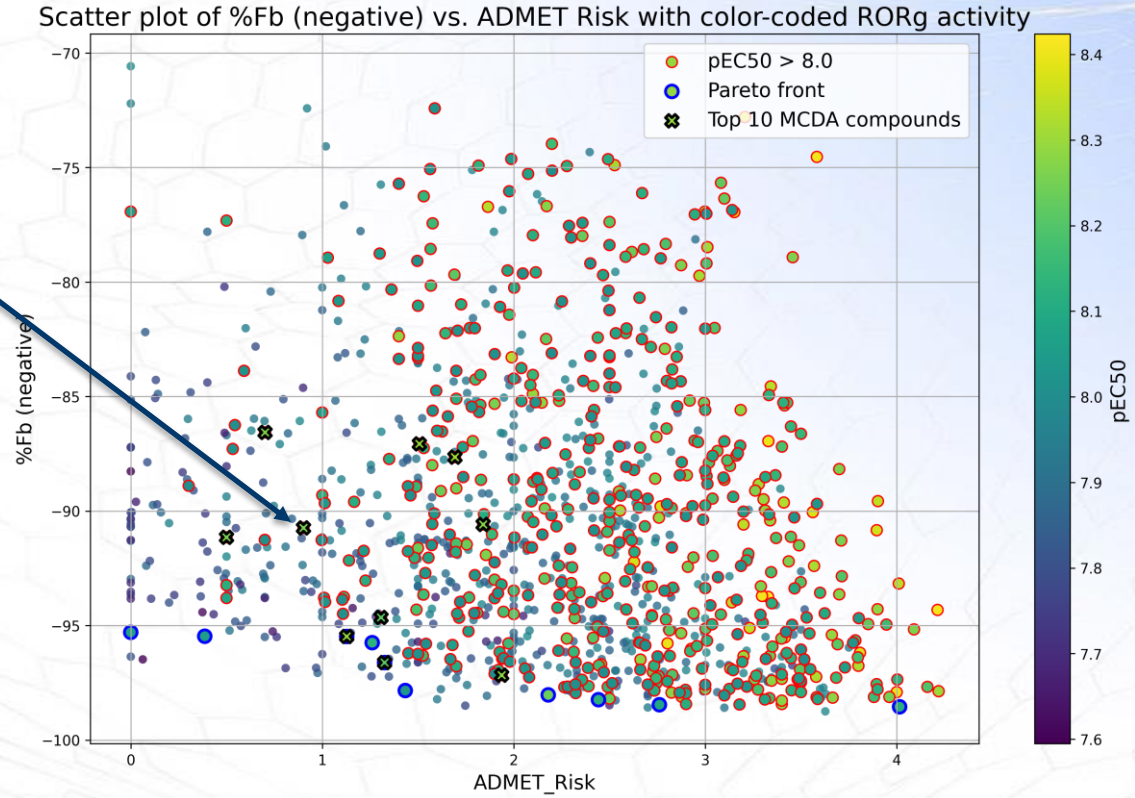
MCDA rank	ADMET_Risk	%Fb	pEC50
0	1.1	95.5	8.29
1	**0.5**	**91.2**	**8.26**
2	0.9	90.7	8.26
3	1.9	97.2	8.29
4	0.7	86.6	8.27
5	1.3	96.6	8.21
6	1.3	94.7	8.22
7	1.7	87.7	8.31
8	1.5	87.1	8.30
9	1.8	90.6	8.28



Use case: RORγ agonist

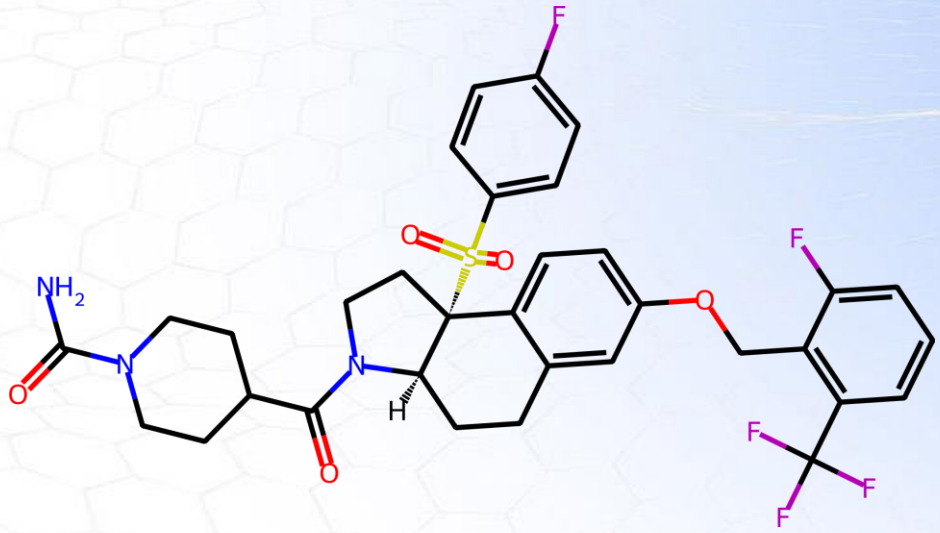


MCDA rank	ADMET_Risk	%Fb	pEC50
0	1.1	95.5	8.29
1	0.5	91.2	8.26
2	**0.9**	**90.7**	**8.26**
3	1.9	97.2	8.29
4	0.7	86.6	8.27
5	1.3	96.6	8.21
6	1.3	94.7	8.22
7	1.7	87.7	8.31
8	1.5	87.1	8.30
9	1.8	90.6	8.28

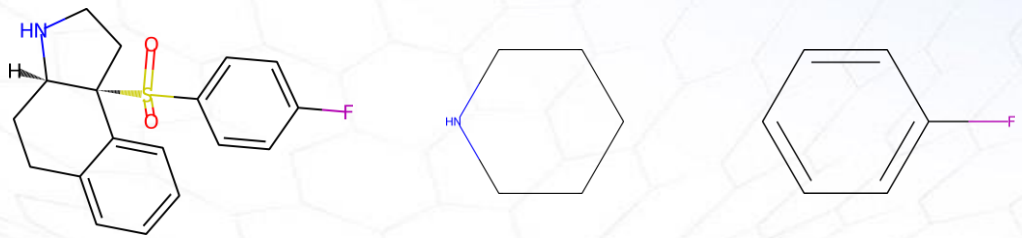


Use case: targeted RORγ agonist optimization

- **Setup:**
  - Seed: well-known strong agonist of RORγ (pEC50: 8.2).
  - Fragmenting procedure: BRICS
  - Sometimes unconstrained minimizing/maximizing is not best approach.
  - Objectives:
    - ADMET_Risk (0.10), direction: min
    - Clearance (0.35), target: 10 L/h
    - Fup (0.15), target: 55%
    - pEC50 (0.40), direction: max
  - HTPK setup:
    - Species: human
    - Route: IR Tablet
    - Dose: 1 mg/kg BW

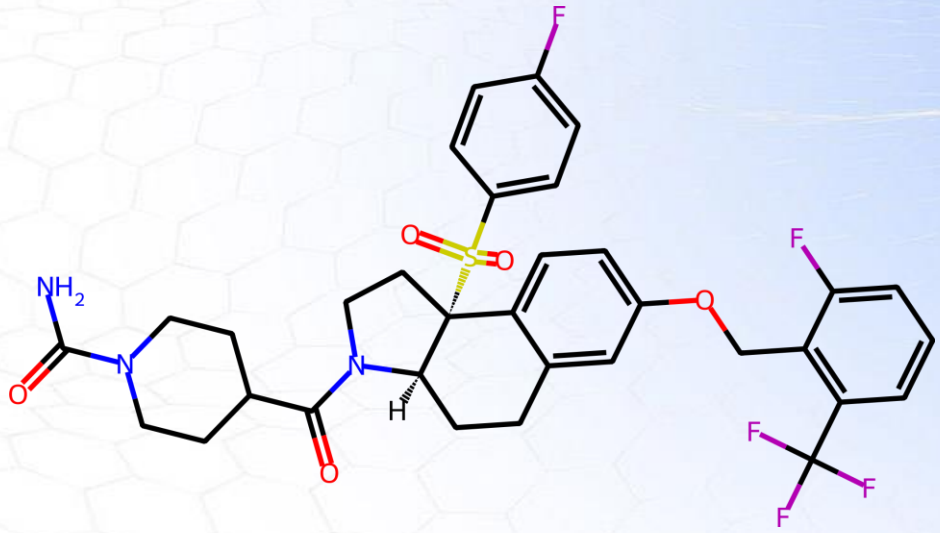


After BRICS fragmentation:

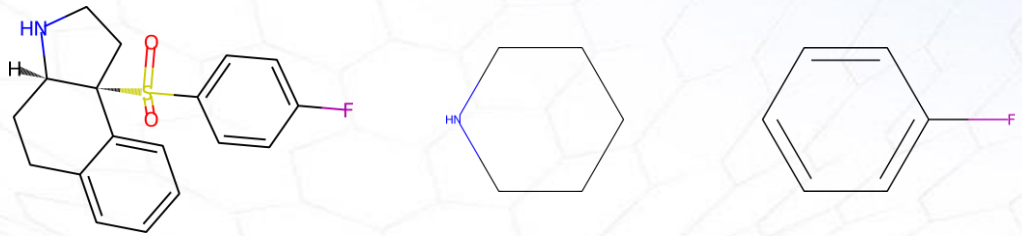


Use case: targeted RORγ agonist optimization

- **Setup:**
  - Seed: well-known strong agonist of RORγ (pEC50: 8.2).
  - Fragmenting procedure: BRICS
  - Sometimes unconstrained minimizing/maximizing is not best approach.
  - Objectives:
    - ADMET_Risk (0.10), direction: min
    - Clearance (0.35), **target: 10 L/h**
    - Fup (0.15), **target: 55%**
    - pEC50 (0.40), direction: max
  - HTPK setup:
    - Species: human
    - Route: IR Tablet
    - Dose: 1 mg/kg BW

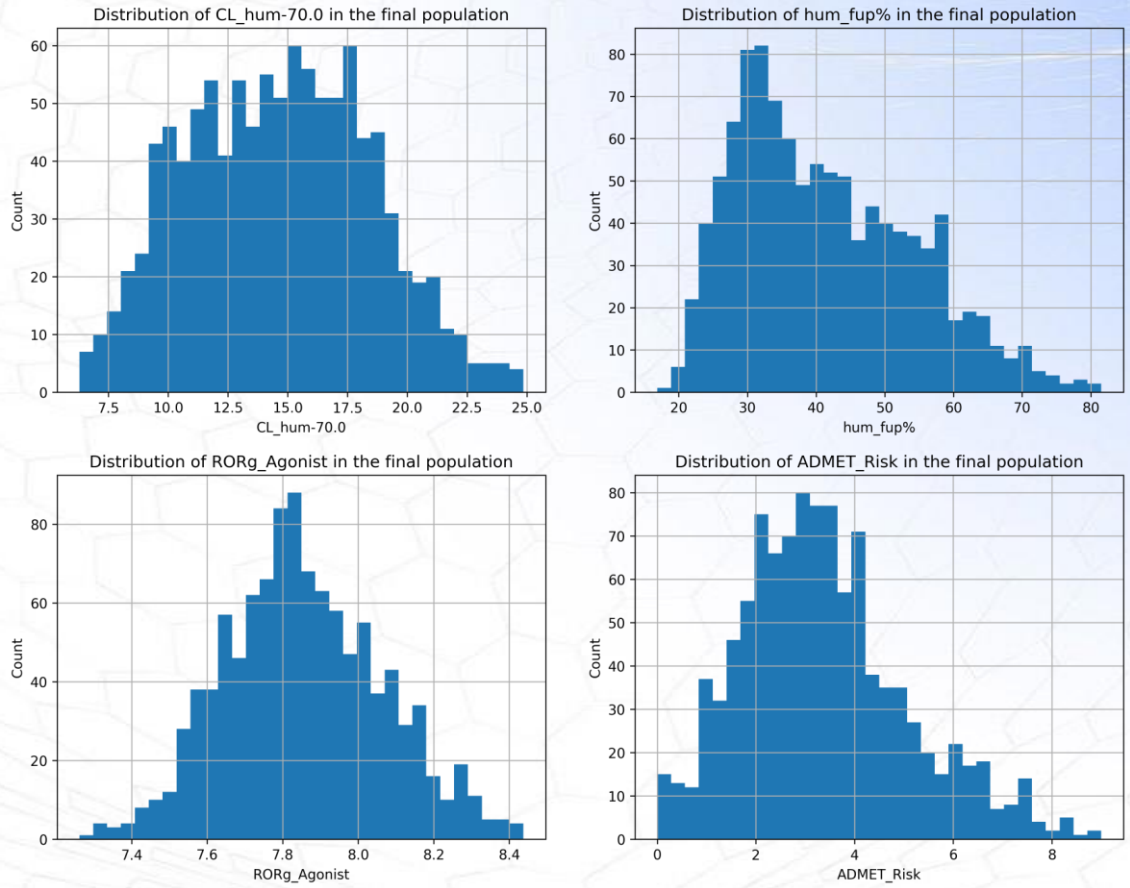


After BRICS fragmentation:



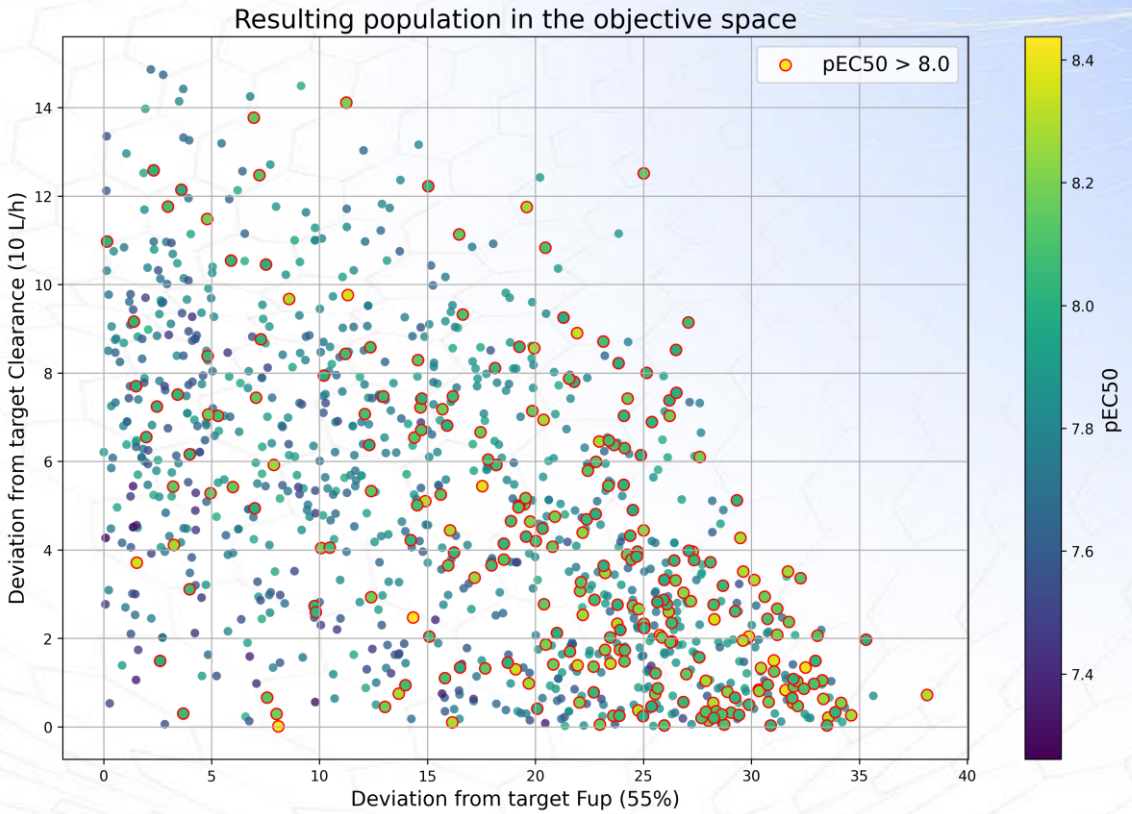
Use case: targeted RORγ agonist optimization

- **Results:**
  - The count of resulting population: 1053.
  - Scatter plot of all compounds.
  - Only predicted as active are important.
  - Pareto front (10 compounds).
  - ε-Pareto optimality: allows for some objective margin (146 compounds).



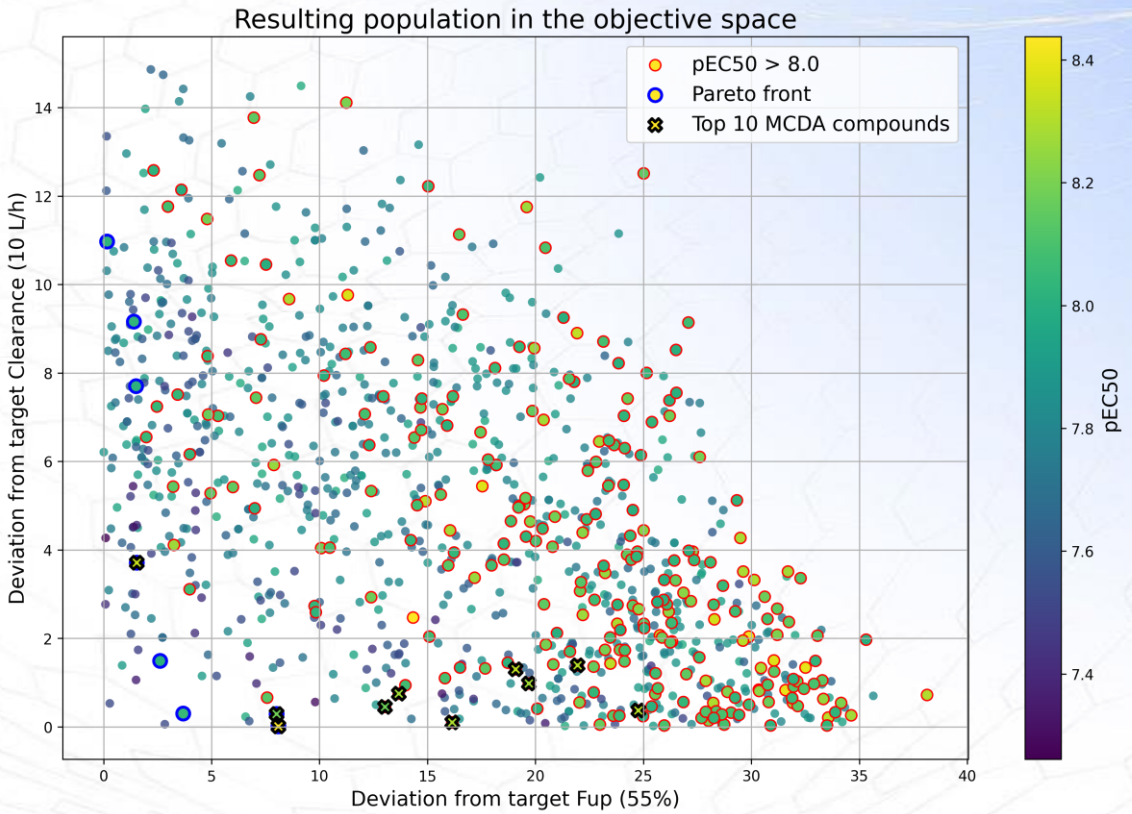
Use case: targeted RORγ agonist optimization

- **Results:**
  - The count of resulting population: 1053.
  - Scatter plot of all active compounds **in the deviation space.**
  - Top MCDA and Pareto-front compounds.
  - Switch back to the **native space.**



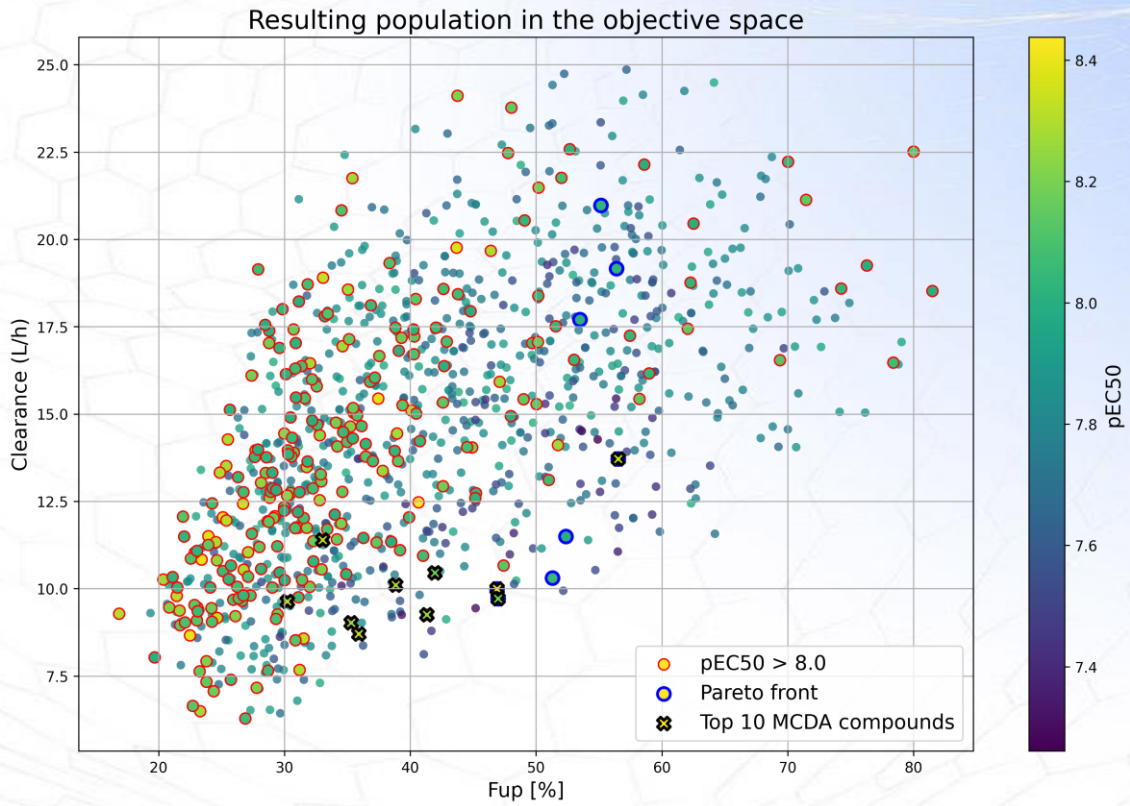
Use case: targeted RORγ agonist optimization

- **Results:**
  - The count of resulting population: 1053.
  - Scatter plot of all active compounds **in the deviation space.**
  - Top MCDA and Pareto-front compounds.
  - Switch back to the **native space.**



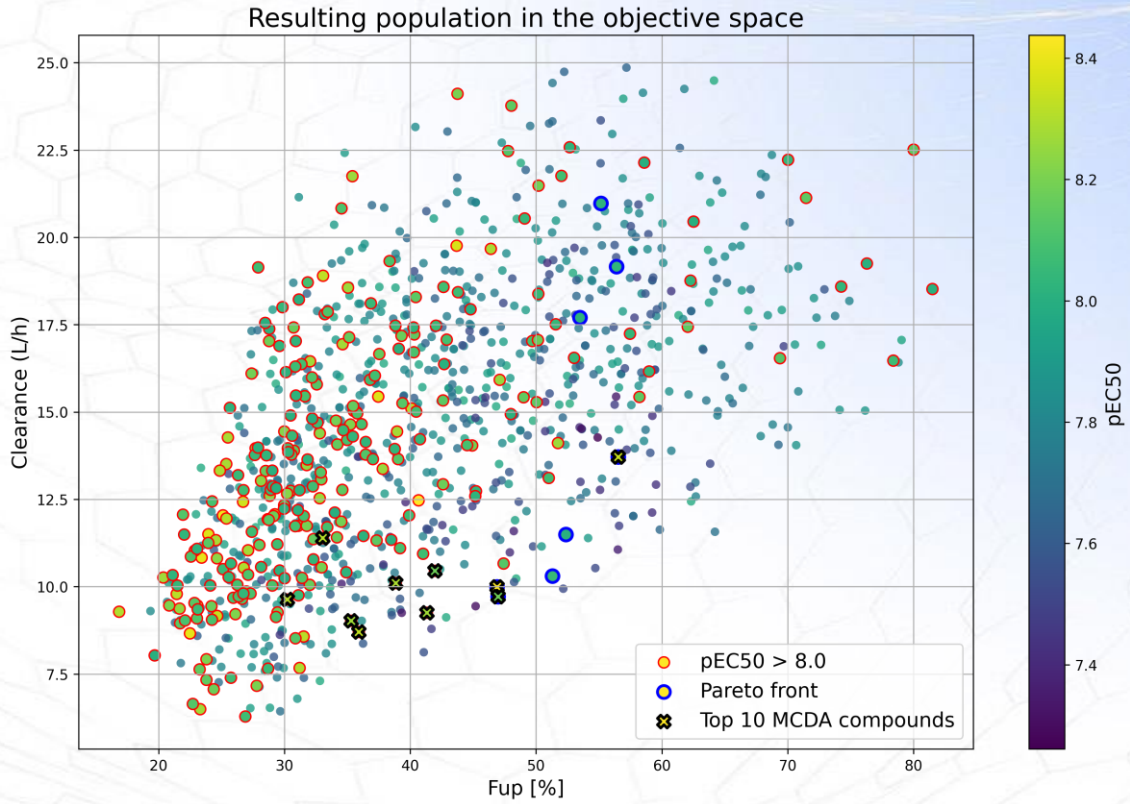
Use case: targeted RORγ agonist optimization

- **Results:**
  - The count of resulting population: 1053.
  - Scatter plot of all active compounds **in the deviation space.**
  - Top MCDA and Pareto-front compounds.
  - Switch back to the **native space.**

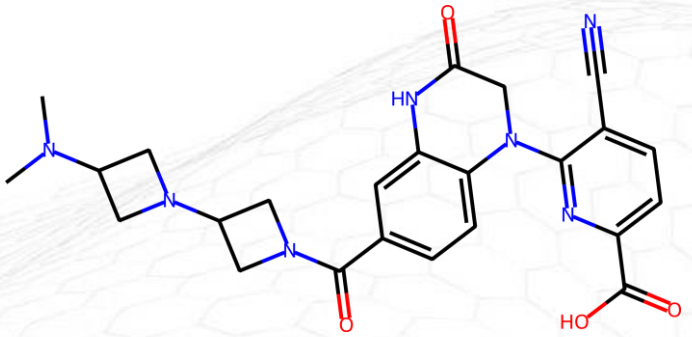


Use case: targeted RORγ agonist optimization

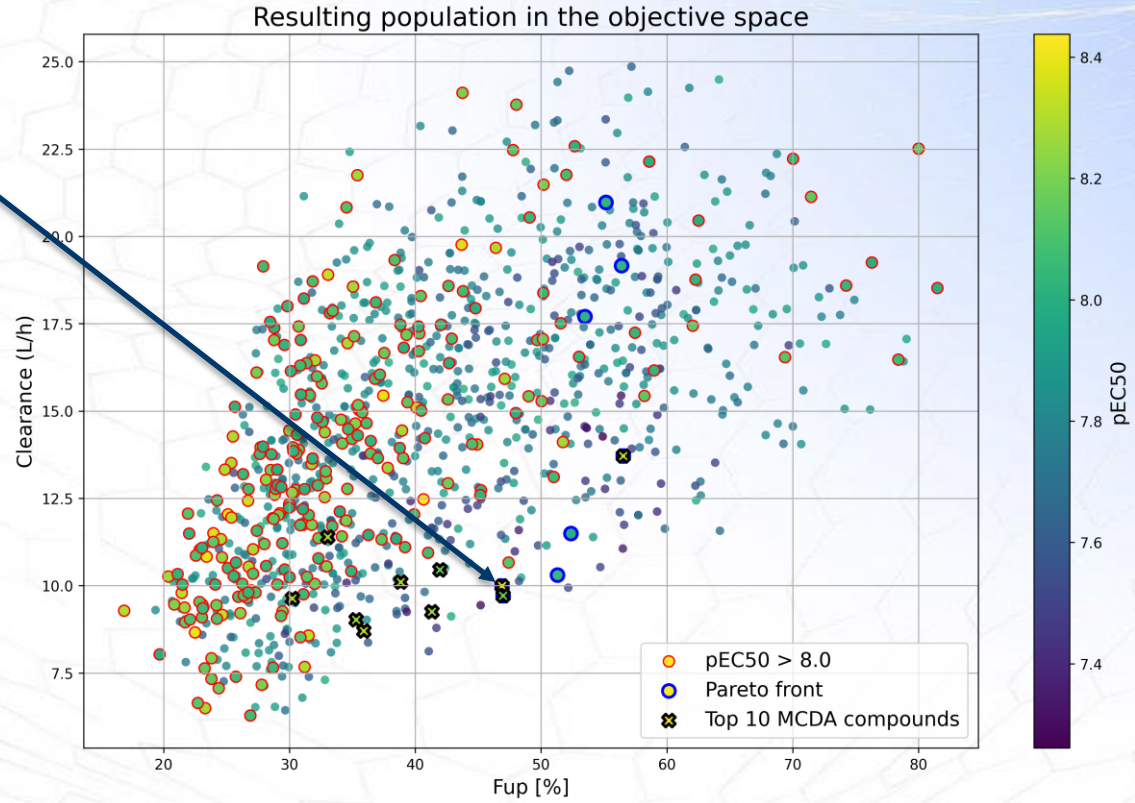
MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	5.4	10.0	46.9	8.44
1	4.8	9.3	41.3	8.28
2	4.8	10.1	38.9	8.29
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	2.0	10.5	42.0	8.18
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



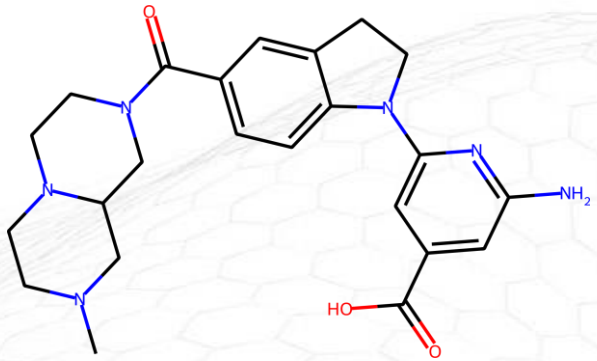
Use case: targeted RORγ agonist optimization



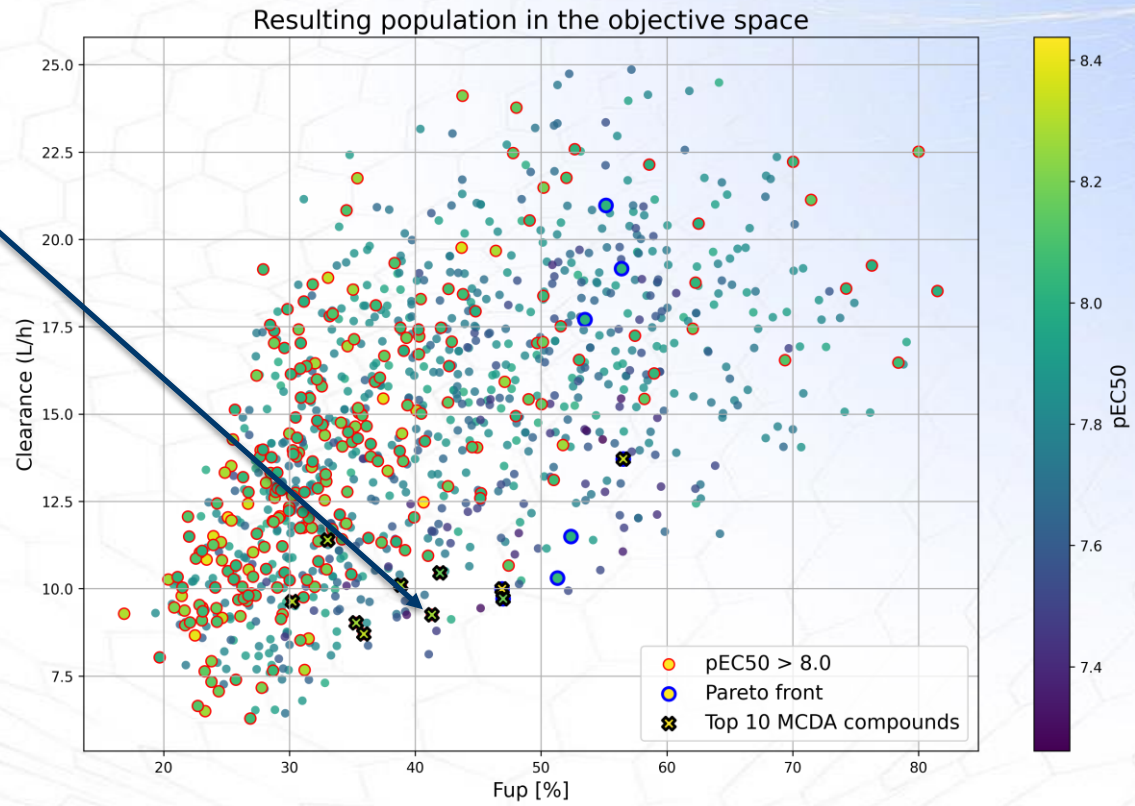
MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	**5.4**	**10.0**	**46.9**	**8.44**
1	4.8	9.3	41.3	8.28
2	4.8	10.1	38.9	8.29
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	2.0	10.5	42.0	8.18
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



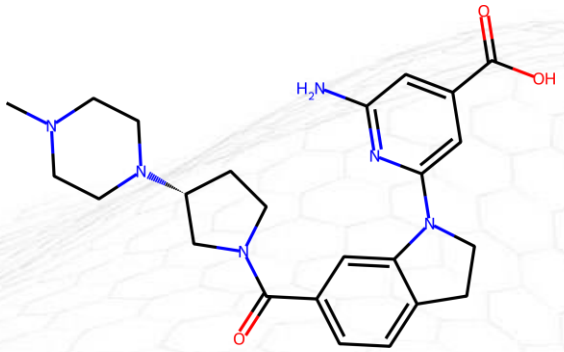
Use case: targeted RORγ agonist optimization



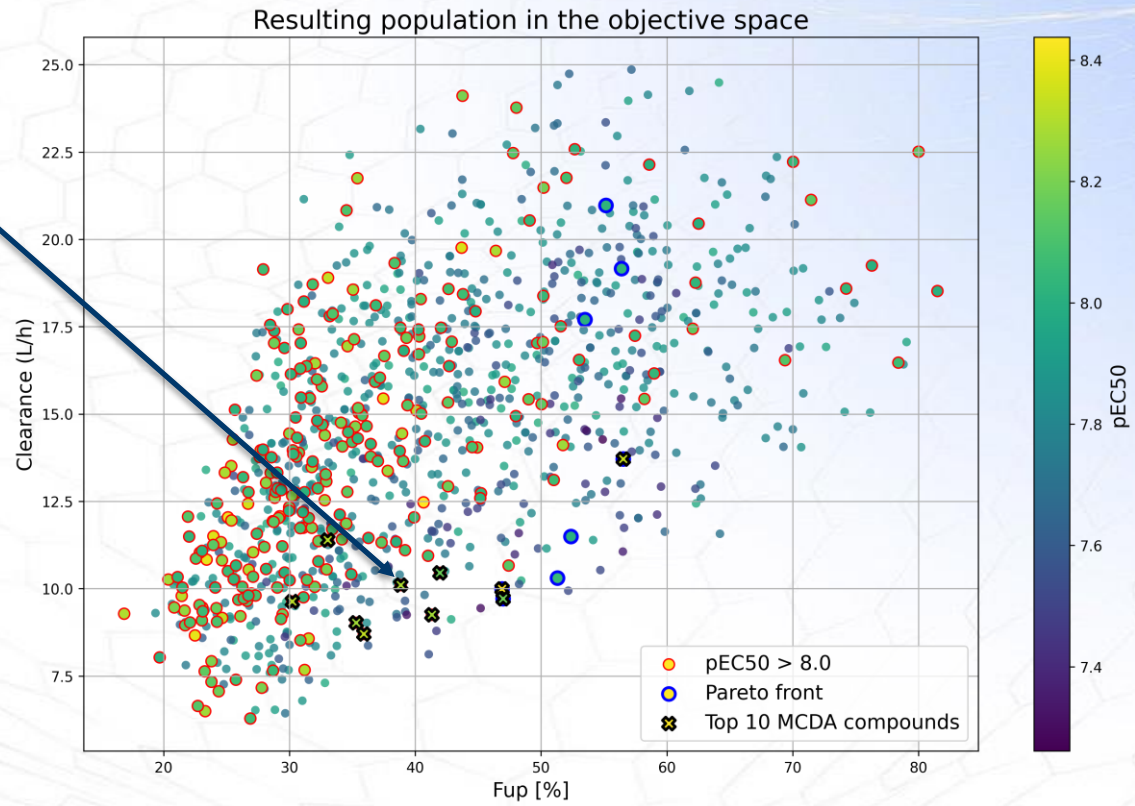
MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	5.4	10.0	46.9	8.44
1	**4.8**	**9.3**	**41.3**	**8.28**
2	4.8	10.1	38.9	8.29
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	2.0	10.5	42.0	8.18
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



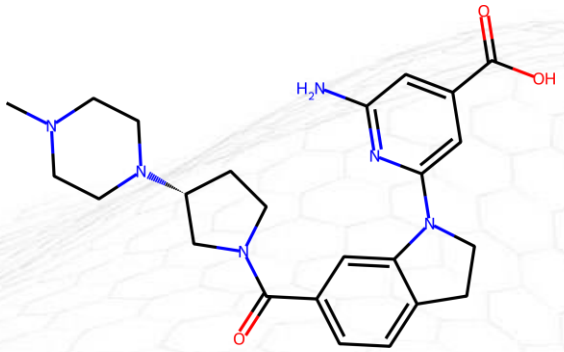
Use case: targeted RORγ agonist optimization



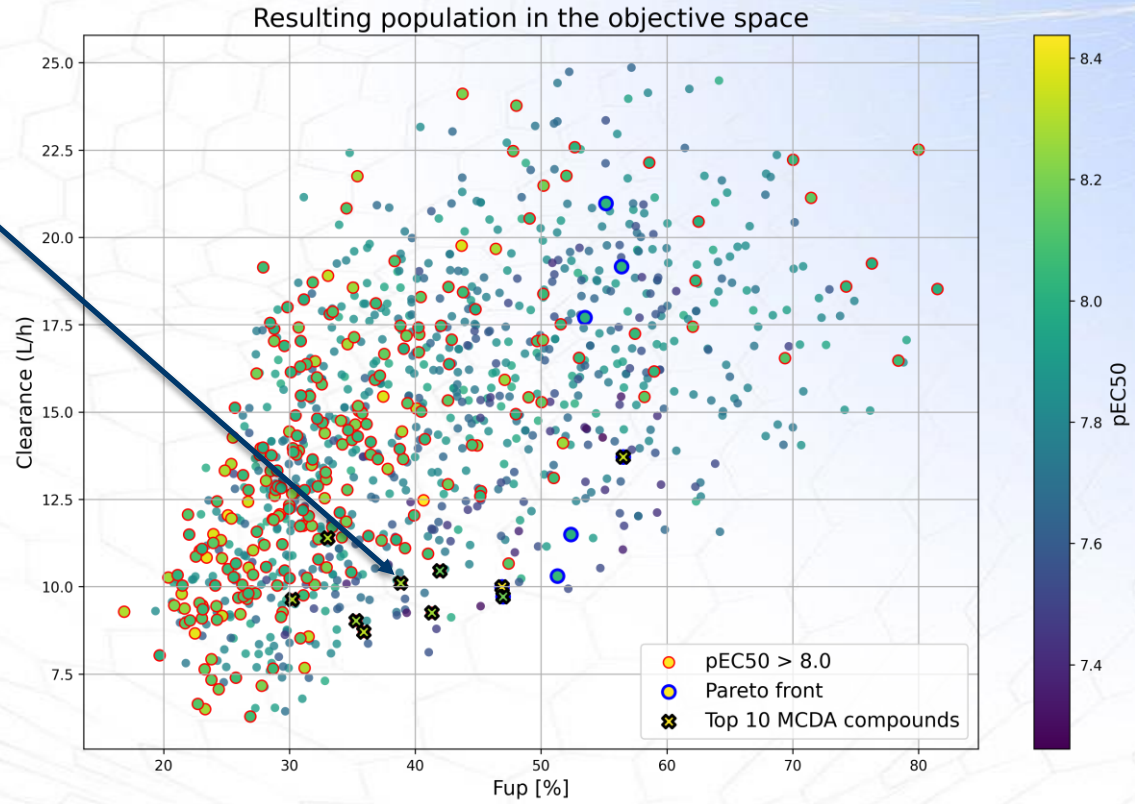
MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	5.4	10.0	46.9	8.44
1	4.8	9.3	41.3	8.28
2	**4.8**	**10.1**	**38.9**	**8.29**
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	2.0	10.5	42.0	8.18
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



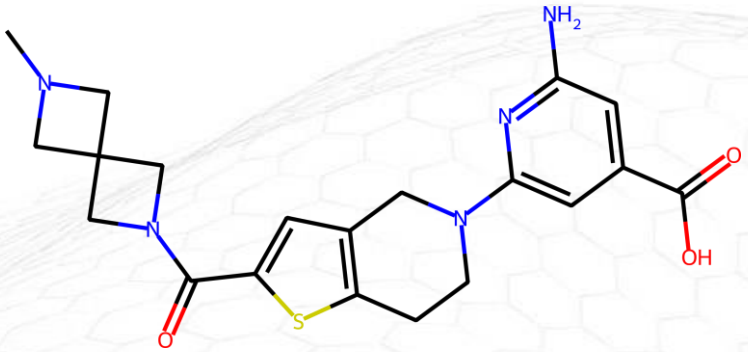
Use case: targeted RORγ agonist optimization



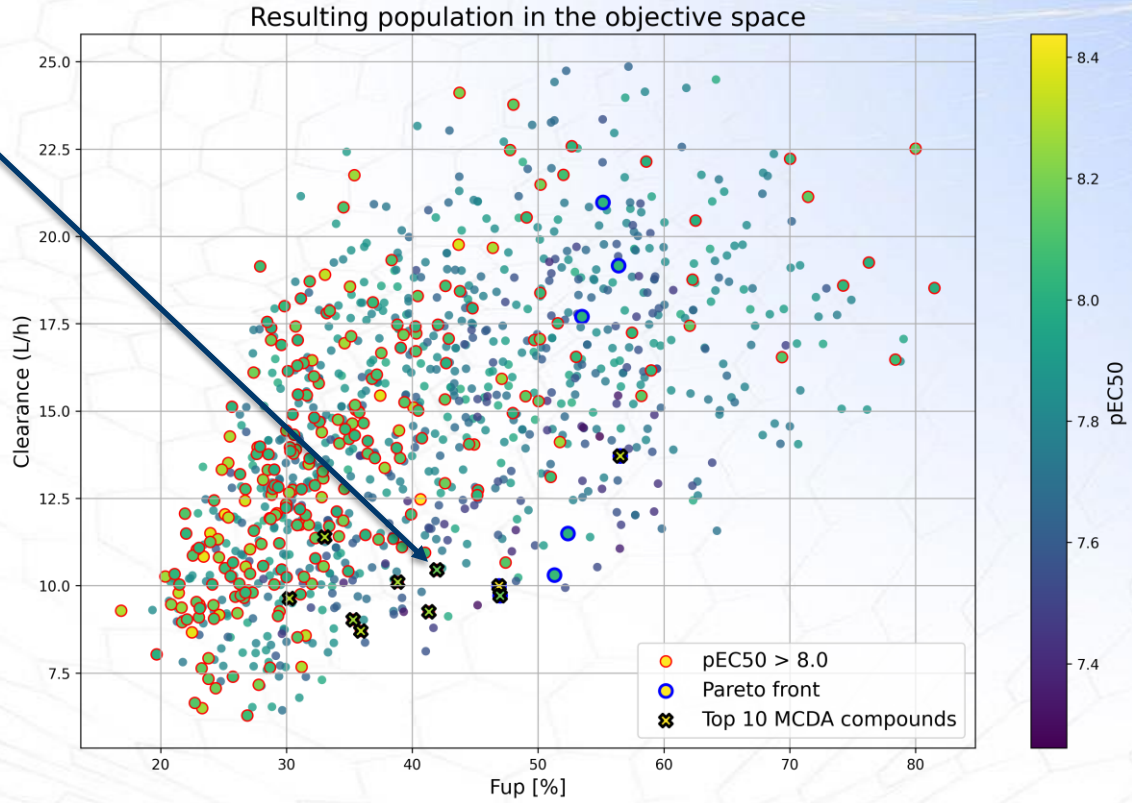
MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	5.4	10.0	46.9	8.44
1	4.8	9.3	41.3	8.28
2	4.8	10.1	38.9	8.29
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	2.0	10.5	42.0	8.18
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



Use case: targeted RORγ agonist optimization

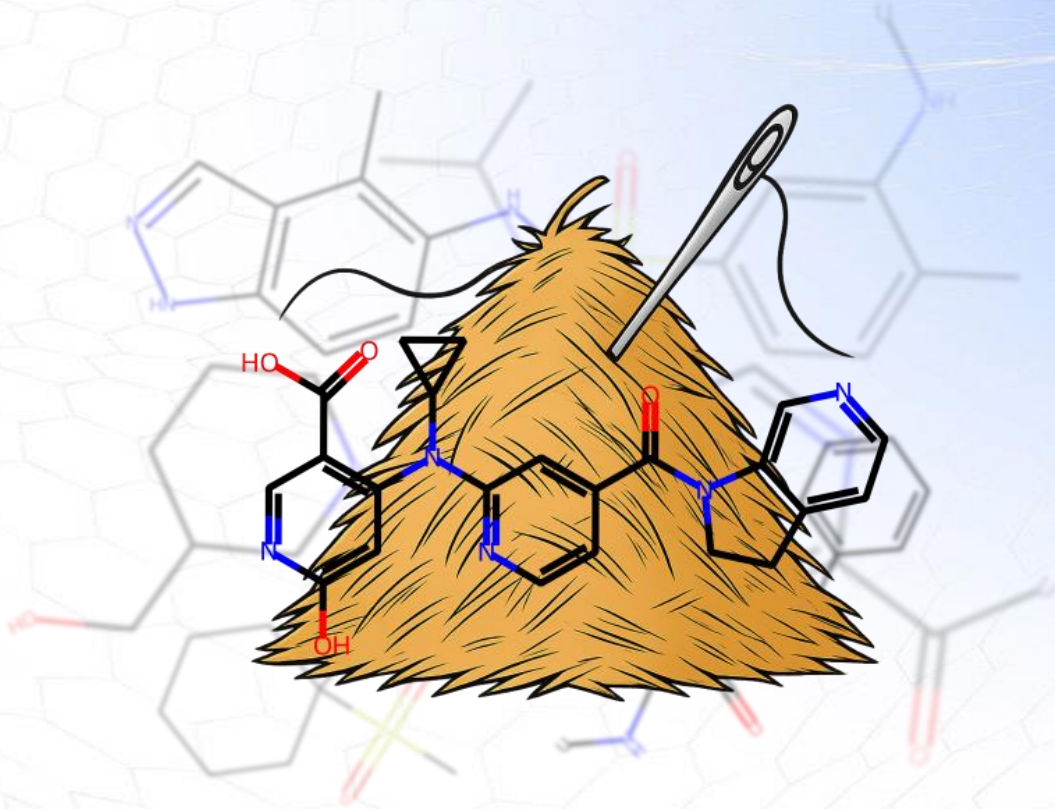


MCDA rank	ADMET_Risk	Clearance [L/h]	Fup [%]	pEC50
0	5.4	10.0	46.9	8.44
1	4.8	9.3	41.3	8.28
2	4.8	10.1	38.9	8.29
3	3.6	8.7	35.9	8.33
4	3.8	9.7	47.0	8.18
5	**2.0**	**10.5**	**42.0**	**8.18**
6	4.5	9.0	35.3	8.28
7	2.3	9.6	30.2	8.32
8	3.0	11.4	33.0	8.31
9	6.3	13.7	56.5	8.34



Conclusions

- Synthon/fragment space: convenient and lightweight.
- **SynthonForgePlus**:
  - …Python-based,
  - …Property/PK-oriented,
  - …MPO,
  - …MCDA-powered,
  - framework for chemical space exploration.
- Freedom Space 3.0 from Chemspace.
- RORγ use cases.



Don't search the haystack—build the needle.

Acknowledgments

- Colleagues from Simulations Plus:
  - Michael Lawless, David Miller, Vladimir Chupakhin, Jeremy Jones, Robert Frączkiewicz, Tom Sanderson, Viera Lukacova, John DiBella
- Chemspace LLC:
  - Olga Tarkhanova, Anna Kapeliukha, Diana Krasko
- Thank you for your attention!