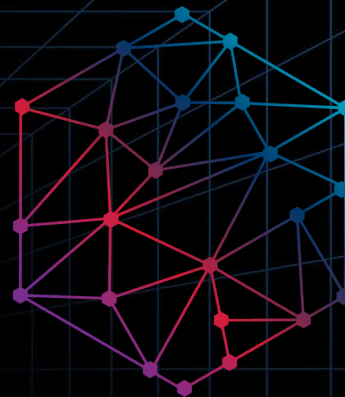


Model-Informed Drug Development

MIDD+

2021 Virtual Conference



**Designing Novel Compounds with Optimized Target
Activity and ADMET Properties Using the AIDD™ Module**

Michael S. Lawless, David Miller, and Marvin Waldman



AIDD Module

Goal – design compounds that have:

- High potency at the primary target
- High synthetic feasibility
- Good ADMET and pharmacokinetic (PK) properties



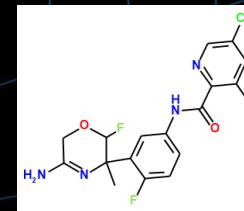
AIDD Workflow

Data Set

- Target Activity

SAR/QSAR

Molecule(s)



Candidates

transformations to randomly-
d compounds from the
population

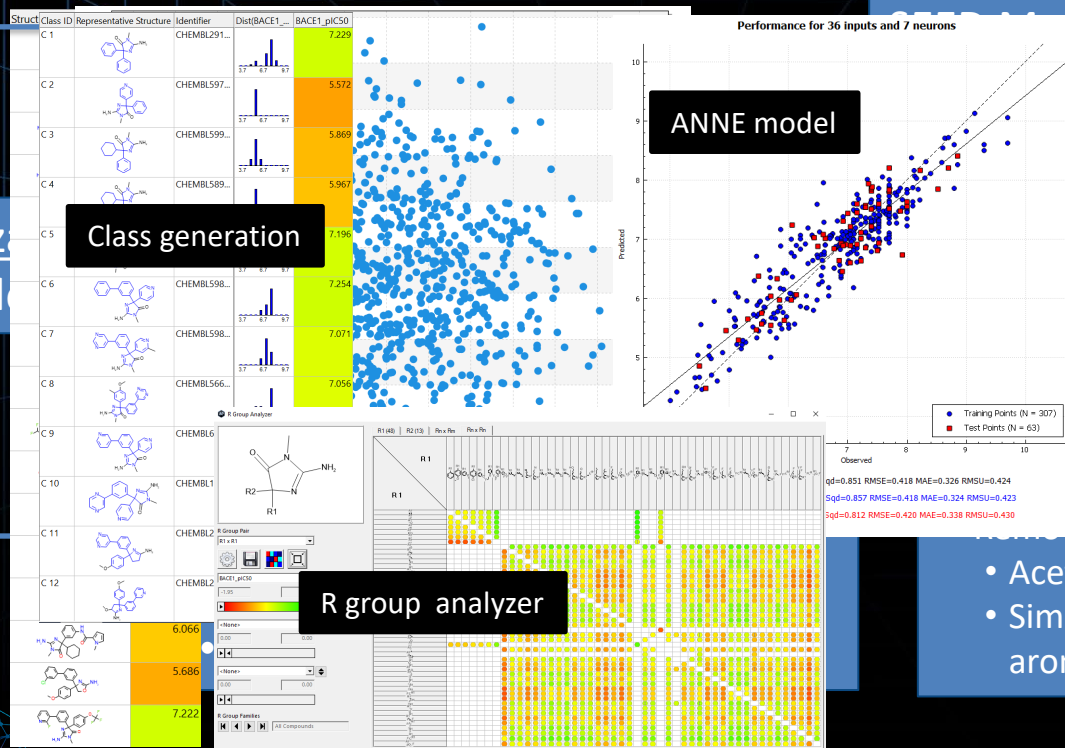
required substructure

the non-druglike

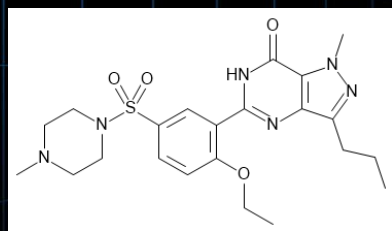
- Acetals, Michael acceptors
- Simple properties, e.g., ≤ 4 aromatic rings

Optimization

- Pareto selection



Generate Molecules Using Transform Rules

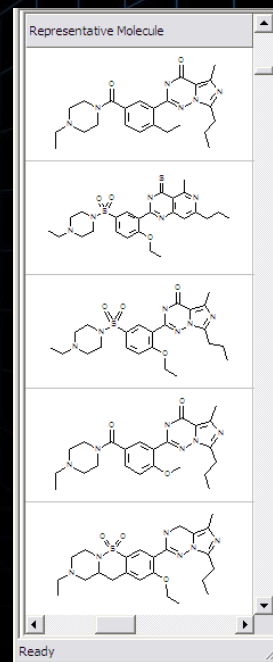


SMIRKS Transforms

- Bioisosteric replacements
- Reactions from literature or in-house expertise
- Chemically-intelligent “mutations”

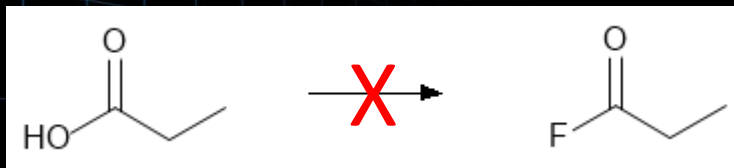
Users can modify the default rules

Generated analogs



Chemically Intelligent Transforms

- Example: Convert Non-fluorine to fluorine
 - Simple version: [!#9:1;D1_S]>>[#9:1]
 - Problem (Need to avoid)



Highly reactive acid halide

- Improved SMIRKS: [!#9;D1_S\$(*~[#6])!\$(*C=[O,N,S]):1]>>[#9:1]
- Currently ~150 transforms



Transforms

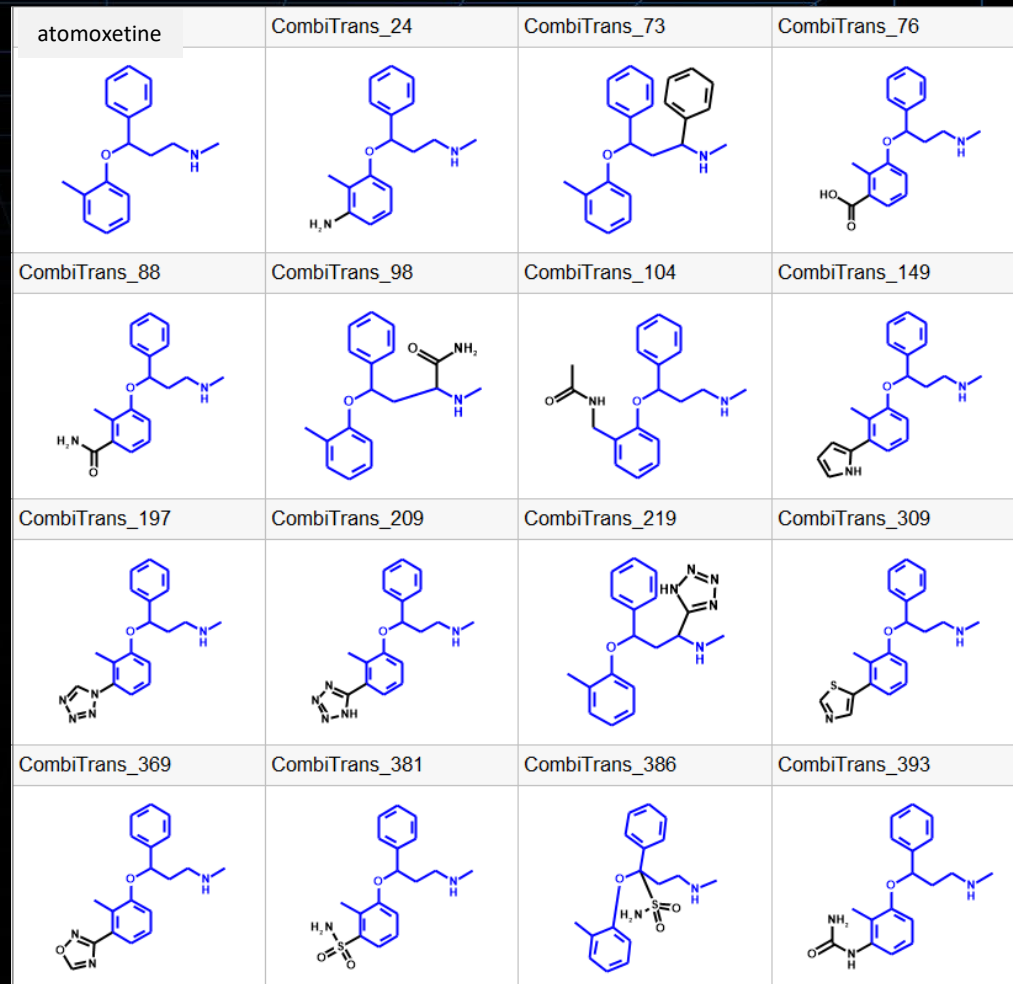
- CHANGE_FUNCTIONAL_GROUP
- CHANGE_CHAIN_LENGTH
- CREATE_RING
- BREAK_RING
- CHANGE_RING_SIZE
- CHANGE_RING_TOPOLOGY
- SHIFT_RING_SUBSTITUENTS
- CHANGE_BOND_ORDER
 - Aromatic_to_single_bond
 - Aromatize_5-membered_ring
 - Aromatize_6-membered_ring
 - De-aromatize_5-membered_ring
 - De-aromatize_6-membered_ring
 - Double_or_triple_to_single_bond
 - Double_to_triple_bond
 - Single_to_double_bond
 - Single_to_triple_bond
 - Triple_to_double_bond
- CHANGE_ATOM_TYPE
 - Non-bromine_to_bromine
 - Non-carbon_to_carbon
 - Non-chlorine_to_chlorine
 - Non-fluorine_to_fluorine
 - Non-iodine_to_iodine
 - Non-nitrogen_to_nitrogen
 - Non-oxygen_to_oxygen
 - Non-sulfur_to_sulfur
- ADD_FUNCTIONAL_GROUP
- ADD_FUSED_RING
- DELETE_FUNCTIONAL_GROUP

- CHANGE_FUNCTIONAL_GROUP
 - 2-Pyridone_to_Phenyl
 - 4-Hydroxypyridine_to_pyridone
 - 4-Pyridone_to_Phenyl
 - Acid_to_aliph_ring
 - Acid_to_arom_ring
 - Acid_to_tetrazole
 - Add_double_bond_oxygen
 - Amide_arom_insertion
 - Amide_reversal
 - Amide_to_hydroxy
 - Amide_to_hydroxy(2)
 - Amide_to_olefin
 - Arom_ring_to_ester(1)
 - Arom_ring_to_ester(2)
 - Arom_ring_to_propyl
 - CF3_to_methyl
 - Carbonyl_to_sulfonyl
 - Catechol_to_imidazole
 - Catechol_to_pyridone
 - Charged_nitrogen_to_carbon
 - Ester_to_amine
 - Ester_to_arom_ring
 - Ester_to_retroamide
 - Ester_to_sulfonamide
 - Ether_to_ethylene
 - Ethylene_to_ether
 - Het_to_sulfone
 - Hydroxy_to_amide

- ADD_FUNCTIONAL_GROUP
 - Add_1-imidazole
 - Add_1-tetrazole
 - Add_1-thiazole
 - Add_2-imidazole
 - Add_2-tetrazole
 - Add_2-thiazole
 - Add_3-piperidine
 - Add_3-tetrazole
 - Add_3-thiazole
 - Add_CF3
 - Add_amide
 - Add_amine
 - Add_bromo
 - Add_carboxylic_acid
 - Add_chloro
 - Add_cyano
 - Add_cyclohexyl
 - Add_cyclopentanone
 - Add_fluro
 - Add_hydroxyl
 - Add_iodo
 - Add_meta_furan
 - Add_meta_pyrrrole
 - Add_meta_thiophene
 - Add_methyl
 - Add_methyl_imide
 - Add_n_sulfonamide
 - Add_nitro
 - Add_ortho_furan



- Starting structure is atomoxetine
- Use transformations from “ADD_FUNCTIONAL_GROUP”
- Creates 403 compounds
- A few diverse compounds are shown on the right



Synthetic Feasibility Assessment

Based on method from Ertl and Shuffenhauer, *J. Cheminfo*, 2009, 1, 8.

$$\text{Score} = \text{fragmentScore} - \text{complexityPenalty}$$



Fragment
frequencies



- ECFP of PubChem compounds
- Does the proposed compound contain fragments of compounds that have been synthesized?



Heavy Atoms
Macrocycles
Stereocenters
Spiro centers
Bridges

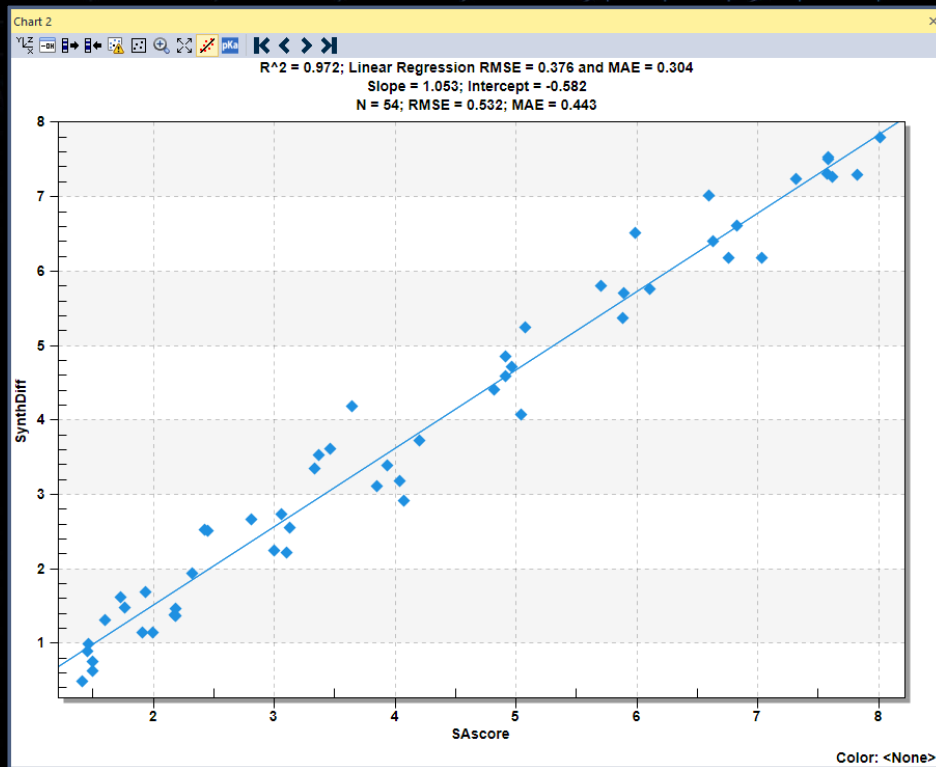


Synthetic Accessibility/Difficulty

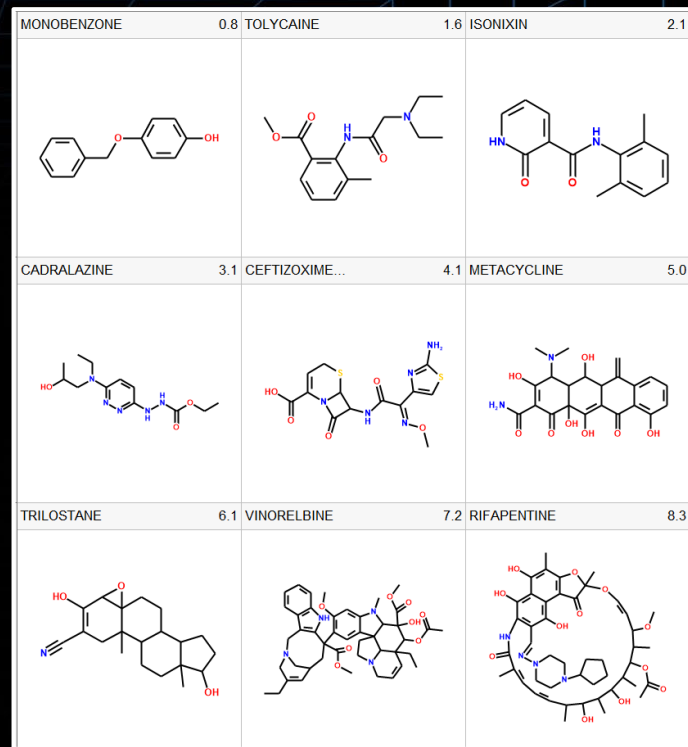
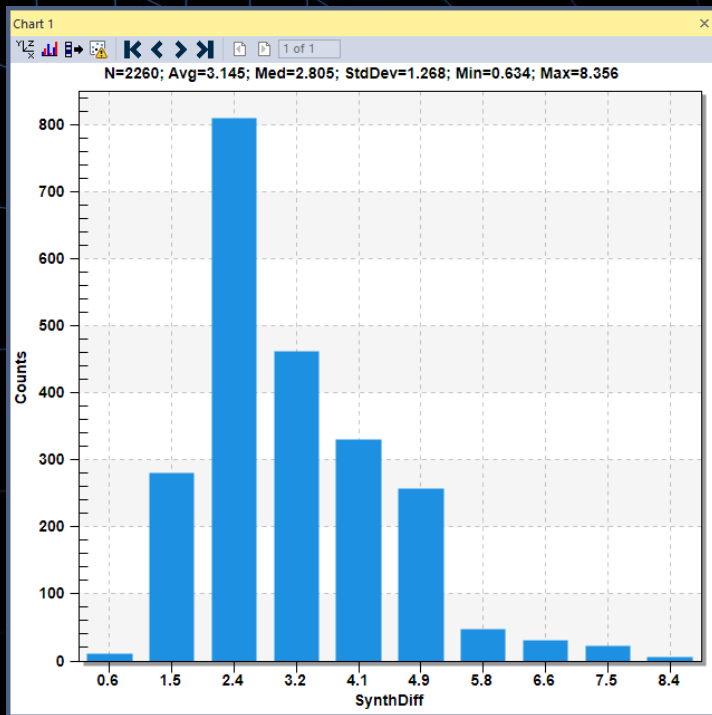
	SA Ertl ¹	SynthDiff ²
Training	~1 million	~47 million
Outer Layer	Any	aromatic vs. aliphatic
Complexity	Same	Same
Range	1-10	0-10

¹Ertl and Shuffenhauer, *J. Cheminfo*, 2009, 1, 8.

²Implemented in ADMET Predictor

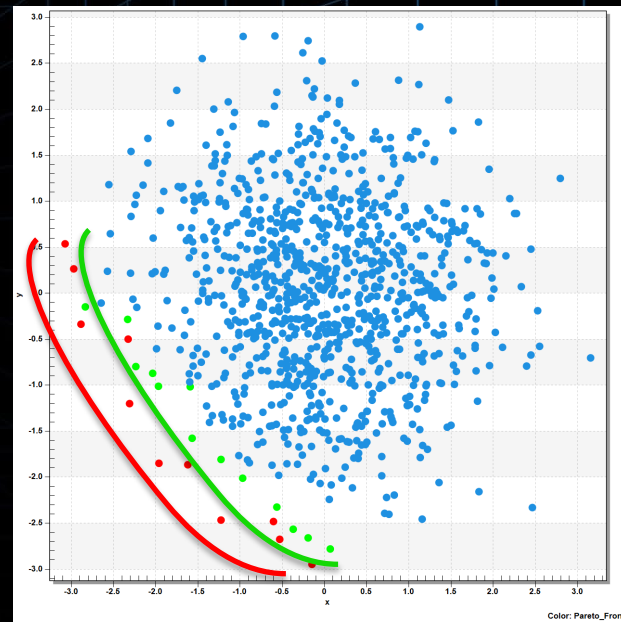
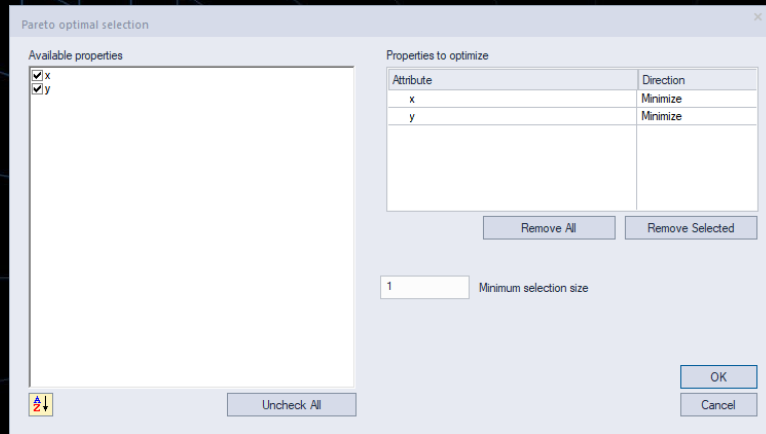


Distribution of SynthDiff Scores



Pareto Optimal Selection

- AIDD Module uses Pareto algorithm to select best molecules
- Pareto selection tool is also available in ADMET Predictor

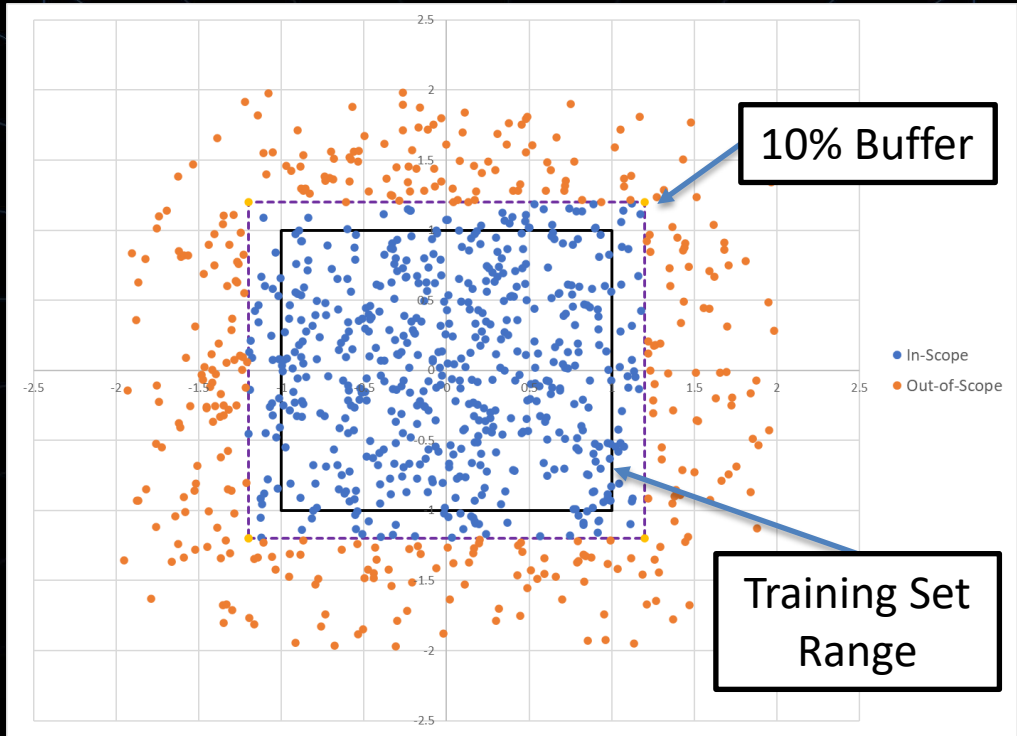


- Red points – 1st Pareto front
- Green points – 2nd Pareto front

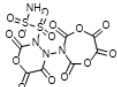
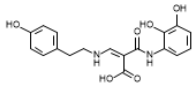
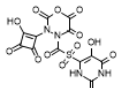
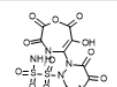
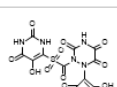
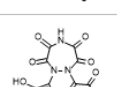


Applicability Domain

Apply penalty to predictions that are outside the scope of the model



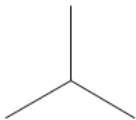
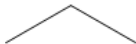



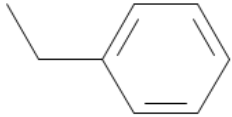
After Applying “Penalties”

	Structure	OBJ_HIVI-ST	ADMET_Risk	SynthDiff	HIVI-ST	HIVI-ST+	ADMET_Risk+	SynthDiff+
1		12.799	3.000	4.256	12.799	2.799	14.000	8.253
2		6.525	3.956	2.313	6.525	-3.475	11.956	6.312
3		12.248	3.500	4.181	12.248	2.248	13.500	7.897
4		12.163	6.000	4.132	12.163	2.163	15.000	7.848
5		11.952	4.000	4.082	11.952	1.952	15.000	6.848
6		11.051	1.694	3.674	11.051	1.051	13.404	7.674

The out of scope penalty results in deprioritization of molecules that are outside the applicability domain of the model.



Flip Side of the Coin: Capping Values

	Structure	OBJ_HIVI-ST	ADMET_Risk	SynthDiff
1		3.645	1.000	0.000
2		3.676	1.000	0.000
3		3.846	1.000	0.000
4		3.506	0.431	0.131
5		3.173	1.000	0.243
6		3.250	1.000	0.384

Properties to optimize

Name	Direction	Capping Value
<Synthetic_Difficulty+>	Minimize	2.5
ADMET_Risk	Minimize	
<Fraction bioavailable (%Fb)>	Maximize	90

Modify Selected

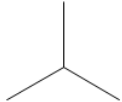
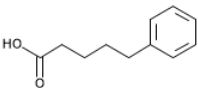

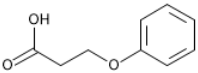

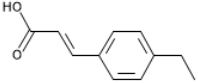


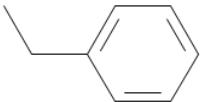
Remove Selected

Remove All

Trivially Simple Molecules:
Very easy to make
Very good in one objective

Assigning a capping value tends to filter out such molecules.
The capping value is assigned as the result when the actual result is "better", because this value is "good enough".

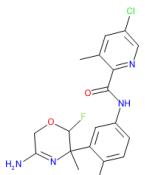
Applying Capping Values: Example

	Structure	OBJ_HIVI-ST	ADMET_Risk	SynthDiffCap		Structure	OBJ_HIVI-ST	ADMET_Risk	SynthDiff+	SynthDiffCap
1		3.645	1.000	2.500	1		4.524	0.000	1.880	2.500
2		3.676	1.000	2.500	2		4.228	0.000	1.907	2.500
3		3.846	1.000	2.500	3		4.701	0.000	2.356	2.500
4		3.506	0.431	2.500						
5		3.173	1.000	2.500						
6		3.250	1.000	2.500						

Molecules on right “dominate” molecules on left after applying capping to SynthDiff



AIDD Setup

Structure	Identifier	Previous Structure	BACE1_pIC50	BACE1
	(2R,3R)-7c		8.161	8.028

Compound optimization

Available properties

- <Synthetic_Difficult>
- <Fraction absorbed (F_a)>
- <Fraction bioavailable (F_b)>
- <Toxic_Risk
- ADMET_Risk
- BACE1_Model
- BBB_Risk
- Bioconcn
- BSEP_IC50
- CYP3A4_ki_midaz
- CYP3A4_ki_testo
- CYP_HLM Clint
- CYP_Risk
- CYP_RLM Clint
- CYPSum Clint
- Daphnia_IC50
- DMT_Cost
- HEP_pClint
- HEP_pClint
- hERG_pIC50
- hVH1_S1
- hVH1_T1
- hVH1_T2
- hVH1_T3
- hVH1_T4
- hVH1_T5
- hVH1_T6
- hVH1_T7
- hVH1_T8
- hVH1_T9
- hVH1_T10
- hVH1_T11
- hVH1_T12
- hVH1_T13
- hVH1_T14
- hVH1_T15
- hVH1_T16
- hVH1_T17
- hVH1_T18
- hVH1_T19
- hVH1_T20
- hVH1_T21
- hVH1_T22
- hVH1_T23
- hVH1_T24
- hVH1_T25
- hVH1_T26
- hVH1_T27
- hVH1_T28
- hVH1_T29
- hVH1_T30
- hVH1_T31
- hVH1_T32
- hVH1_T33
- hVH1_T34
- hVH1_T35
- hVH1_T36
- hVH1_T37
- hVH1_T38
- hVH1_T39
- hVH1_T40
- hVH1_T41
- hVH1_T42
- hVH1_T43
- hVH1_T44
- hVH1_T45
- hVH1_T46
- hVH1_T47
- hVH1_T48
- hVH1_T49
- hVH1_T50
- hVH1_T51
- hVH1_T52
- hVH1_T53
- hVH1_T54
- hVH1_T55
- hVH1_T56
- hVH1_T57
- hVH1_T58
- hVH1_T59
- hVH1_T60
- hVH1_T61
- hVH1_T62
- hVH1_T63
- hVH1_T64
- hVH1_T65
- hVH1_T66
- hVH1_T67
- hVH1_T68
- hVH1_T69
- hVH1_T70
- hVH1_T71
- hVH1_T72
- hVH1_T73
- hVH1_T74
- hVH1_T75
- hVH1_T76
- hVH1_T77
- hVH1_T78
- hVH1_T79
- hVH1_T80
- hVH1_T81
- hVH1_T82
- hVH1_T83
- hVH1_T84
- hVH1_T85
- hVH1_T86
- hVH1_T87
- hVH1_T88
- hVH1_T89
- hVH1_T90
- hVH1_T91
- hVH1_T92
- hVH1_T93
- hVH1_T94
- hVH1_T95
- hVH1_T96
- hVH1_T97
- hVH1_T98
- hVH1_T99
- hVH1_T100

Properties to optimize

Name	Direction	Capping Value
<Synthetic_Difficult>	Minimize	4.5
<Fraction bioavailable (F _b)>	Maximize	90
ADMET_Risk	Minimize	
BACE1_Model	Maximize	
BBB_Risk	Minimize	

Parameters for selected properties

- 1 Out-of-scope factor for risk models
- 10 Out-of-scope penalty for standard models
- 7.4 pH for pH-dependent properties
- Species for pharmacokinetic properties
 - Rat
 - Human
- 100 Dose for pharmacokinetic properties [mg]

Filter rules

Scaffold query (optional)

```
SLQ [NH2][C]:z5,x6|[R1][R1]@N[_S](~*)(~*)c1[cH]c([NH]C
```

- Define scaffold
- Select filter criteria

Input file containing product filter criteria (optional)

I:\Simulations Plus\ADMET Predictor\AIDD\AIDD-example\ Browse

Clear File Open File Query Wizard

Successful compounds must Pass Fail every query filter Cancel

- Select transform rules
- Define no. of optimization cycles
- Select output folder

Run parameters

Input file containing transform rules

C:\Users\mlawless\AppData\Local\Simulations Plus, Inc\ Browse

View Transforms Enable or disable individual transform rules

5 Number of optimization generations 12345 Random seed

500 Number of candidate molecules per generation 10 Intermediate file frequency (%)

1000 Size of initial population Use multithreading for product generation

500 Minimum size of population after each pruning cycle

Folder for result files

I:\Simulations Plus\ADMET Predictor\AIDD\AIDD-example\ Browse

Write input compound(s) to the output

Display results in new window

Cancel < Back Run

- Select seed molecule(s)
- Select properties to optimize
- Define capping values



A Few Results

The screenshot displays the ADMET Predictor software interface. On the left, there is a 'Spreadsheet Controls' panel with various input fields and dropdown menus. The main area is a grid of 20 cells, each containing a chemical structure and its corresponding ADMET value. The values are: 2561, 1756, 2306, 2371, 1702, 2631, 2777, 2124, 2041, 2611, 1519, 2623, 2181, 1811, 1594, 2010, 2822, 2967, 2661, 2832, 1806, 1530, 2521, and 1732. The bottom of the interface shows a navigation bar with tabs for 'Component', 'Class', 'Tables', 'Risks', and 'Logs'. The status bar at the very bottom indicates 'Ready' and shows the number of selected items (0).

ADMET Value	Chemical Structure
2561	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1756	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2306	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2371	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1702	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2631	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2777	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2124	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2041	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2611	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1519	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2623	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2181	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
1811	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1594	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2010	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2822	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2967	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2661	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
2832	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1806	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1530	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Br)cc(Cl)c2</chem>
2521	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>
1732	<chem>CC1=CC=C(C=C1)C(=O)Nc2cc(Cl)cc(Cl)c2</chem>



AIDD Module

Goal – design compounds that have:

- High activity based on a QSAR model
 - Penalizes out of scope predictions
- High synthetic feasibility
 - Based on fragment counts of PubChem compounds
- Good ADMET and PK properties
 - Incorporates absorption, CYP metabolism, toxicity, and oral bioavailability



Q & A

Questions & Answers

Model-Informed Drug Development

MIDD+

2021 Virtual Conference



Michael Lawless
mlawless@simulations-plus.com



 **Learn More!** www.simulations-plus.com

S+ *SimulationsPlus*
Cognigen | DILIsym Services | Lixoft