Adding Reactions into the Chemistry Commons, Retrosynthetic Analoging, and Reaction Enumeration

UGI By Maksim DuNova

Chemistry Commons

The goal of Chemistry Commons is to make an interactive space for chemists to facilitate collaboration on reaction curation and to share their own reactions for the purpose of exploring chemical space!

Chemistry Commons Reactions Synthons

Login/Register -

Welcome to Chemistry Commons!

The Chemistry Commons is an effort to facilitate the connections between synthetic organic chemists biologists seeking new reagents for biology as well as medicinal chemists and computational chemists.

The Chemistry Commons is a platform for sharing chemistry. Synthetic organic chemists looking to make their chemistry available to explore. Chemists use the chemistry commons to articulate and enumerate libraries of accessible compounds. With our assistance, libraries in 2D and 3D are prepared and made available. Medicinal chemists and biologists seeking new chemistry for their biological targets use our tools to search, download, and dock these libraries. Compounds are prioritized for synthesis, and an agreement is made between the synthetic chemist and the biology lab testing the compounds. Ideally, compounds are produced, tested, and, ideally, some compounds will work.

Reaction Database

Chemistry Commons Reactions Synthons

.ogin/Register 🔻

Filter By Name

Short Name	Description (Product : Reagents)	RXN SMARTS
ACETYLATION	Amide: amine + acid anhydride	[#6:1][NH:2][#6:3].[C:4][C:5](=O)OC(=O)C>>[#6:1][N:2]([#6:3
AC-SUZUKI-CI	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both comp	[c:1]B(0)0.[c:2][C1]>>[c:1]-[c:2]
AC-SUZUKI-Br	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both comp	[c:1]B(0)O.[c:2][Br]>>[c:1]-[c:2]
AC-SUZUKI-I	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both comp	[c:1]B(0)O.[c:2][I]>>[c:1]-[c:2]
AMIDE	Amide: amine + acid + CMPI	[C:1](=[0:2])[0D1].[N:3]>>[N:3][C:1]=[0:2]
Primary AMIDE	Amide: primary amine + acid + CMPI	[C:1](=[0:2])[OD1].[N:3]>>[N:3][C:1]=[0:2]
Secondary AMIDE	Amide: amine + acid + CMPI	[C:1](=[O:2])[OD1].[N:3]>>[N:3][C:1]=[O:2]
Primary Boc-AMIDE	Amide: primary BOC-amine + acid	[C:1](=[O:2])[OD1].[N:3]>>[N:3][C:1]=[O:2]
Secondary Boc-AMIDE	Amide: secondary BOC-amine + acid	[C:1](=[O:2])[OD1].[N:3]>>[N:3][C:1]=[O:2]
AMIDE1	Amide: amine + acid + EDC	[C:1](=[O:2])[OD1].[N:3]>>[N:3][C:1]=[O:2]

Rows per page: 10 - 1-10 of 109

Synthon Database

Chemistry Commons Reactions Synthons

Login/Register

Synthons

Show entries		Search
10		
Synthon ID	Synthon Name	SMARTS
1	Acetylenes_AlkyneCH	C#[CH]
2	Acid_Aromatic_Acid	[OD1]C(=O)c1cccc1

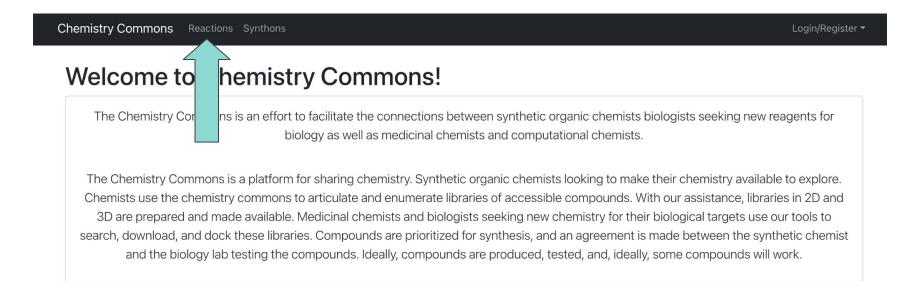
Reaction Details

Chemistry Commons Reactions Synthons

ogin/Register

	Reaction Details			
	Reaction Details			
Reaction ID	CC-1			
Reaction Description	Amide: amine + acid anhydride			
Reaction SMARTS	[#6:1][NH:2][#6:3].[C:4][C:5](=O)OC(=O)C>>[#6:1][N:2]([#6:3])[#6:5](=O)[#6:4]			
SMARTS Description				
	Synthons			
	Gynanono			
AnhydridesAcyclic_Anhydrid	des			
Inclusion Smarts				
[#6]-C(=O)!@-O-C(-[#6])=O Arthor				
Exclusion Smarts				
C=[NH]				
[SiH,SiH2,SiH3,SiH4,PH,PH2,	PH3,SH,SH2,BH,BH2,BH3,BH4]			
[N;\$([NH,NH2]);!\$(NC=[C,N,O,P,S])!\$(N=[C,N,O,S])!\$(N[P,S])!\$(N#[C])]				
[S,P](=[O,N])[F,CI]				
[OH,SH,S-,O-]				

How to add a Reaction into the Chemistry Commons.



About | Contact | Privacy Policy | Terms of Use | Copyright © 2024 Irwin and Shoichet Labs and UC Regents

How to add a Reaction into the Chemistry Commons.

Chemis	try Commons	Reactions Synthons		Kha	inh Tang
Sort By	Tag ▼ 20 ▼ Per	Page		search	
		« < 1 2 3 4 5 >			$\left\{ \right.$
Rxn ID	Short Name	Product: Reagents	Rxn SMARTS		Act
CC-1	Acetylation	Amide: amine + acid anhydride	[C:1][NH:2][C:3].[C:4][C:5](=O)OC(=0 [C:5](=O)[C:4]	D)C>>[C:1][N:2]([C:3])	
CC- 2.11	Suzuki-Cl-CA-1	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both components contain carboxylic group in a side chain.	[c:1]B(O)O.[c:2][Cl]>>[c:1]-[c:2]		0
CC- 2.12	Suzuki-CI-CA-2	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both components contain carboxylic group in a side chain.	[c:1]B(O)O.[c:2][Cl]>>[c:1]-[c:2]		0
CC- 2.13	Suzuki-Cl-CA-3	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both components contain carboxylic group in a side chain.		0	
CC- 2.21	Suzuki-Br-CA-1	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both components contain carboxylic group in a side chain. [c:1]B(O)O.[c:2][Br]>>[c:1]-[c:2]		0	
CC-	Suzuki-Br-CA-2	Biaryl (Suzuki): aryl boronic acid + aryl halide; one or both components	[c:1]B(O)O.[c:2][Br]>>[c:1]-[c:2]		0

Enter in Reaction Details and Submit!

Register Reaction	Synthons	
Reaction Name*	Synthon 1	
Enter Reaction Name		
Reaction Description	Synthon Name*	
	Enter Synthon Name	
Enter Reaction Description	InSMARTS*	
	Enter SMARTS	
Reaction SMARTS*		
Enter SMARTS		
SMARTS Description	ExSMARTS*	
Enter SMARTS	Enter SMARTS	

Submit Reaction and Synthons

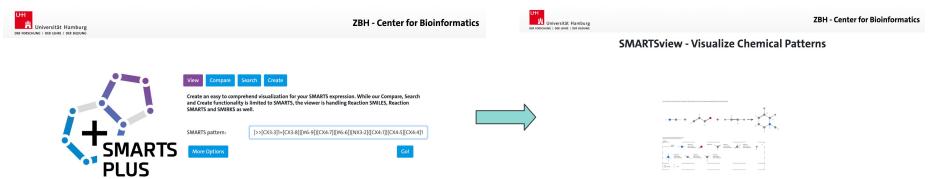
Enter in Reaction Details and Submit!

Register Reaction	Synthons	\oplus
Reaction Name*	Synthon 1 🛞 Synthon 2 🛞 Synthon 3 🛞 Syntho	n4 ⊗
Ugi-New V		
Reaction Description	Synthon Name*	
Covalent Ugi Reaction	alphaAllylNitrile-UGI	~
	InSMARTS*	
	C=C-[C](=O)-[OD1]	
Reaction SMARTS*		
[C:10]=[C:9]-[C:1](=[O:11])-[OD1].[CX4:8]-[ND1:2].[CX4:7]-[CD2:3]=O.[C:6]#[N:5]>>[C:10]=[C:9]-[C:1] (=[O:11])-[N:2]([C:8])-[C:3]([CX4:7])-[C:6]-[N:5]		
SMARTS Description/References	ExSMARTS*	
Alpha-allyl carboxylate plus primary amine plus aldehyde/ketone plus isonitrile gives the product.		
	[C][CH]=O	
	[C]#[N]	
Add Reaction Tag		
4-reagent × Covalent ×		

First Step is to learn SMARTS Resources

- <u>Smarts Plus</u>
- <u>Daylight SMARTS</u>

These are useful resources in helping learn the syntax of SMARTS, visualize SMARTS, as well as seeing if the SMARTS patterns are valid or not



First Step is to learn SMARTS Resources

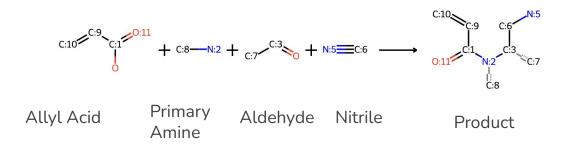
BBFilter module on tldr.docking.org

This is a module on from our own tool that helps verify your inclusion and exclusion SMARTS

TLDR Start Jobs	
bbfilter	
I Filter SMARTS pattern from BB list.	
Inputs	
included_SMARTS Include this SMARTS pattern	
excluded_SMARTS Exclude this SMARTS pattern	
BB_library bb_50	~
HAC_range Min 6 Max 25	
Enter memo for your job Memo	
Submit Job	

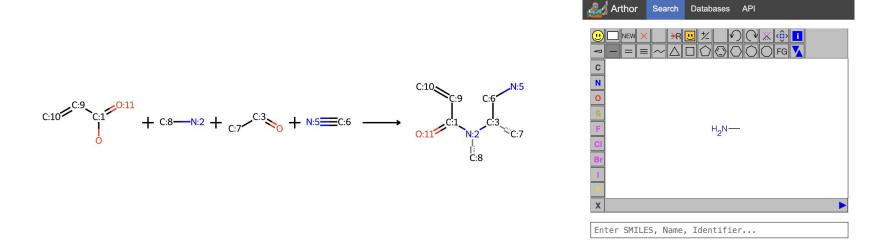
Transforming the Reaction into SMARTS

- Figure out the reaction that you want to transform into SMARTS!
- In this example we are transforming, UGI
- UGI is a 4-Component Reaction



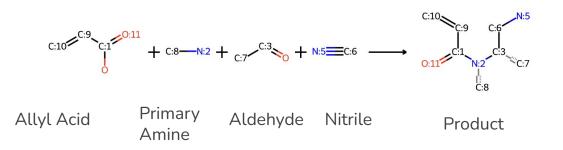
Transforming the Reaction into SMARTS: Finding the SMILES of each Reactant

- First step is to figure out the SMILES of the reactant(s) and product
- You can do this simply by drawing the molecule in our tool <u>Arthor</u> and copying the SMILES output by the drawing tool



Transforming the Reaction into SMARTS: R Group SMARTS

- After you figure out the SMILES of each component, now specify the R-group bond order/connections using SMARTS notation if this is required for your reaction
- Example: C-N -> [CX4]-[NH2]



Transforming the Reaction into SMARTS: Daylight Resource

• Again <u>Daylight SMARTS</u> is a good resource if you need to learn SMARTS notation

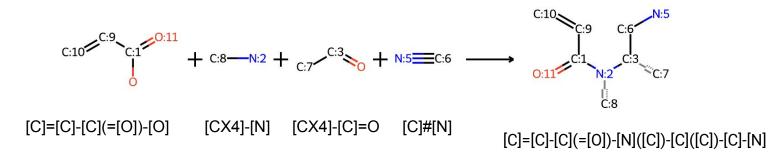
SMARTS Atomic Primitives					
Symbol	Symbol name	Atomic property requirements	Default		
*	wildcard	any atom	(no default)		
а	aromatic	aromatic	(no default)		
A	aliphatic	aliphatic	(no default)		
D <n></n>	degree	<n> explicit connections</n>	exactly one		
H <n></n>	total-H-count	<n> attached hydrogens</n>	exactly one ¹		
h <n></n>	implicit-H-count	<n> implicit hydrogens</n>	at least one		
R <n></n>	ring membership	in <n> SSSR rings</n>	any ring atom		
r <n></n>	ring size	in smallest SSSR ring of size <n></n>	any ring atom ²		
v <n></n>	valence	total bond order <n></n>	exactly one ²		
X <n></n>	connectivity	<n> total connections</n>	exactly one ²		
x <n></n>	ring connectivity	<n> total ring connections</n>	at least one ²		
- <n></n>	negative charge	- <n> charge</n>	-1 charge (is -2, etc)		
+ <n></n>	positive charge	+ <n> formal charge</n>	+1 charge (++ is +2, etc)		
#n	atomic number	atomic number <n></n>	(no default) ²		
@	chirality	anticlockwise	anticlockwise, default class ²		
00	chirality	clockwise	clockwise, default class ²		
@ <c><n></n></c>	chirality	chiral class <c> chirality <n></n></c>	(nodefault)		
@ <c><n>?</n></c>	chiral or unspec	chirality <c><n> or unspecified</n></c>	(no default)		
<n></n>	atomic mass	explicit atomic mass	unspecified mass		

SMARTS Atomic Primitives

Transforming the Reaction into SMARTS: Inclusion SMARTS

• At the step you should have the some simplified SMARTS of each component

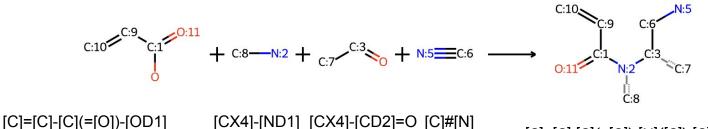
1. Allyl Acid 2. Primary 3. Aldehyde 4. Nitrile Amine



Transforming the Reaction into SMARTS: Bond SMARTS

- Save the SMARTS from the previous step as these will be our inclusion SMARTS
- Next step is to specify bonds!

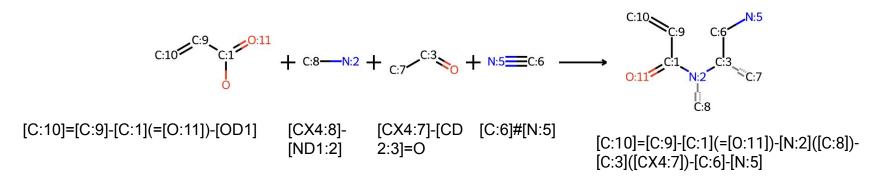
1. Allyl Acid 2. Primary 3. Aldehyde 4. Nitrile Amine



C]=[C]-[C](=[0])-[N]([C])-[C]([CX4])-[C]-[N]

Transforming the Reaction into SMARTS: Mapping Atoms

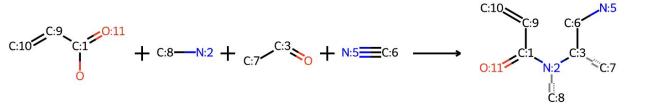
- Next is to **Map** the atoms in the reactants into the product (long step)
 - 1. Allyl Acid 2. Primary 3. Aldehyde 4. Nitrile Amine



Transforming the Reaction into SMARTS: Combining it all together

• Now these parts can be all put together to get your reaction SMARTS!

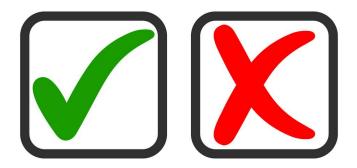
1. Allyl Acid 2. Primary 3. Aldehyde 4. Nitrile Amine



 $\label{eq:c:10} [C:10] = [C:9] - [C:10] = [C:10] - [C:1$

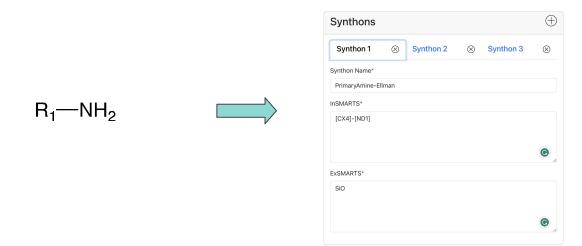
Transforming the Reaction into SMARTS: Filtering the Reagents

- Now it's time to think about what should and should not be allowed for each reactant
- These will be our Inclusion/Exclusion SMARTS



Transforming the Reaction into SMARTS: Inclusion/Exclusion SMARTS

- Inclusion SMARTS were already made earlier though making the reaction SMARTS!
- As for Exclusion SMARTS it can be and string of molecules that should not work as a reactant. Example: Anything with SiO in it



Transforming the Reaction into SMARTS: Testing Inclusion/Exclusion SMARTS

• You can use the <u>bbfilter module</u> in TLDR to test Inclusion/Exclusion SMARTS to see if you get sensible molecules or if you're filtering too many!

TLDR Start Jobs	Welcome MAKSIM -
bbfilter	
i Filter SMARTS pattern from BB list.	
Inputs	
included_SMARTS [CX4]-[ND1]	
excluded_SMARTS [Si]	
BB_library bb_50	~
HAC_range Min 6 Max 25	
Memo Enter memo for your job	
Submit Job	

Enumeration and Debugging!

• Once you feel more or less confidant with your Reaction and Inclusion/Exclusion SMARTS it is time to run some Enumeration to test your SMARTS!

-bash-4.	1\$ ls]
bb50	chbrbb1123	Filte	red NirM	ADE	s	cripts		
BU	chbrbb1223	innov	abb otav	abb	s	ialbb		
chbr	ChemistryComm	ons junk	READ	ME	s	ialbb1123		
chbrbb	emolbb	MADE	REAL	_Space_22	Q3_29B T	est		
-bash-4.	1\$ cd bb50]
-bash-4.	1\$ ls]
CC-10	CC-122.22	CC-140	CC-36.2	CC-4.4	CC-69.2	CC-80.2	CC-93.3	
CC-108	CC-122.23	CC-142	CC-39.1	CC-53	CC-69.3	CC-92.1	CC-94.1	
CC-122.1	1 CC-122.31	CC-142.1	CC-39.2	CC-68.1	CC-70.1	CC-92.2	CC-94.2	
CC-122.1	2 CC-122.32	CC-143	CC-4.1	CC-68.2	CC-70.2	CC-92.3	CC-94.3	
CC-122.1	.3 CC-122.33	CC-148	CC-4.2	CC-68.3	CC-70.3	CC-93.1		
CC-122.2	21 CC-138	CC-36.1	CC-4.3	CC-69.1	CC-80.1	CC-93.2		
-bash-4.	1\$							

Exporting Chemistry Commons Reaction

-Once you have made your reaction on Chemistry commons it's time to export your JSON containing the reaction into the command line!

-bash-4.1\$ cd /nfs/exj/Fe/scripts/common_db/
-bash-4.1\$ ls
<pre>export export_cc.bash log</pre>
-bash-4.1\$ bash export_cc.bash

Finding Reaction and Synthon JSONs

-Once run the export script you can find the Reactions.json and Synthons.json in the export directory

-bash-4.1% cd export		
[-bash-4.1\$ ls		
commons_20231002.sql	commons_20231208.sql	commons_20231216.sql
commons_20231009.sql	commons_20231209.sql	commons_20231217.sql
commons_20231017.sql	commons_20231210.sql	commons_20231218.sql
commons_20231031.sql	commons_20231211.sql	commons_20231219.sql
commons_20231115.sql	commons_20231212.sql	Reactions.json
commons_20231204.sql	commons_20231213.sql	Synthons.json
commons_20231206.sql	commons_20231214.sql	
commons_20231207.sql	commons_20231215.sql	

Picking library to Enumerate

- From experience it's best to pick a larger library to Enumerate to make sure there are enough molecules in the database, such as **bb50**.
- However, if you need quick results I recommend **Sialbb**.

-bash-4.1\$ cd /nfs/exj/Fe/Filtered/								
-bash-4.1\$ ls								
bb50	emolbb	innovabb	otavabb-v	sialbb				
chbrbb	indexed	MADE	REAL_Space_22Q3_29B	Test				
-bash-4.1\$								

Filtering Synthons

- Go into the library of your choosing and run the filtering command below for each synthon

[-bash-4.1\$ cd sialbb/ [-bash-4.1\$ ls sialbb_synthon_100_ReagentsForOlefination-Wittig sialbb_synthon_101_ReagentsForOlefination-HornerWadsworthEmmonsReagents sialbb_synthon_102_SecondaryAmines-SecAminesAliphaticPolycyclic sialbb_synthon_103_SecondaryAmines-secHet-Anilines sialbb_synthon_104_SecondaryAmines-Cyc-Aliphatic sialbb_synthon_105_SecondaryAmines-Cyc-Benzylic sialbb_synthon_106_SecondaryAmines-SecAminesAliphaticFluorinated sialbb_synthon_107_SecondaryAmines-diArylAmines sialbb_synthon_108_SecondaryAmines-Cyc-Het-benzylic sialbb_synthon_108_SecondaryAmines-Cyc-Anilines sialbb_synthon_109_SecondaryAmines-Cyc-Anilines sialbb_synthon_109_Anhydrides-AcyclicAnhydrides

sialbb_synthon_9_Acylhalides-N-Acylhalides
-bash-4.1\$ bash /nfs/exj/Fe/scripts/bb_wrapper.bash /nfs/exj/Fe/Filtered/indexed
/<catalog>/<catalog>.smi <id> /nfs/exj/Fe/scripts/common_db/export/Synthons.json

Enumeration: Part 1 of 3

1. First step is to source the environment:

-bash-4.1\$ source /nfs/home/ak87/exa/UCSF/SynthI/BESPOKE/arthor-env/bin/activate

2. Next step is to go into Library Enumeration Directory or create your own!

(arthor-env) -bash-4.1\$ cd /nfs/exj/Fe							
(arthor	-env) -bash-4.	1\$ ls				2 Hz	
bb50	chbrbb1123	Filt	ered N:	irMADE		scripts	
BU	chbrbb1223	inno	vabb o	tavabb		sialbb	
chbr	ChemistryComm	ons junk	RI	EADME		sialbb112	3
chbrbb	emolbb	MADE	RI	EAL_Space_2	2Q3_29B	Test	
(arthor-env) -bash-4.1\$ cd sialbb							
(arthor	-env) -bash-4.	1\$ ls					
CC-10	CC-122.22	CC-140	CC-4.1	CC-68.2	CC-70.2	CC-92.3	CC-94.3
CC-108	CC-122.23	CC-143	CC-4.2	CC-68.3	CC-70.3	CC-93.1	
CC-122.	11 CC-122.31	CC-36.1	CC-4.3	CC-69.1	CC-80.1	CC-93.2	
CC-122.	12 CC-122.32	CC-36.2	CC-4.4	CC-69.2	CC-80.2	CC-93.3	
CC-122.	13 CC-122.33	CC-39.1	CC-53	CC-69.3	CC-92.1	CC-94.1	
00 100	21 00 120	00 20 2	00 40	1 00 70 1	00 00 0	00 04 0	

Enumeration: Part 2 of 3

- Inside create a directory using the following Syntax: mkdir CC-<Reaction ID>
- Go inside the directory and run the first enumeration script:

(arthor-env) -bash-4.1\$ python3 /nfs/exj/Fe/scripts/Enumeratorsplit.py -i <Synthonfile1> <Synthon file2> -r CC-<Reaction ID> -o <library> -e /mnt/nfs/exj/Fe/Filtered/Test/EXAMPLE_Library_Enumerat ion_File_hg2_noProtected_logP350_HAC35_RotB7_V3_A-AK1.json -a /nfs/exj/Fe/scripts/common_db/expor t/Reactions.json

Enumeration: Part 3 of 3

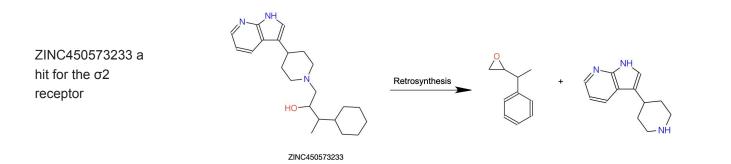
- Now after first script is finished running run the final script

(arthor-env) -bash-4.1\$ python3 /nfs/exj/Fe/scripts/Enumeratorcombine.py -i <pairing file> -r CC-<Reaction ID> -o <library> -e /mnt/nfs/exj/Fe/Filtered/Test/EXAMPLE_Library_Enumeration_File_hg2_ noProtected_logP350_HAC35_RotB7_V3_A-AK1.json -a /nfs/exj/Fe/scripts/common_db/export/Reactions.j son

Analog By Building Block Tool or AB3

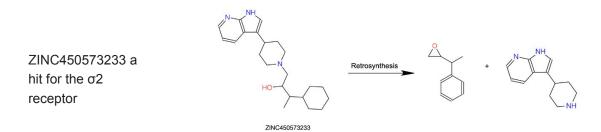
-Retrosynthesis is the process of taking a molecule and breaking it apart into building blocks

-We can use this concept to access a new area of purchasable chemical space!



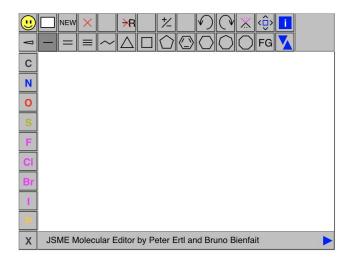
Analog By Building Block Tool or AB3

- 1. Use a sturdy retrosynthetic reaction to break down a molecule into building blocks
- 2. Find purchasable analogs of the building blocks by doing a similarity search
- 3. The respective building blocks can now be recombined with a forward reaction to find a final analog of the starting molecule





Home Reactions Contact



SMILES



Interface

-Draw your molecule of interest or just put in the SMILES in the input box

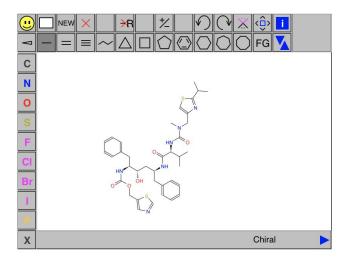


Home Reactions Contact

Interface

-Draw your molecule of interest or just put in the SMILES in the input box

-Set your TC value and hit submit!

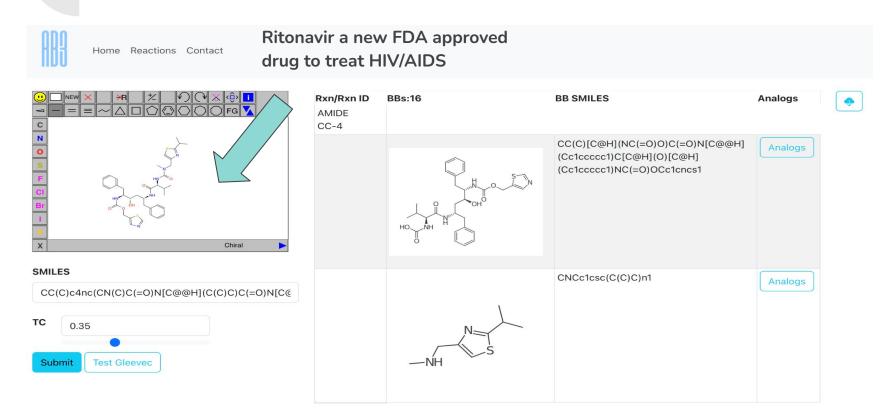


SMILES

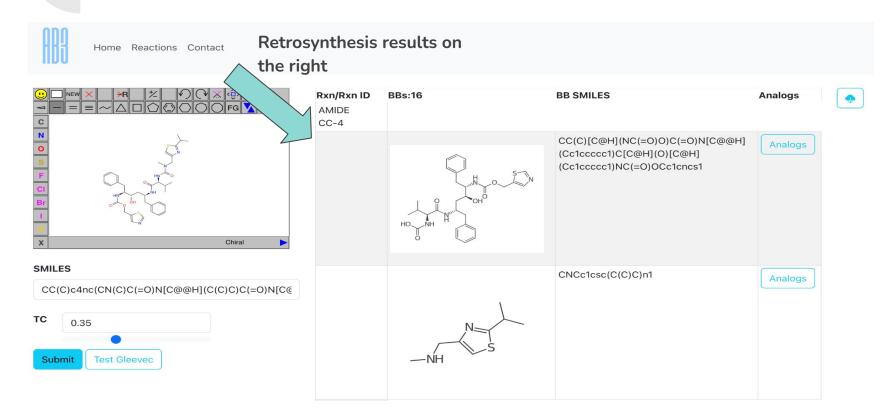
CC(C)c4nc(CN(C)C(=O)N[C@@H](C(C)C)C(=O)N[C@



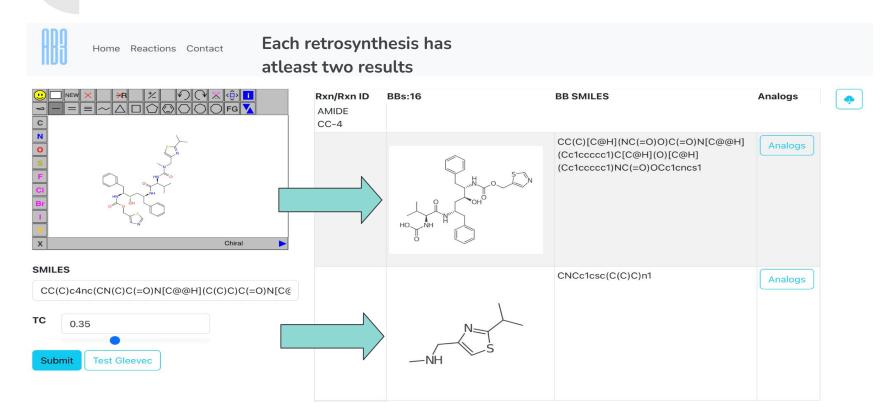
Finding Analogs of Novel Molecules: 1 of 8



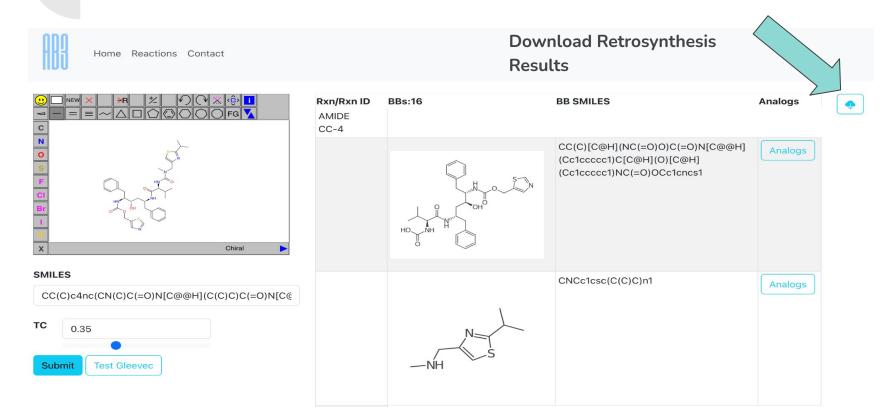
Finding Analogs of Novel Molecules: 2 of 8



Finding Analogs of Novel Molecules: 3 of 8



Finding Analogs of Novel Molecules: 4 of 8

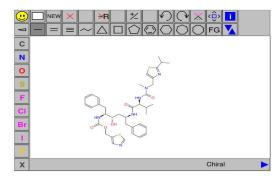


Finding Analogs of Novel Molecules: 5 of 8



Home Reactions Contact

Press Analogs to View Purchasable Analogs of BB

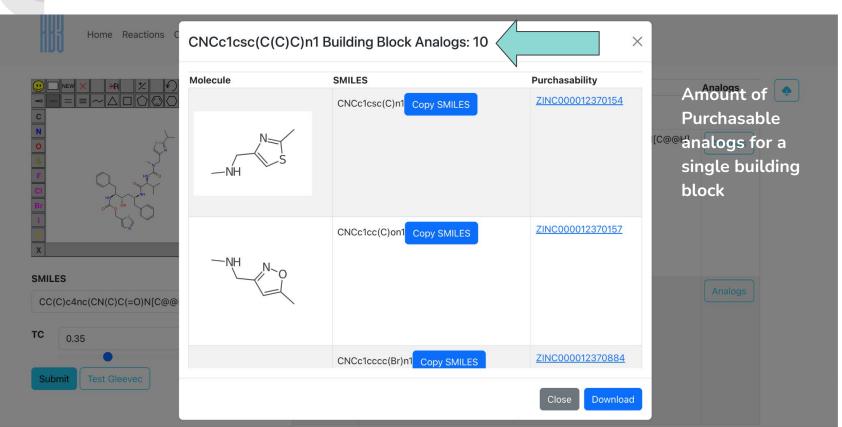


SMILES

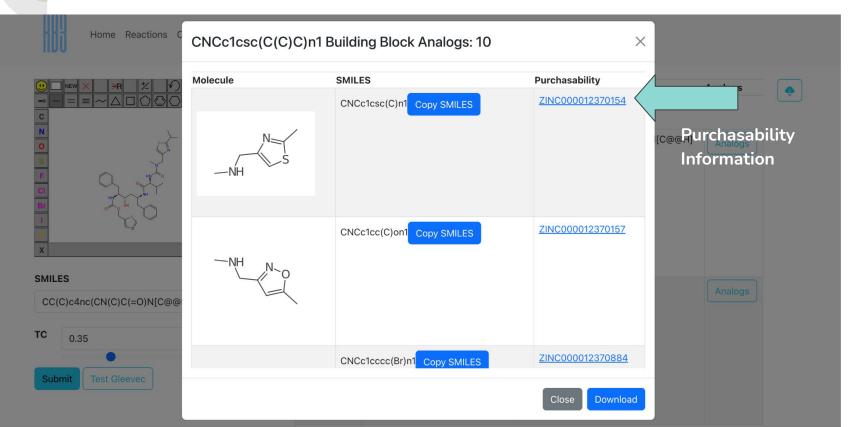


Rxn/Rxn ID AMIDE CC-4	BBs:16	BB SMILES	Analogs
		CC(C)[C@H](NC(=O)O)C(=O)N[C@@H] (Cc1ccccc1)C[C@H](O)[C@H] (Cc1ccccc1)NC(=O)OCc1cncs1	Analogs
	-NH NH	CNCc1csc(C(C)C)n1	Analogs

Finding Analogs of Novel Molecules: 6 of 8



Finding Analogs of Novel Molecules: 7 of 8



Finding Analogs of Novel Molecules: 8 of 8

