

MedChemica

CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

MCPairs Online AI tools to accelerate projects Generic training Course

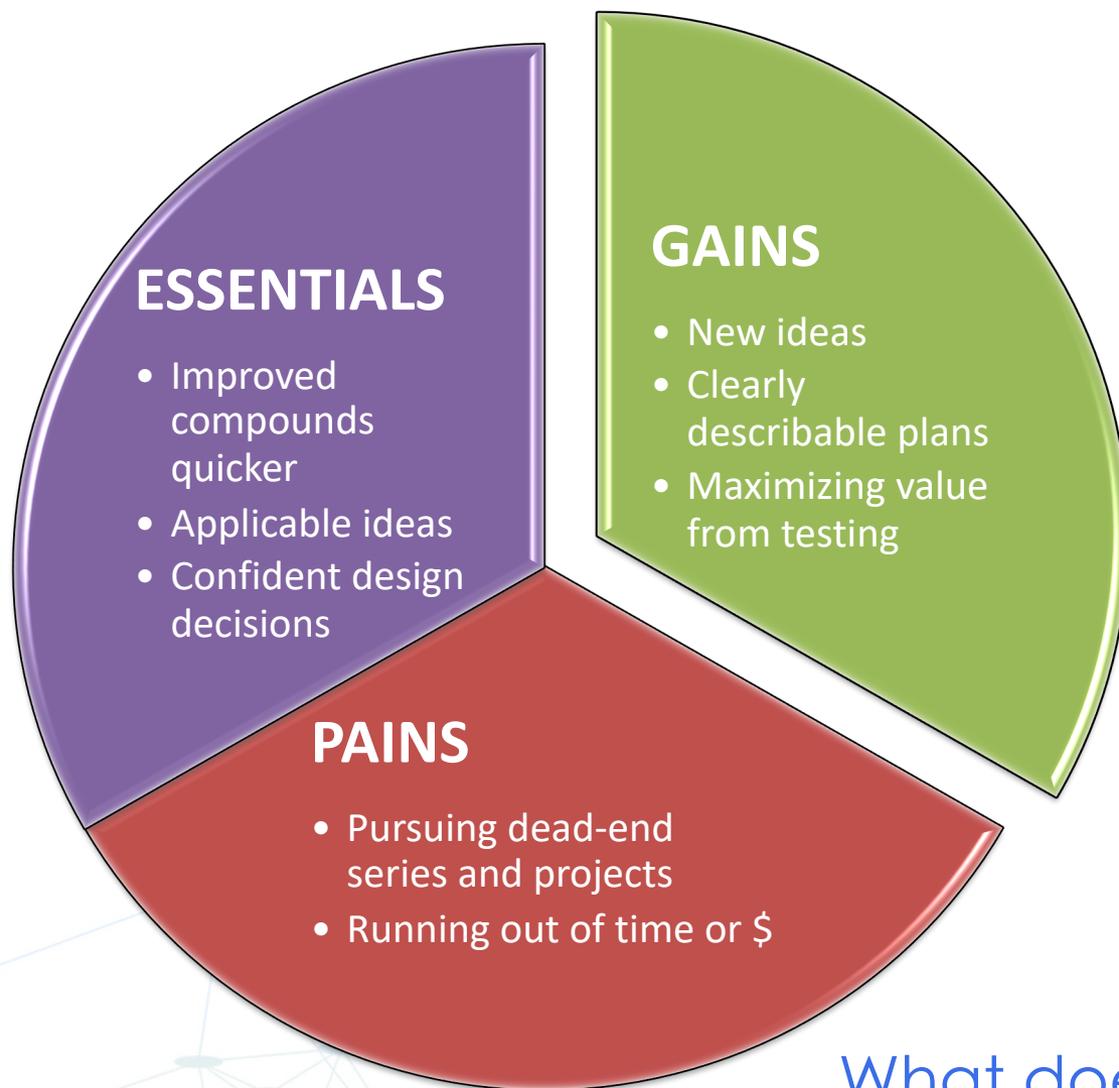
Twitter @MedChemica
Twitter @covid_moonshot

Twitter #BucketListPapers
www.medchemica.com/bucket-list/

Last update – 21st June 2020

For the training today:

- Going to present Powerpoint slides, then live demos with MCPairs Online in Google Chrome
- What MCPairs gives to the experienced medicinal chemist?
- From SAR to Matched Molecular Pair Analysis
- How MCPairs works?
- SAR analysis with Compound-To-Pairs
- Testing your ideas with SpotDesign™
- **Ideas and new directions with RuleDesign**



What does a project need?

Training will use a literature example

Journal of
**Medicinal
Chemistry**

Article

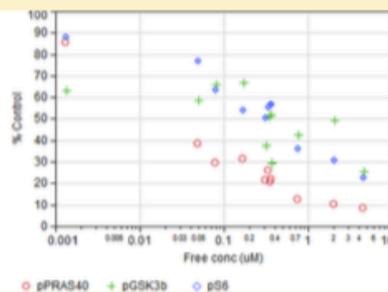
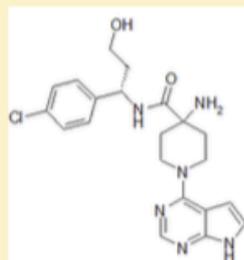
pubs.acs.org/jmc

Discovery of 4-Amino-*N*-[(1*S*)-1-(4-chlorophenyl)-3-hydroxypropyl]-1-(7*H*-pyrrolo[2,3-*d*]pyrimidin-4-yl)piperidine-4-carboxamide (AZD5363), an Orally Bioavailable, Potent Inhibitor of Akt Kinases

Matt Addie, Peter Ballard, David Buttar, Claire Crafter, Gordon Currie, Barry R. Davies, Judit Debreczeni, Hannah Dry, Philippa Dudley, Ryan Greenwood, Paul D. Johnson, Jason G. Kettle,* Clare Lane, Gillian Lamont, Andrew Leach, Richard W. A. Luke, Jeff Morris, Donald Ogilvie,† Ken Page, Martin Pass, Stuart Pearson, and Linette Ruston

Oncology iMed, AstraZeneca, Alderley Park, Macclesfield SK10 4TG, United Kingdom

Supporting Information



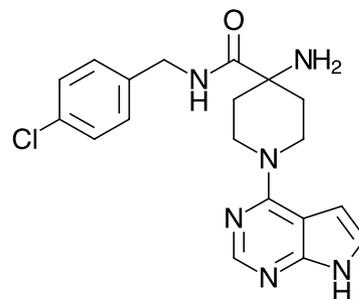
ABSTRACT: Wide-ranging exploration of analogues of an ATP-competitive pyrrolopyrimidine inhibitor of Akt led to the discovery of clinical candidate AZD5363, which showed increased potency, reduced HERG affinity, and higher selectivity against the closely related AGC kinase ROCK. This compound demonstrated good preclinical drug metabolism and pharmacokinetics (DMPK) properties and, after oral dosing, showed pharmacodynamic knockdown of phosphorylation of Akt and downstream biomarkers in vivo, and inhibition of tumor growth in a breast cancer xenograft model.

<http://dx.doi.org/10.1021/jm301762v>

Capivasertib (AZD5363) – AKT inhibitors

AKT pIC₅₀ 7.89 (13nM)
 LogD 2.9
 Sol (pSol) -5.3 (5 – 150µM)
 hERG pIC₅₀ 5.2 (5.2 µM)

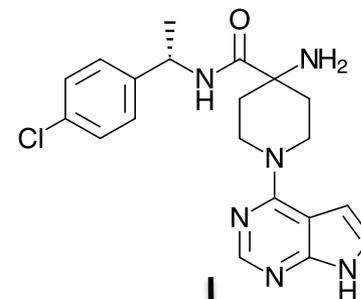
Potent enough
hERG and improved
solubility



CHEMBL598194

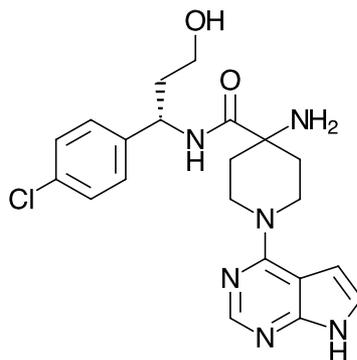
Δ pIC₅₀ + 0.2
 Δ LogD - 0.2 (?)
 Δ pSol + 0.5
 Δ hERG ~ 0.2

CHEMBL2325742

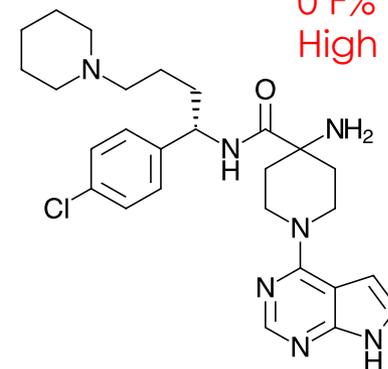


Δ pIC₅₀ ~ 0.0
 Δ LogD - 0.3
 Δ pSol + 1.4
 Δ hERG - 0.5

0 F%
 High Cl



CHEMBL2325741

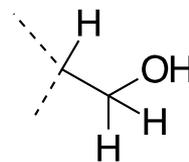
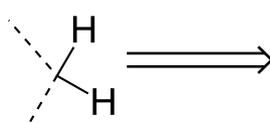


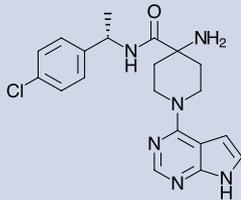
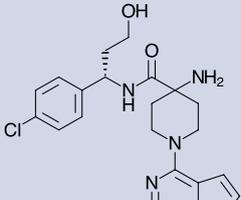
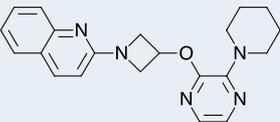
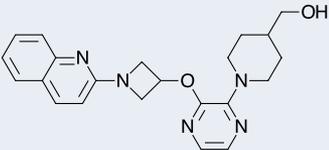
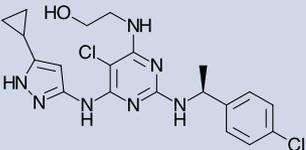
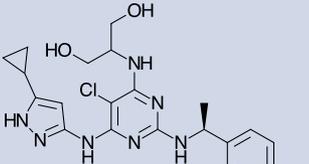
CHEMBL2325729

AKT pIC₅₀ 8.4 (3nM)
 LogD 2.5
 Sol (pSol) -3.1 (780µM)
 hERG pIC₅₀ <4.0 (>100 µM)

Potent
No hERG and improved
solubility

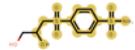
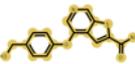
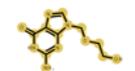
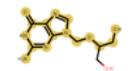
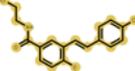
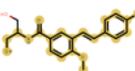
From SAR to MMPA.....



A	B	pSol A (μM)	pSol B (μM)	ΔpSol
 CHEMBL2325742	 CHEMBL2325741	- 4.1 (77 μM)	- 3.1 (870 μM)	1.0
 CHEMBL3356658	 CHEMBL218767	- 6.0 (1.0 μM)	- 3.7 (178 μM)	2.3
 CHEMBL456802	 CHEMBL456322	-5.7 (2.0 μM)	- 4.1 (82 μM)	1.6
				3 pairs +ve Sol Median 1.6

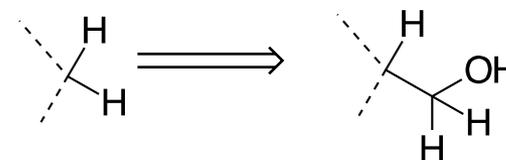
MCPairs Rule finder required 6 matched pairs for 95% confidence

From SAR to MMPA.....

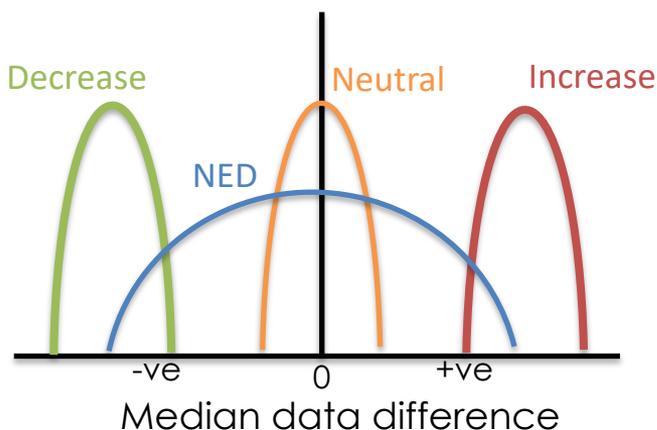
compound name A	compound name B	Depiction A	Depiction B	Aqueous Solubility at pH 7.4 solubility [CHEMBL2362975] unit	=	Aqueous Solubility at pH 7.4 solubility [CHEMBL2362975] measurement A	=	Aqueous Solubility at pH 7.4 solubility [CHEMBL2362975] measurement B	Aqueous Solubility at pH 7.4 solubility [CHEMBL2362975] measurement delta
CHEMBL104459	CHEMBL316800			log10(M)	=	-1.6042	=	-1.2971	0.3071
CHEMBL118022	CHEMBL115462			log10(M)	=	-4.63875	=	-4.3343	0.30445
CHEMBL161956	CHEMBL165547			log10(M)	=	-9.9586	=	-9.7447	0.2139
CHEMBL165864	CHEMBL166093			log10(M)	=	-10.699	=	-9.6778	1.0212
CHEMBL184	CHEMBL182			log10(M)	=	-5.2353	=	-4.7115	0.5238
CHEMBL184521	CHEMBL439660			log10(M)	=	-5.0168	=	-3.3704	1.6464
CHEMBL1949786	CHEMBL1949790			log10(M)	=	-4.31675	=	-3.1588	1.15795

Actual Rule from MCPairs
Endpoint:
Aqueous Solubility at pH 7.4
[CHEMBL2362975]

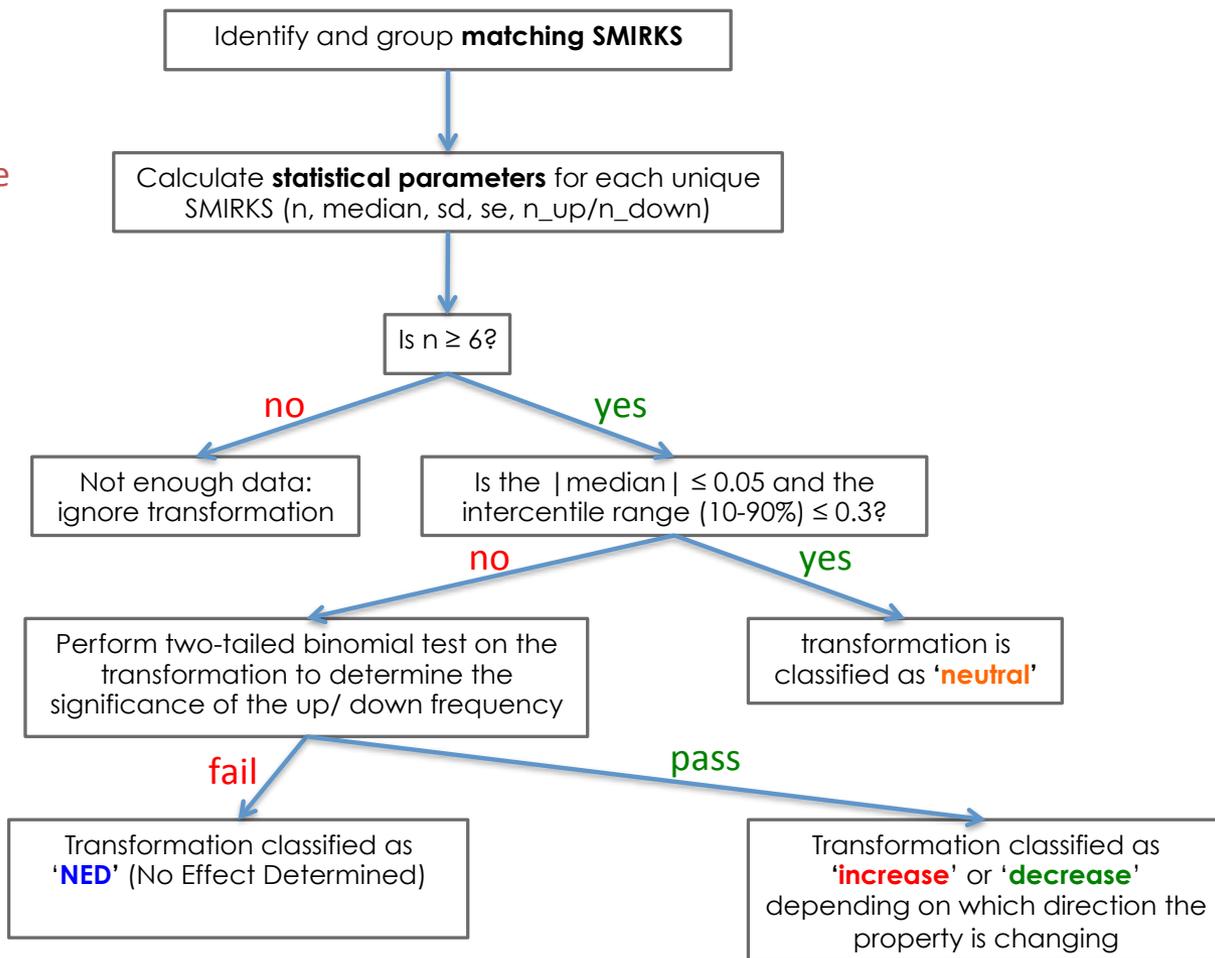
n-qual	69
n-qual-up	47
n-qual-down	21
median Δ pSol	0.26
std dev +/-	0.636



Rule selection

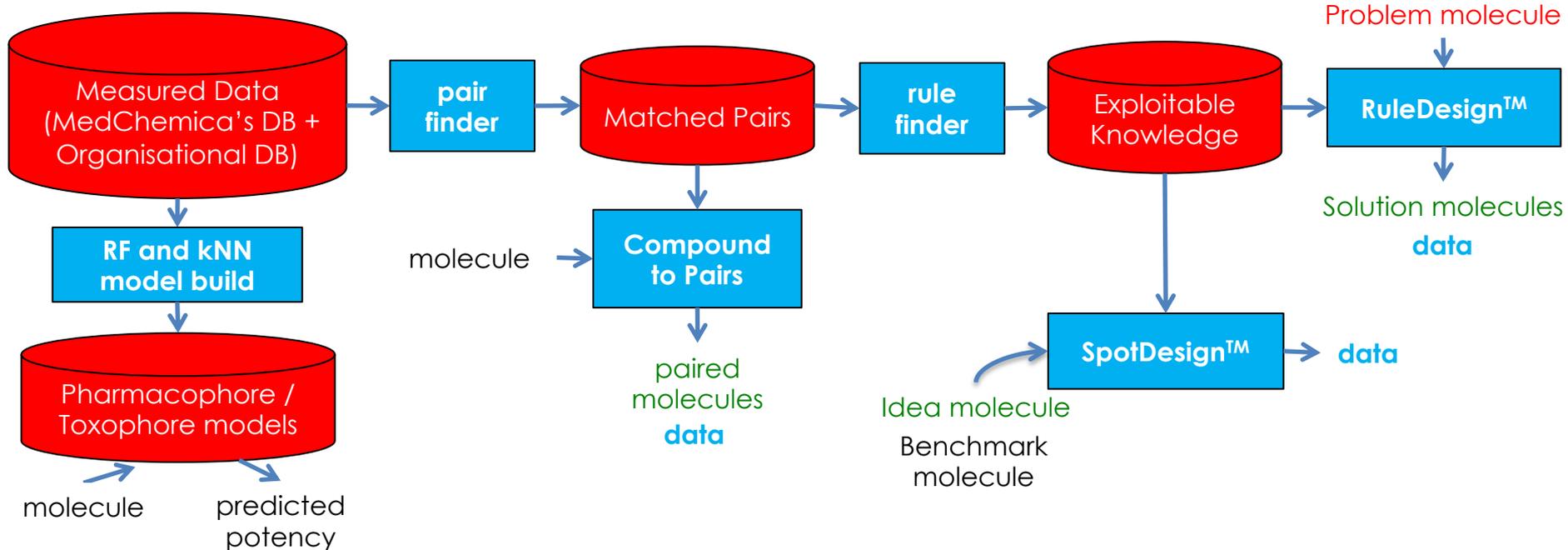
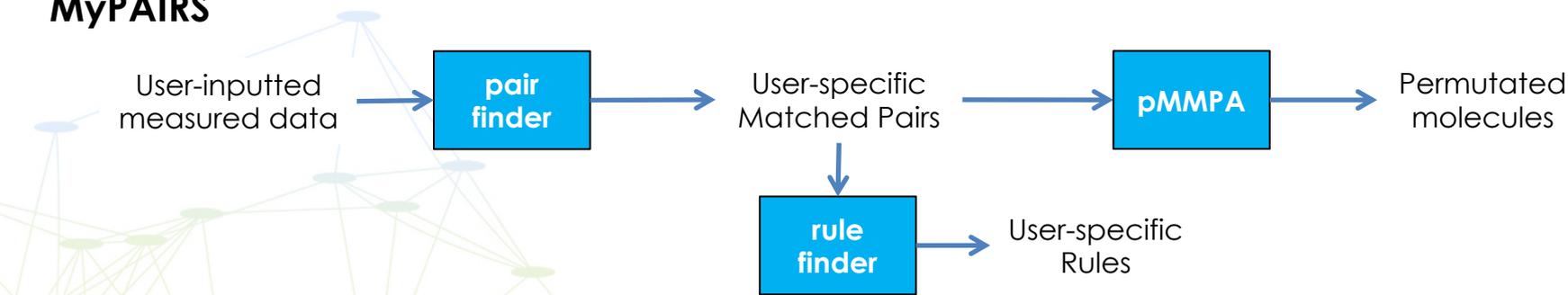


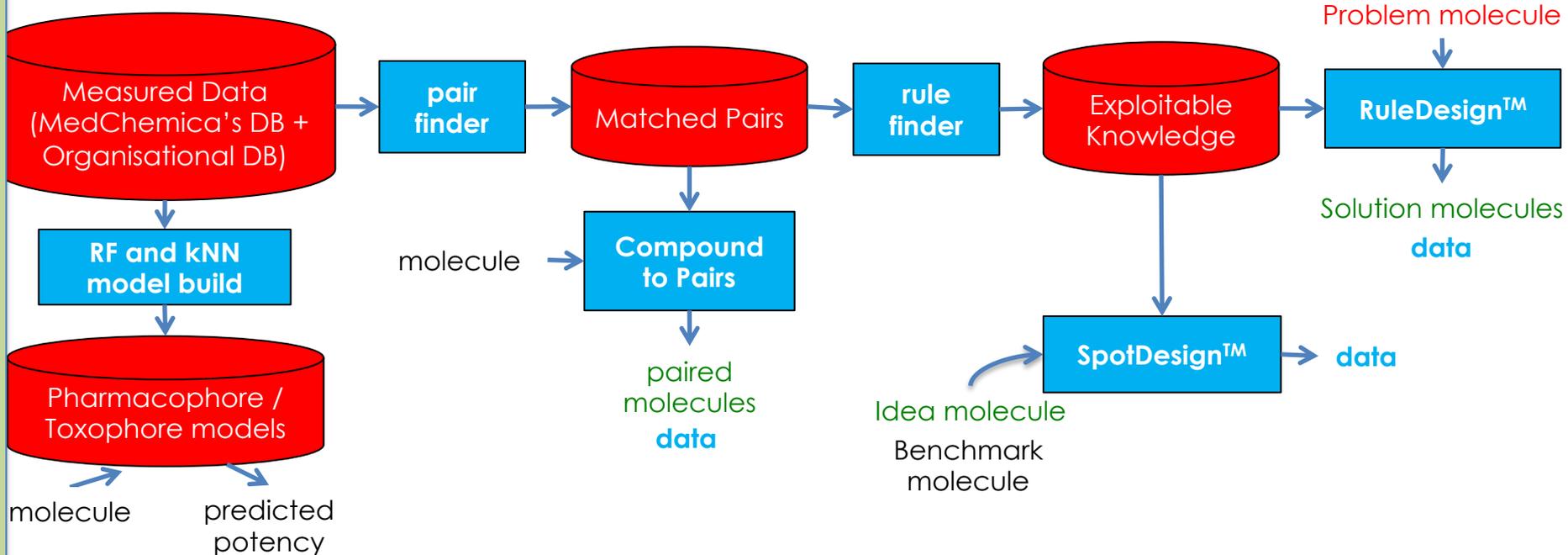
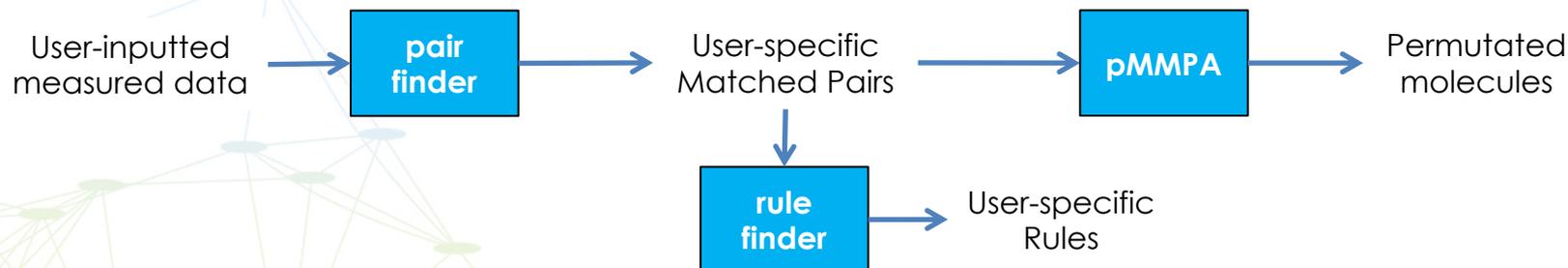
- No assumption of normal distribution
- Manages 'censored' = qualified / out-of-range data



Leach et al. *J. Chem. Inf. Model.* **2017**, **57**, 2424 - 2436

Exploiting data derived ADMET Knowledge

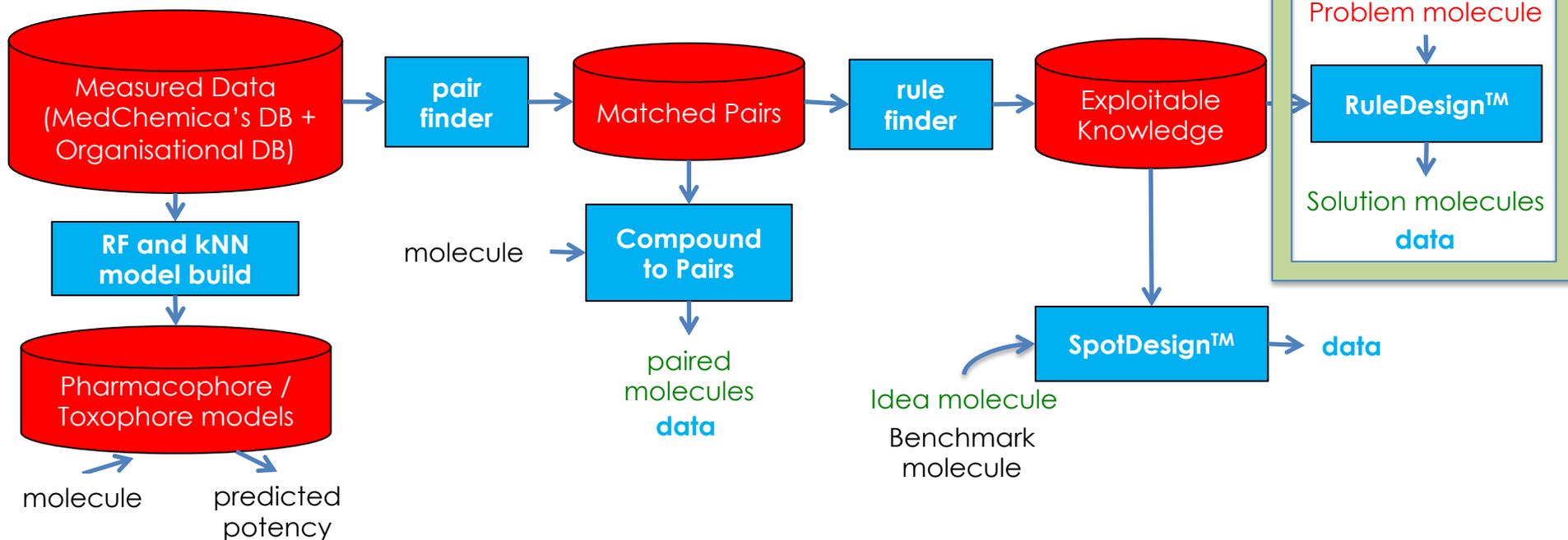
DATABASE TOOLS

MyPAIRS


Exploiting data derived ADMET Knowledge **MedChemica****DATABASE TOOLS****MyPAIRS**

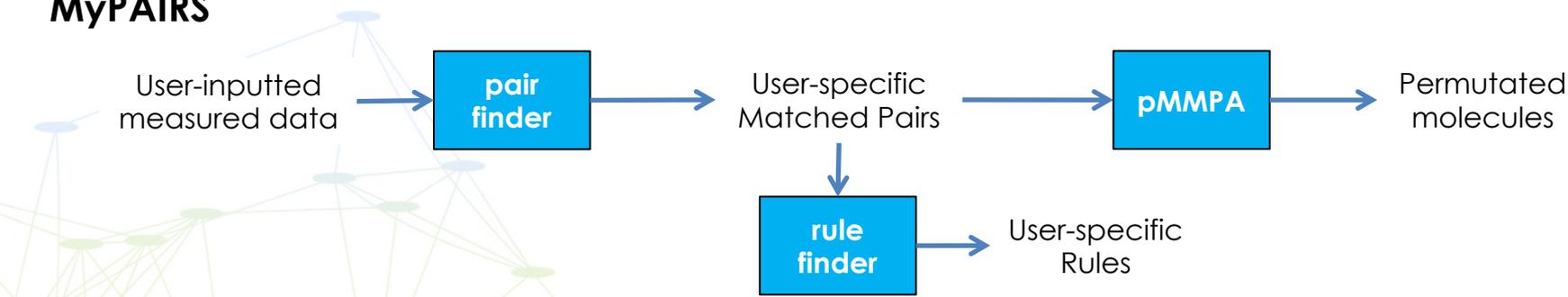
Exploiting data derived ADMET Knowledge

MedChemica
CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

DATABASE TOOLS



MyPAIRS



Exercise – New Idea based on rules (Show me the Gems'!)

- Input chemical name for lead molecule CHEMBL2325997
 - Choosing the property to improve (setting the Goal)
 - Find Phys_Prop -> Solubility (Broad Goal)
 - Filtering the results down by substructure and phys-props
 - Export the results – controlling the output
 - Exploring results in Excel
 - what do the numbers mean?
 - How do I use these to make a knowledge based decision?
 - Drilling back to the original source (looking at the matched pairs)

Matched Pairs

Icon indicates that these tools use the central database of knowledge

Access to MCPairs Database Matched Molecular Pairs Search



RULEDESIGN™

Formerly Known as Compounds From Rules

Submit compound(s), enumerate products using Rules from the central database



SPOTDESIGN™

Submit a reference compound and explore your ideas seeking support from the central database



COMPOUND TO PAIRS

Search the central database of the matched pairs of a compound



COMPOUND TO MEASUREMENTS

Search the central database for the current aggregated data of a compound

Rule Design

MedChemica
Rule Design
Modules My Account

Input Molecule*

Single
Multi

Compound Name:

SMILES*:

Goal*

Direction*:

Endpoint*:

Select

Specificity:

Advanced Filters

Molecular charge:

HBA:

HBD:

CLogP:

RMM:

PSA:

Substructure Lock:

Input Compound Phys. Properties

Anions	0	Cations	1
HBA	5	HBD	3
LogP	2.79	PSA	99.93
RMM	398.90		

GO!

Submit

Copied from SpotDesign

Check the direction

Goal from SpotDesign

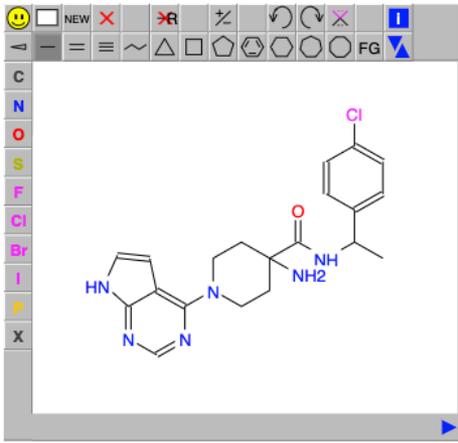
Rule Design

Input Molecule*

Single | **Multi**

Compound Name:

SMILES*:



Goal*

Direction*:

Endpoint*:
 Select

Specificity:

Advanced Filters

- Molecular charge:**
- HBA:**
- HBD:**
- CLogP:**
- RMM:**
- PSA:**
- Substructure Lock:**

Input Compound Phys. Properties

Anions	0	Cations	1
HBA	5	HBD	3
LogP	2.79	PSA	99.93
RMM	398.90		

Focus your design

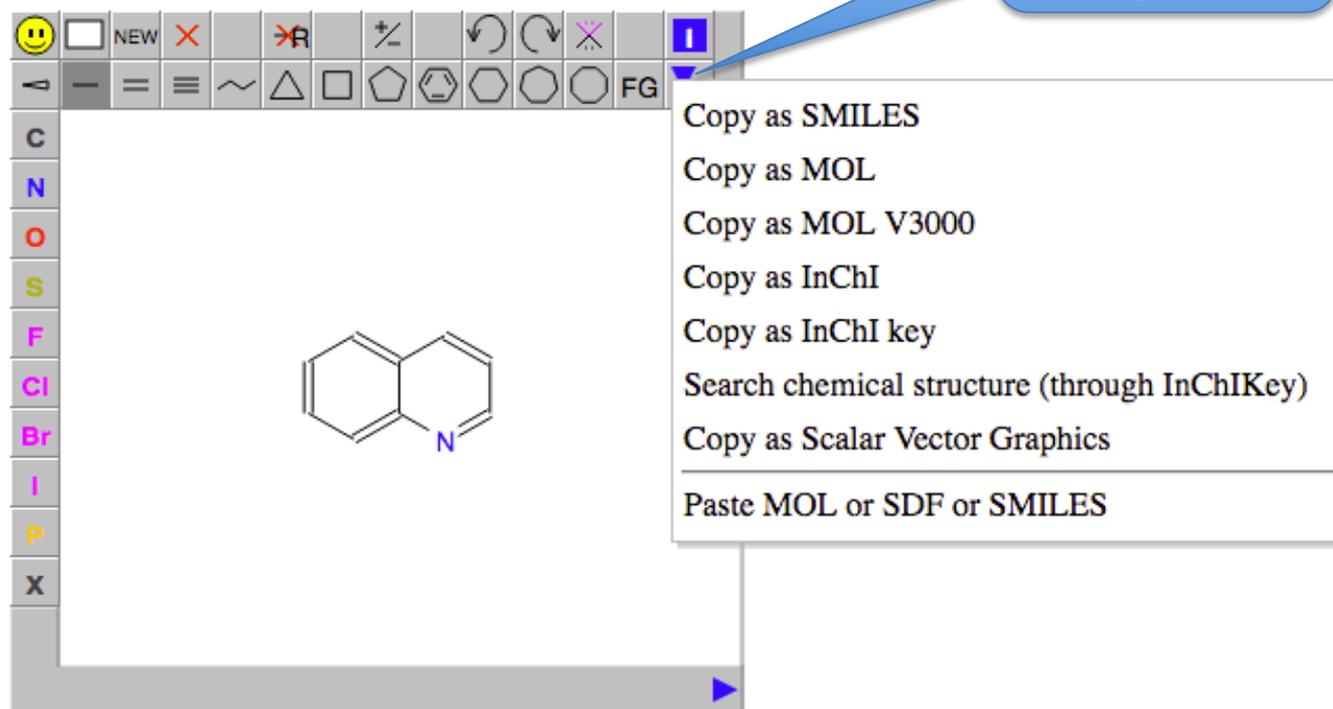
Save Your results

Submit

#	Timestamp	Status	Number of Products	
0	Feb 1, 2021, 11:46:13 AM	Complete	612	<div style="border: 1px solid #007bff; border-radius: 5px; padding: 2px 5px; background-color: #007bff; color: white; font-weight: bold; font-size: 0.8em;"> Save </div>

TopTip : Other methods to get a structure in

- **Pasting structures in and out**



The image shows a screenshot of a chemical software interface. The main window displays a chemical structure of a benzimidazole derivative. The interface includes a toolbar at the top with various icons for file operations, editing, and drawing. A vertical toolbar on the left contains element selection buttons for C, N, O, S, F, Cl, Br, I, P, and X. A context menu is open over the structure, listing several options for copying and pasting the structure. A blue callout box points to a blue triangle icon in the toolbar, with the text "Click blue triangle square".

Click blue triangle square

- Copy as SMILES
- Copy as MOL
- Copy as MOL V3000
- Copy as InChI
- Copy as InChI key
- Search chemical structure (through InChIKey)
- Copy as Scalar Vector Graphics

- Paste MOL or SDF or SMILES

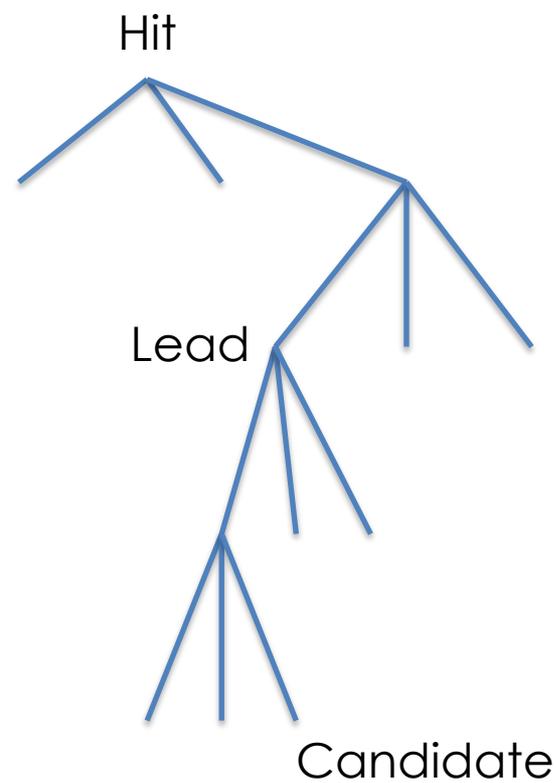
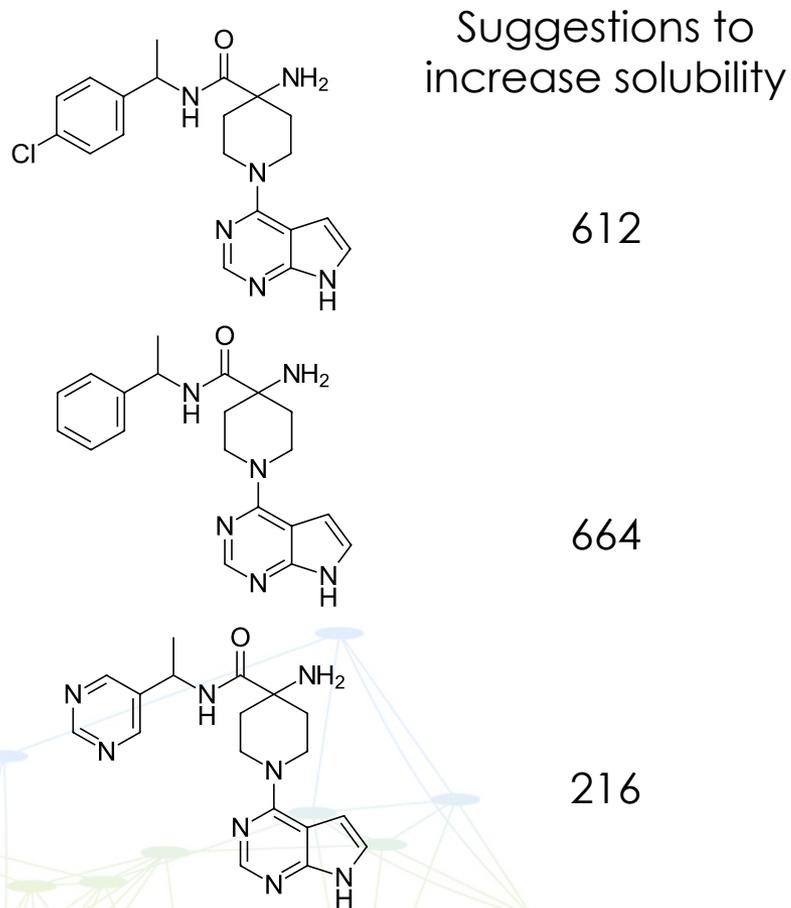
TopTip : Other methods to get a structure in

- **Upload a file of multiple compounds**
 - File can contain compounds + measurements or just compounds

The screenshot shows a web interface for uploading multiple compounds. The main heading is "Input Molecule*". Below it, there are two tabs: "Single" and "Multi", with "Multi" selected. A blue button labeled "+ Add Compounds" is visible. A dashed box highlights the text "Available file types: Compound Data file or SMILES file" and "Maximum 50 compounds". A callout box points to this text, stating "Click here to input a multi-compound file". Another callout box points to the "Available file types" text, stating "Hover over for info on the different file formats accepted". Below this, a pop-up box titled "Compound Data file" provides details: "Accepted Extensions: .txt .tsv .csv", "Files are tab (\t) or comma (,) separated, and can be submitted without headers but the order of the columns must be fixed as: compound_name, compound_structure, measurement, qualifier. Measurement is on the log scale (e.g. pIC50), and the qualifier is either ">", "<" or "=".", and "Check the documentation for more details." To the right, there is a "Direction*" section with a dropdown menu set to "Increase" and a "Select" button. Below that, there is a "Specificity" checkbox and a "solubility" filter with a close button. At the bottom, there is an "Advanced Filters" section with several checkboxes: "Molecular charge:", "HBA:", "HBD:", "CLogP:", "RMM:", "PSA:", and "Substructure Lock:".

What molecule should I put in?

The more specific the 'seed' molecule, the fewer rules will apply to it



Filtering down the output

- Adding a locked structure (eg pharmacophore)
 - Highlighting the atoms to changes
 - Blocking H substitution
 - Variable atoms
 - 'Complete SMARTS patterns'
 - Also see Tips'n'Tricks document for a detailed work flow
- Charge control
Filtering by the charge of the final compounds (neutral, acidic, basic)
- Simple Lipinski property filters



Highlight the part you want to change

- Add a locked structure with Substructure-Lock (eg pharmacophore)

The image shows two sequential screenshots of the 'Sub-structure Lock' dialog box. In the first screenshot, the 'Enable Highlight Tool' checkbox is checked, and a blue circle is visible on the highlighter tool icon in the toolbar. A text box with the instruction 'Enable tool, click Blue circle' has arrows pointing to the checkbox and the blue circle. The chemical structure in the main window is not yet highlighted. In the second screenshot, the same dialog box is shown, but the blue circle is no longer present, and the phenyl ring of the chemical structure is highlighted in cyan. A large blue arrow points from the first screenshot to the second, indicating the progression of the action.

Sub-structure Lock

SMARTS:
CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)

Enable Highlight Tool

Result:

Sub-structure Lock

SMARTS:
[CH3:1][CH:1](NC(=O)C3(N)CCN(c1ncnc2[nH]ccc12)CC3)[

Enable Highlight Tool

Result:
[cH]1[cH][nH][c]2[c]1[c]([n][cH][n]2)[N]3[CH2][CH2][C]([C

Enable tool, click Blue circle

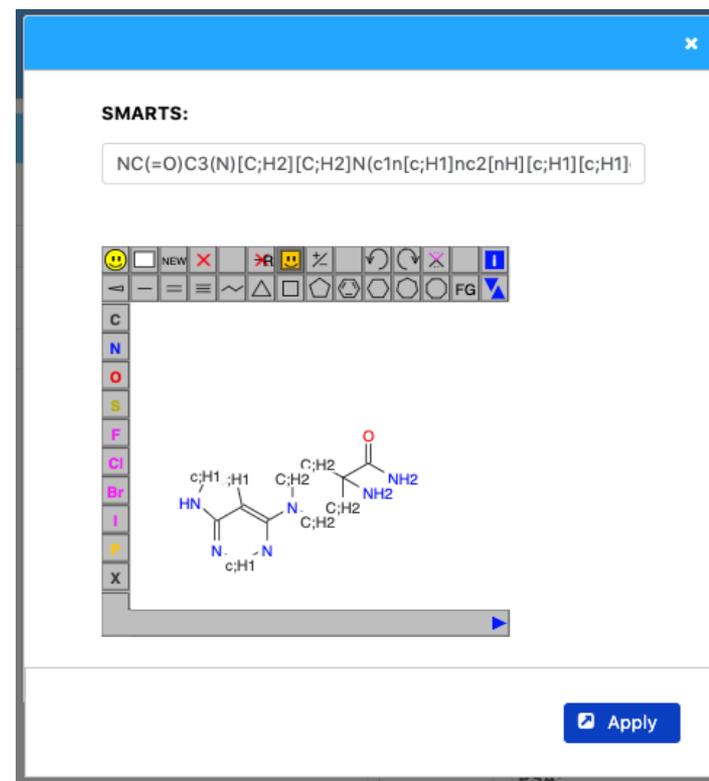
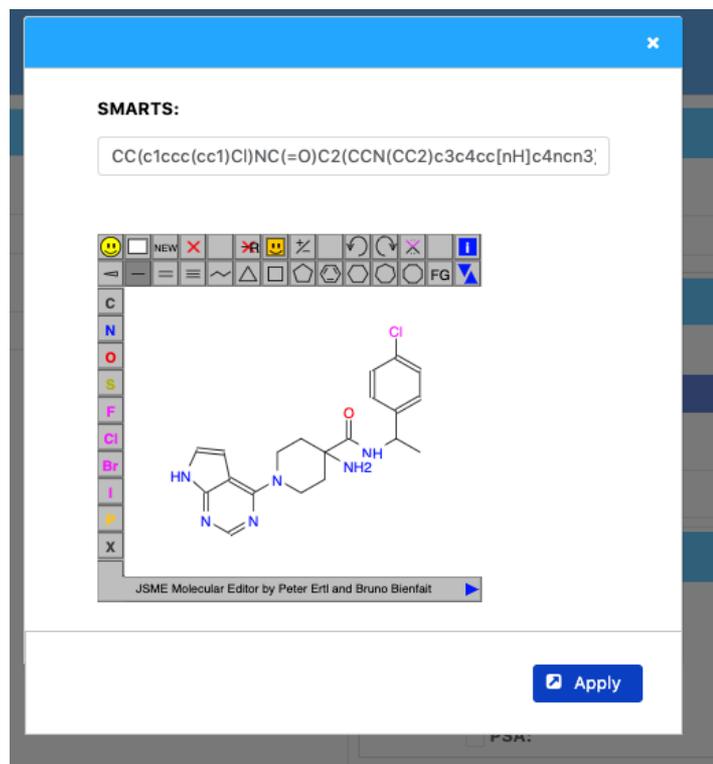
Click atoms with highlighter tool

Alternatively lock out specific hydrogens



Draw

- Add a locked structure with Substructure-Lock (eg pharmacophore)



Delete to groups you want to change
Lock out hydrogens to stop substitution

Defining H-atoms to block out

Blocking H

Atom/Bond Query

Atom type : Any Any except Halogen

Or select one or more from the list :

C N O S P F Cl Br I

Number of hydrogens : 1

Number of connections : Any (H's don't count.)

Atom is : **Aromatic** Nonaromatic Ring Nonring

Bond is : Any Aromatic Ring Nonring

c;H1 Reset Close

Variable atoms

Atom/Bond Query

Atom type : Any Any except Halogen

Or select one or more from the list :

C **N** O S P F Cl Br I

Number of hydrogens : Any

Number of connections : Any (H's don't count.)

Atom is : **Aromatic** Nonaromatic Ring Nonring

Bond is : Any Aromatic Ring Nonring

c,n Reset Close

Close the box, then click on atoms in the picture to change their filter types



Sub-structure Lock

SMARTS:

NC(=O)C3(N)[C;H2][C;H2]N(c1n[c;H1]nc2[nH][c;H1][c;H1]c12)[C;H2][C;

Draw

Combined Filters

Goal Selection

Charge and Calc
Props

Sub-Structure Lock

History Box

1st run Solubility ideas	612
2 nd run with Sub-Structure Lock	338
3 rd Run no neutral molecules	1
4 th with single cation	123

MedChemica

CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

Goal*

Direction*:
Increase

Endpoint*:
solubility ✕ Select

Specificity:

Advanced Filters

Molecular charge: cation —|—○

HBA:
 HBD:
 CLogP:
 RMM:
 PSA:
 Substructure Lock:

NC(=O)C3(N)[C;H2][C;H2]N(c1n[c;H1]nc2[nH][c;H1][c;H1]c12)[C;H2][C;H2]3 ✕ Draw

Input Compound Phys. Properties

Anions	0	Cations	1
HBA	5	HBD	3
LogP	2.79	PSA	99.93
RMM	398.90		

Submit

	Number of Products	
612		Save
338		Save
1		Save
123		Save

Common Goal Patterns

Example	Goal	Directions
1	Solubility	increase
2	Metabolism	decrease
3	HLM_Clint_generic_uL.min-1.mg-1	decrease
4	Absorption	increase
5	Caco-2_A_to_B_perm_CHEMBL614058	increase
6	Transporter	decrease
7	MDCK-dog-perm-log(ER)-dog	decrease
8	Fraction unbound	increase
9	PPB_hum_log(proportion_Free)	increase
10	LogD_TM	decrease (usually)
11	Anti-target_cardiac	decrease
12	hERG_hum_inhib_pIC50	decrease
13	Anti-target_DDI	
14	CYP_inhibs_3A4_pIC50_hum	
15	CCR5_human_pIC50_CHEMBL274	decrease (reduce off target)

How many pairs? – deeper Goal setting

Specific Goal settings

'All Rules'

- all of the Increase and Decrease Rules for all datasets
- warning output can be large
- not suitable for Excel spreadsheet

'Hit to Lead'

- most frequent transformations chemists perform

Non-rules transformations from pair counts

'Min 3 pair Trans'

- all transformations with 3 OR MORE matched pairs

'Min 6 pair Trans'

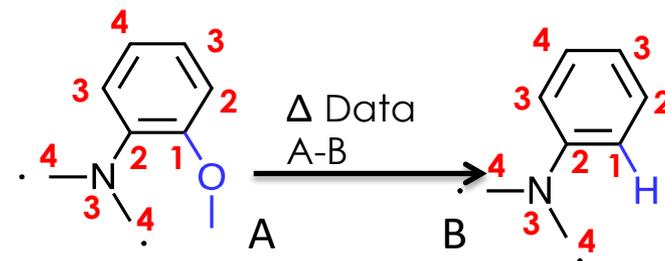
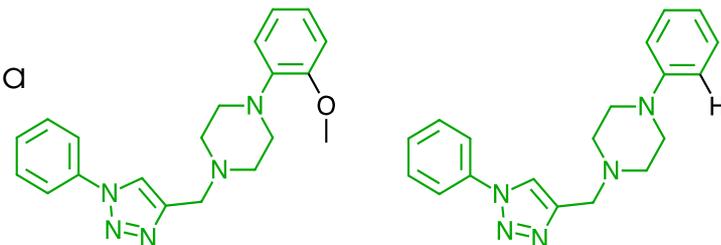
- all transformations with 6 OR MORE matched pairs
- Actually Increase, Decrease, Neutral and NED

The screenshot shows the 'Goal*' configuration window. The 'Direction*' dropdown menu is open, displaying 'Increase' as the selected option. Below the dropdown, the 'All Rules' section is visible, with 'All Rules' selected. A 'Select' button is present to the right. The 'Advanced' section is partially visible below.

The screenshot shows the 'Goal*' configuration window. The 'Direction*' dropdown menu is open, displaying '3 Pair Trans Only' as the selected option. Below the dropdown, the 'Min Pair Rules' section is visible, with 'Min 3 Pair Trans' selected. A 'Select' button is present to the right. The 'Advanced' section is partially visible below.

How much environment is captured?

- **Matched Molecular Pairs** – Molecules that differ only by a particular, well-defined structural transformation
- **Capture the change and environment** – MMPs can be recorded as transformations from A → B
- **Statistical analysis to define “medicinal chemistry rules”**
Defined transformations with high probability of improving properties of molecules
- High specificity are environment size 3 and 4, low is 1 and 2.



Griffen, E. *et al.* *J. Med. Chem.* **2011**, **54**(22), pp.7739-7750.

Specificity of Transformation

Specificity setting

'Specificity slider'

- Range 1 to 4
- 4 being the most specific
- setting 3-4 (or 3,4) only apply the most specific rules.
- The number refers to the amount of environment captured with the change in the group

Useful combinations

- Min 3 pairs Trans with high specificity
 - with Kinase – best rules for kinase potency
- All Rules with high specificity
 - The most specific ADMET solutions

Goal*

Direction*:
Min 3 Pair Trans

Endpoint*:
solubility x Select

Specificity: [3 - 4]

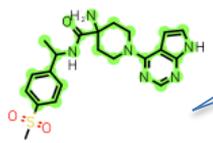
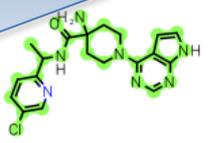
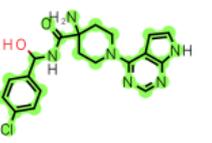
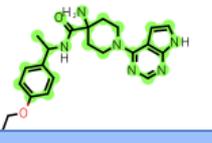
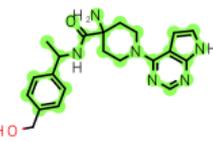
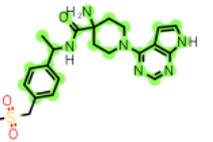
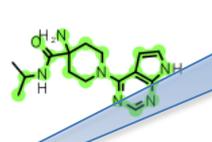
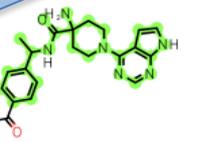
Click the Blue circle to see run results in Browser

#	Timestamp	Status	Number of Products	Actions
0	Jul 26, 2021, 2:56:50 PM	Complete	108	<input type="checkbox"/> Save

Products: No Sort

Hover mouse over to see larger structures

Results are paginated

<input type="checkbox"/> CHEMBL2325990  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_2  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_3  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_4  <input type="button" value="Goals"/>
<input type="checkbox"/> 1627307787759_5  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_6  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_7  <input type="button" value="Goals"/>	<input type="checkbox"/> 1627307787759_8  <input type="button" value="Goals"/>

1 2 3 4 5

Sort the results – select option here

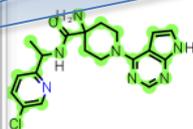
Submit

#	Timestamp	Status	Number of Products	Actions
0	Jul 26, 2021, 2:56:50 PM	Complete	108	Save

Products: No Sort Save Selection

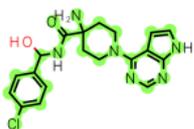
- LogP
- PSA
- RMM
- Aq_Solubility_pH_7.4_[CHEMBL2362975]
- Aq_Solubility_generic_log(M)
- Aq_Solubility_pH_7.4_[CHEMBL612558]

7787759_2



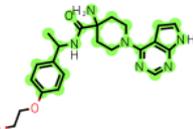
[Goals](#)

1627307787759_3



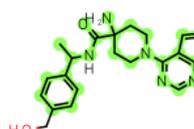
[Goals](#)

1627307787759_4



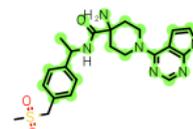
[Goals](#)

1627307787759_5



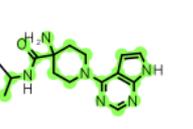
[Goals](#)

1627307787759_6



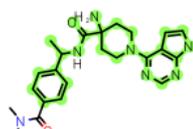
[Goals](#)

1627307787759_7



[Goals](#)

1627307787759_8



[Goals](#)

1 2 3 4 5

Select one of the dataset chosen by the Goal

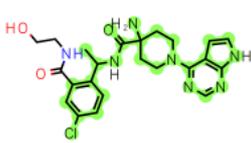
Selected Dataset / Goal

Select sort order – Median Change is the size of the change expected

#	Timestamp	Status	Number of Products	Actions
0	Jul 26, 2021, 2:56:50 PM	Complete	108	Save

Products: Sort: [Save Selection](#)

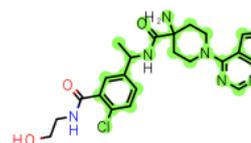
1627307787759_74



1.15
Med. Change

[Goals](#)

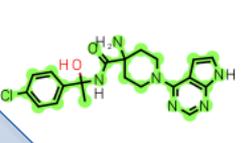
1627307787759_73



1.15
Med. Change

[Goals](#)

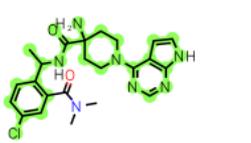
1627307787759_70



1.09
Med. Change

[Goals](#)

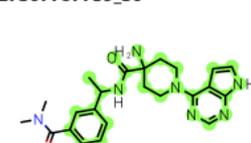
1627307787759_31



1.00
Med. Change

[Goals](#)

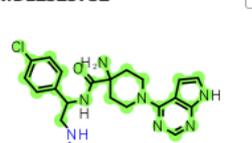
1627307787759_30



1.00
Med. Change

[Goals](#)

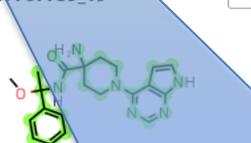
CHEMBL2325732



0.82
Med. Change

[Goals](#)

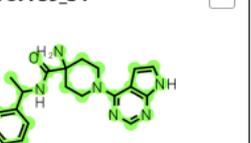
1627307787759_19



0.78
Med. Change

[Goals](#)

1627307787759_34



0.63
Med. Change

[Goals](#)

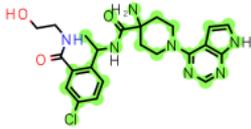
1 2 3

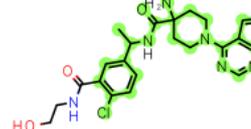
Select preferred example to export

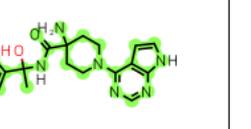
Click the Goal view to see more detailed data

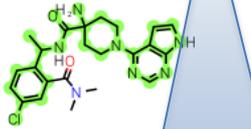
#	Timestamp	Status	Number of Products	Actions
0	Jul 26, 2021, 2:56:50 PM	Complete	108	Save

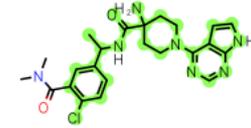
Products: Aq_Solubility_pH_7.4_[CHEMBL2362975] | Median Change | [Save Selection](#)

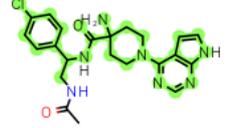
1627307787759_74  1.15 Med. Change [Goals](#)

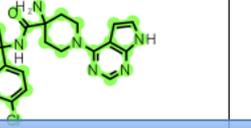
1627307787759_73  1.15 Med. Change [Goals](#)

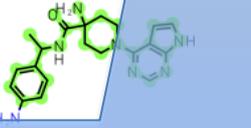
1627307787759_70  1.09 Med. Change [Goals](#)

1627307787759_31  1.00 Med. Change [Goals](#)

1627307787759_30  1.00 Med. Change [Goals](#)

CHEMBL2325732  0.82 Med. Change [Goals](#)

1627307787759_19  Med. Change

1627307787759_34  Med. Change

Save Selection is available once compounds have been selected

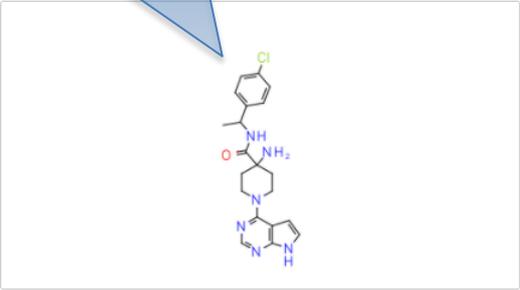
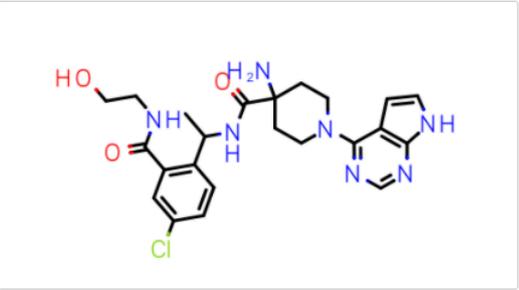
1 2 3

Input molecule and suggested molecule

Results Viewer

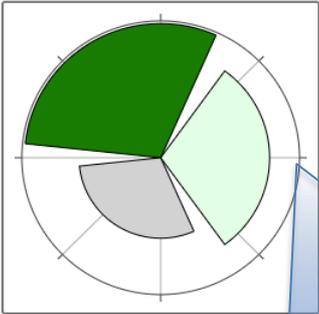
Reference: 1627307787759

Product: 1627307787759_74

Shown Goal: solubility

Aq_Solubility_pH_7.4 [CHEMBL2362975]
1.15 Med. Change

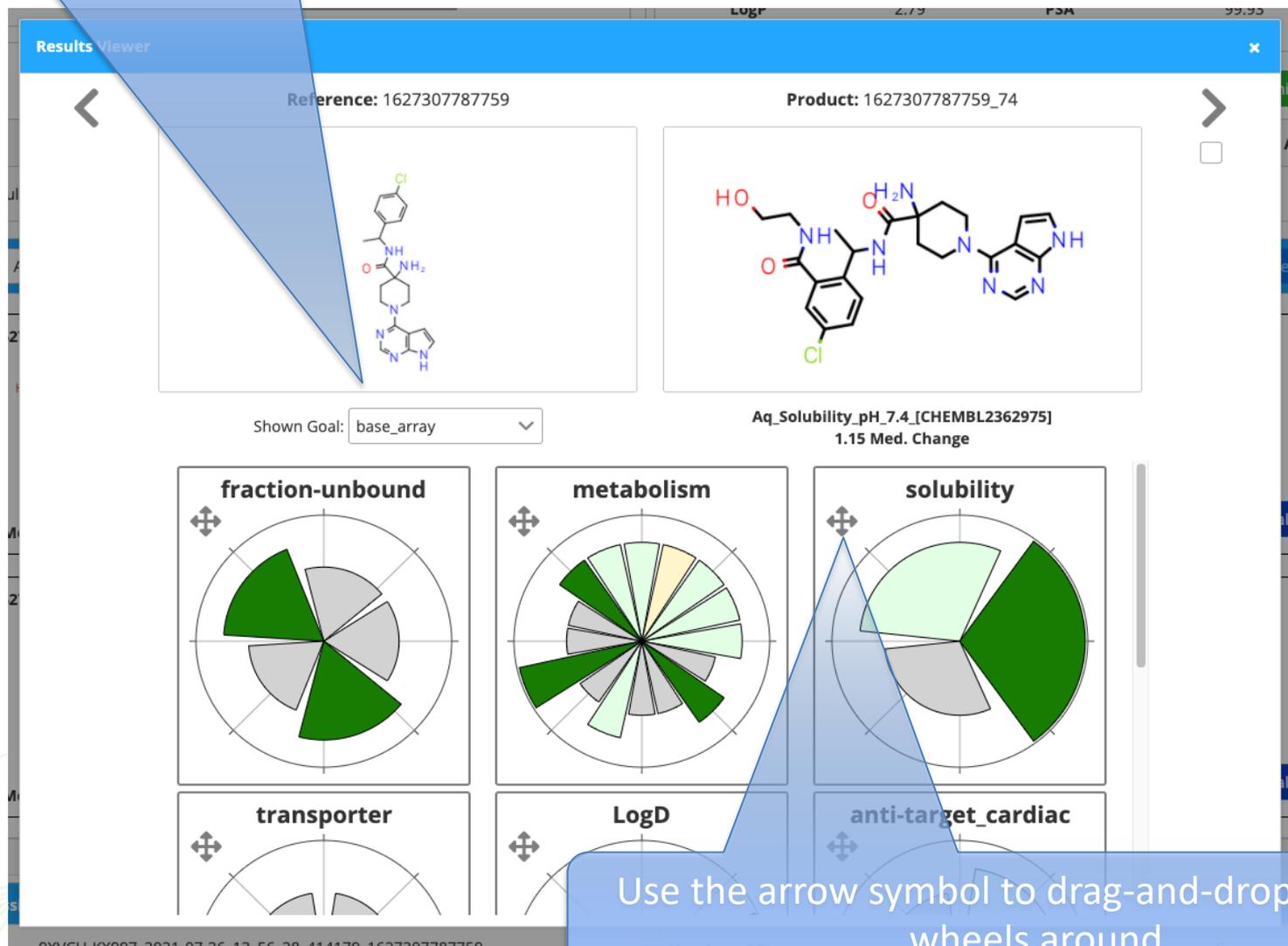


CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4ccfnH1c4ncn3)N

Select compound. Use Right and Left arrows to browse through.

Hover over to see more detailed matched pair data

Select base_array to get the basic ADMET properties



Export results

The screenshot shows a 'Save' dialog box with the following sections:

- Stats Fields***: 16 items selected
- Goal Endpoints***: Aq_Solubility_generic_log(M), Aq_Solubility_pH_7.4_[CHEMBL23...
- Base Array Endpoints***: 30 items selected
- Other Endpoints**: No Endpoints Selected
- Selected Compounds**: 4 compounds(s) selected
- File Type***: Spreadsheet (.xlsx) (selected), Spreadsheet (.xlsx), Flat File (.txt), CSV (.csv), Spotfire Text Data (.stdf)

Callouts point to the Stats Fields, Goal Endpoints, Base Array Endpoints, and File Type dropdown menus.

Select preferred stats fields

Customise which datasets are exported. Default is your Goal and the Base Array

Select output file style

Using the results

- Download the Excel file
- Inspect the structures
 - Filtering and sorting is possible as structures locked to cells
- Look at the breadth of data
- Unhide columns for more data
- Or export csv and use Spotfire, Vortex, Datawarrior....

Looking at the results

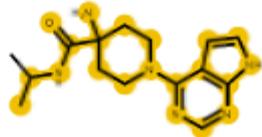
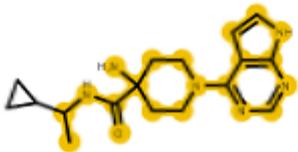
MedChemica

CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

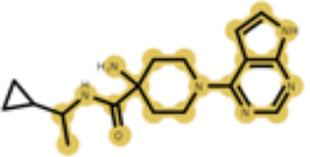
Results sorted in increasing RMM (Mol Weight)

One column per assay
 – colour and direction
 - LogD decrease, Sol increase

Yellow highlight is the overlap with the input compound

	G	H	I	J	K	L		AD	AT	
							LogD TM	Aq Solubility pH 7.4 [CHEMBL236297 5]	Aq Solubility pH 7.4 [CHEMBL612558]	
		pair data	CLogP	HBA	HBD	PSA	RMM	direction	direction	direction
		matched pair data	0.35	7	3	99.93	302.4	decrease	NED	.
		matched pair data	0.72	7	3	99.93	328.4	NED	.	.

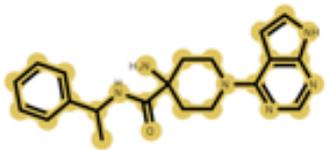
Looking at the Stats

G	H	I	J	K	L	M	BJ	BQ	BR	BS	BT	BU	BV	BW
Depiction	pair data	ClogP	HB A	HB D	PSA	RMM	Aq Solubility comb patent data log(M) direction	Aq Solubility comb patent data log(M) n quant	Aq Solubility comb patent data log(M) n qual	Aq Solubility comb patent data log(M) n qual up	Aq Solubility comb patent data log(M) n qual down	Aq Solubility comb patent data log(M) median change	Aq Solubility comb patent data log(M) std dev	Aq Solubility comb patent data log(M) std dev
	matched pair	0.4	7	3	100	302.4	increase							
<chem>C1[CCN(CC1)c2c3cc[nH]c3ncn2]N</chem>	matched pair	0.7	7	3	100	328.4	increase	9	9	9	0	1.701	0.586	

Click the hyperlink to see the source data

n-qual – number of matched pairs used in analysis
 n-qual-up – number of +ve delta measurements
 n-qual-down – number of -ve delta measurements
 n-median result – middle delta measurement
 Std-dev – standard deviation – give an idea of the spread of data

Looking at the Stats

G	H	I	J	K	L	M	QO	QK	QL	QM	QN	QO	QP	QQ	QR
Depiction	pair data	ClO gP	HB A	HB D	PSA	RMM	hERG hum inhib pIC50 direction	hERG hum inhib pIC50 n quant	hERG hum inhib pIC50 n qual	hERG hum inhib pIC50 n qual up	hERG hum inhib pIC50 n qual down	hERG hum inhib pIC50 median change	hERG hum inhib pIC50 std dev	hERG hum inhib pIC50 std err	hERG hum inhib pIC50 binom test
 <chem>NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)N</chem>	matched pair	1.6	7	3	100	364.4	decrease	133	133	19	114	-0.411	0.357	0	0

Looking at this example – Cl to H to decrease hERG binding

133 matched pairs

114 pairs are -ve so pIC50 is lower for the phenyl

19 pairs are +ve so pIC50 is higher for the phenyl

-0.411 – so half of the pairs are -0.411 or greater improvement

0.357 – std dev – small spread – very good chance of 3 fold improvement

Summary

- Choose a 'reasonable' start point
- Focus your results by iterative runs
- Use Substructure locking

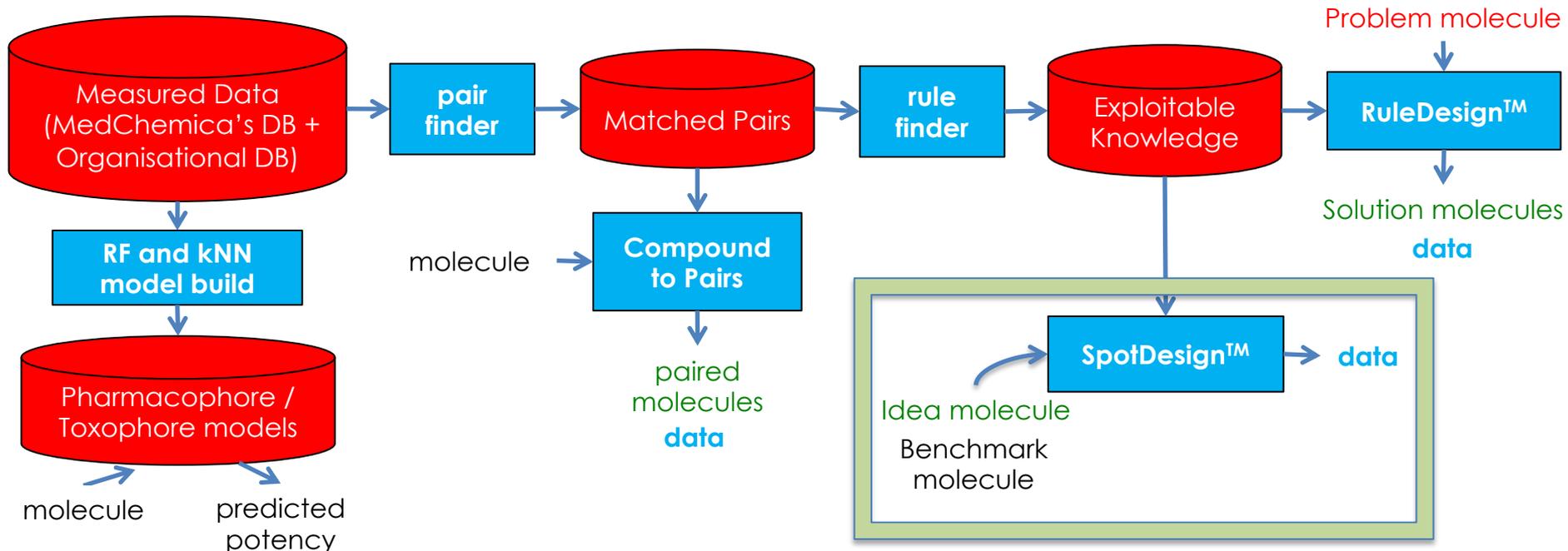
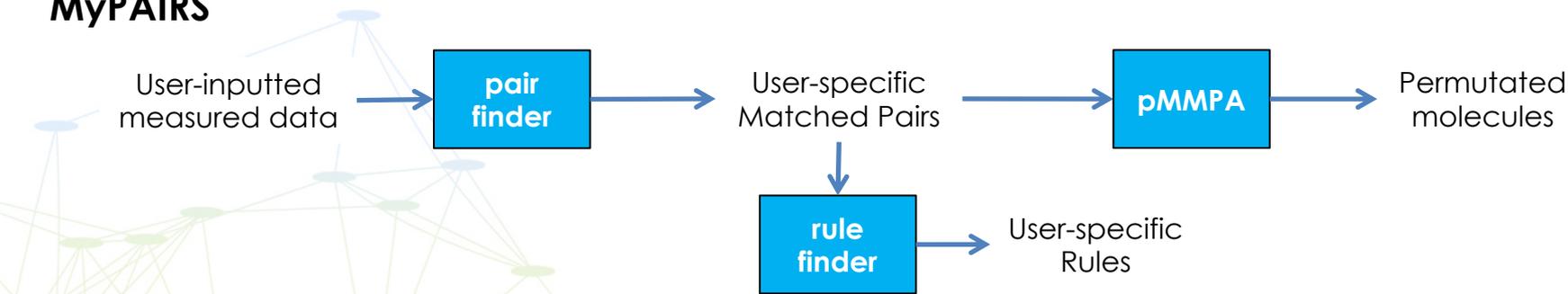
Remember:

Always download your results – server stores nothing!

Questions on RuleDesign?

Please use Webinar tools if you are on-line

Exploiting data derived ADMET Knowledge

DATABASE TOOLS

MyPAIRS


Exercise – Explore ideas to improve solubility

- Input chemical name for lead molecule CHEMBL2325997
 - Choosing the property to improve (setting the Goal)
 - Find Phys_Prop -> Solubility (Broad Goal)
 - Returning results, exploring results
 - what do the numbers mean?
 - Drilling back to the original source (looking at the matched pairs)
 - Running multiple ideas
- Search – Inspect – Export
 - Save with structures or as .csv for other applications
 - (Structures saved as SMILES)

Matched Pairs

Access to MCPairs Database Matched Molecular Pairs Service



RULEDESIGN™

Formerly Known as Compounds From Rules

Submit compound(s), enumerate products using Rules from the central database



SPOTDESIGN™

Submit a reference compound and explore your ideas seeking support from the central database



COMPOUND TO PAIRS

Search the central database of the matched pairs of a compound



COMPOUND TO MEASUREMENTS

Search the central database for the current aggregated data of a compound

SpotDesign : How good is my idea?

Entering a reference compound

Reference Molecule*

Compound Name:

CHEMBL2325997

SMILES*:

CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)N

Copy as SMILES
Copy as MOL
Copy as MOL V3000
Copy as InChI
Copy as InChI key
Search chemical structure (through InChIKey)
Copy as Scalar Vector Graphics
Paste MOL or SDF or SMILES

Next >>

Compound name?

SMILES structure

Other methods

SpotDesign : How good is my idea?

Selecting a Goal

Endpoint*:

Goal*

Endpoint*:

solu

Phys_Props (50316)

- ✓ Solubility (12138)
 - ✓ Aq_Solubility_generic_log(M) (10832)
 - ✓ Aq_Solubility_pH_7.4_[CHEMBL2362975] (1)
 - ✓ Aq_Solubility_pH_7.4_[CHEMBL612558] (20)

Apply

Text search

Select Goal

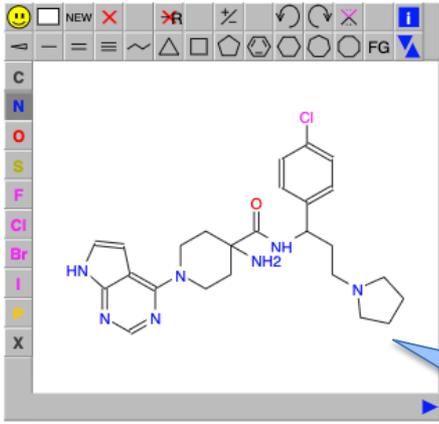
SpotDesign : How good is my idea?

Idea Molecule*

Single Multi

SMILES*:

```
NC5(C(=O)NC(CCN1CCCC1)c2ccc(Cl)cc2)CCN(c3ncnc4[nH]ccc34)CC5
```



Draw your idea molecule

	Idea	Delta	
ClogP	2.77	0.51	↑
HBA	8.00	1.00	↑
HBD	3.00	0.00	●
PSA	103.17	3.24	↑
RMM	482.02	83.13	↑

Submit

Rules

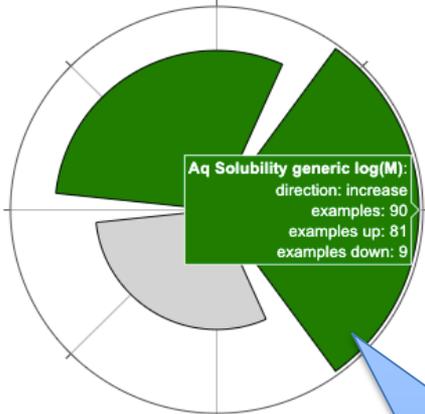
Phys_Props > Solubility

Endpoints with Rules:

2 / 3

Other Rules:

increase	2
decrease	2
remaining	37
total	41



Aq Solubility generic log(M):
 direction: increase
 examples: 90
 examples up: 81
 examples down: 9

Tip! Clicking an endpoint

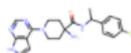
GO!

Hover to explore results
Click to drill deeper

Quick tip...

Reference Molecule* and Goal*

SMILES*:

CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)N


Endpoint*:

Solubility ✕

Select

Copy to Rule Design

Idea Molecule*

Single

Multi

SMILES*:

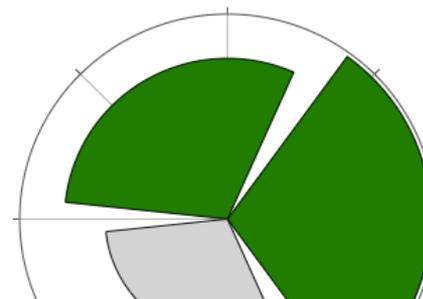
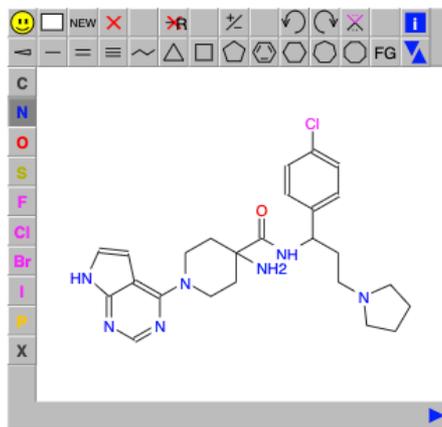
NC5(C(=O)NC(CCN1CCCC1)c2ccc(Cl)cc2)CCN(c3ncnc4[nH]ccc34)CC5

Shortcut to RuleDesign

2 / 3

Other Rules:

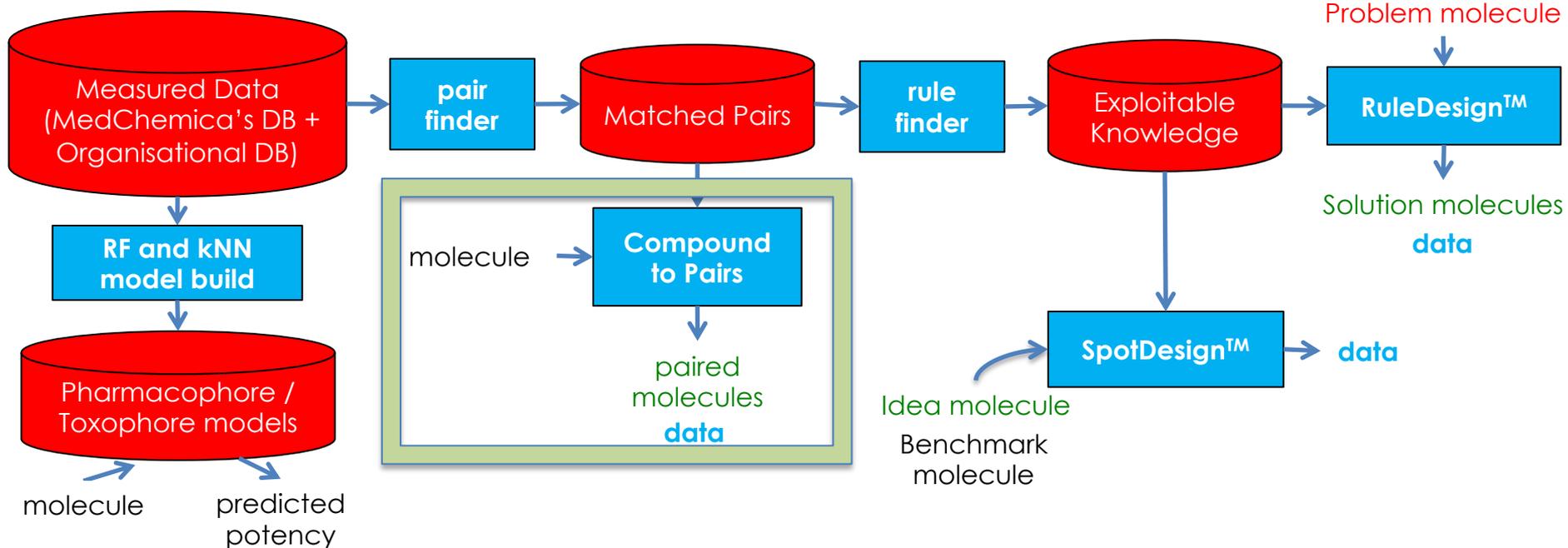
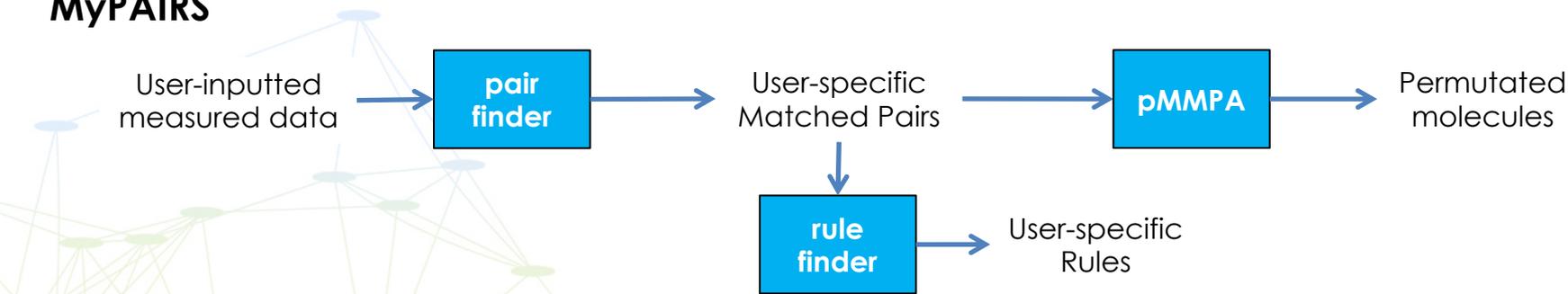
increase	2
decrease	2
remaining	37
total	41



Questions on SpotDesign™?

Please use Webinar tools if you are on-line

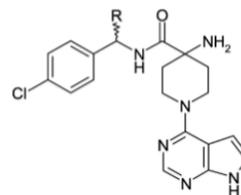
Exploiting data derived ADMET Knowledge

DATABASE TOOLS

MyPAIRS


Fast SAR understanding via Matched pairs

Table 2. Akt Enzyme and Cell Potency, Selectivity, and hERG Activity for α -Alkyl-Substituted Benzylamide Analogues

Table 3. Akt Enzyme and Cell Potency, Selectivity, and hERG Activity for Selected Aryl-Substituted Benzylamide Analogues

Table 4. Akt Enzyme and Cell Potency, Selectivity, and hERG Activity for α -Substituted Benzylamide Analogues Carrying a Basic Side Chain

 3
 32
 33
 34
 35
 36
 37
 38

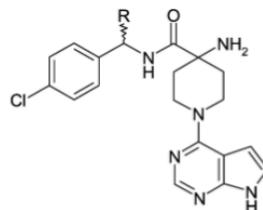
^aAll IC₅₀
 of phospho
 IonWorks

	R
39	
40	
41	
42	
43	
44	

	R	IC ₅₀ (nM)						log D	solubility ^e (μ M)
		Akt1 ^a	Akt2 ^a	Akt3 ^a	Cell ^b	ROCK2 ^c	hERG ^d		
45		2	14	7	96	31 [16]	>100000		
46		3	15	4	126	71 [25]	>100000	1.4	> 1700
47		4	96	36	209	104 [29]	>100000	1.3	> 1500
48		5	42	15	156	134 [25]	74807	2.7	1400
49		5	30	10	81	112 [24]	>33300		
50	S-	3			110	34 [12]	>33300		2320
51	S-	4	31	18	43	39 [10]	21030		>2470
52	S-	4	13	6	78	34 [10]	29367	2.4	>2090

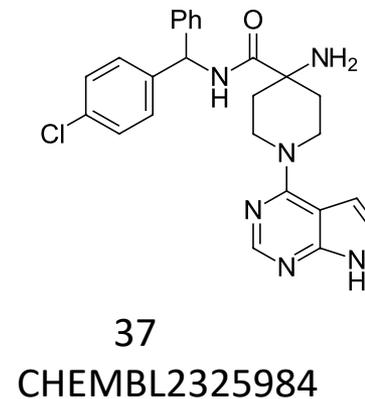
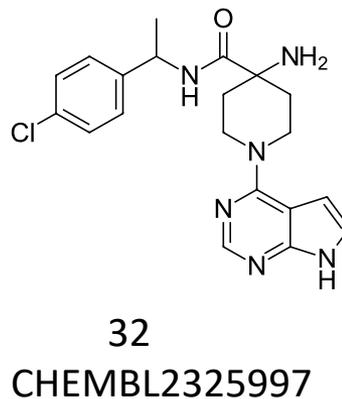
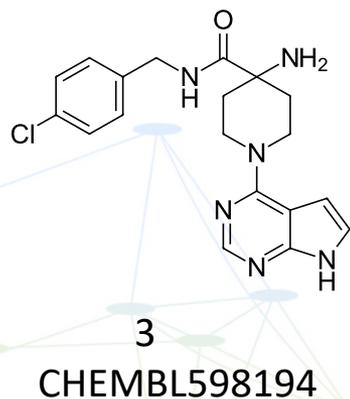
^aAll IC₅₀ data are the mean of at least 3 experiments
 mediated by Akt in MDAMB468 cells
 solubility in 0.1 M phosphate buffer

SAR worked example: AKT inhibitors

Table 2. Akt Enzyme and Cell Potency, Selectivity, and hERG Activity for α -Alkyl-Substituted Benzylamide Analogues


	R	IC ₅₀ (nM)					log <i>D</i>	solubility ^e (μ M)	
		Akt1 ^a	Akt2 ^a	Akt3 ^a	cell ^b	ROCK2 ^c			hERG ^d
3	H	13	66	57	328	66 [5]	5235	2.9	150
32	Me	8	40	30	197	101 [13]	7200	2.7	180
33	<i>R</i> -Me	276	836	523	4594	1396 [5]	9092	3.3	77
34	<i>S</i> -Me	4	20	16	134	55 [15]	6747	2.7	110
35	Et	7	23	15	144	126 [19]	6495	3.5	13
36	<i>c</i> -Pr	5	30	24	208	261 [52]	2600	3.4	31
37	Ph	41	210	270	1620	576 [14]	1600	4.1	<1
38	Bn	31	190	150	1650	586 [19]	3500	4.1	<1

^aAll IC₅₀ data are the mean of at least $n = 2$ independent measurements. Each has a standard error of measurement (SEM) ± 0.2 log unit. ^bInhibition of phosphorylation of GSK3 β mediated by Akt in MDAMB468 cells. ^cValue in brackets indicates enzyme selectivity ratio to Akt1. ^dCHO cells, IonWorks assay. ^eThermodynamic solubility in 0.1 M phosphate buffer at pH 7.4 (25 °C).



Exercise – Rapid exploration of SAR for literature compounds

- Input chemical name to return matched pairs (CHEMBL598194)
 - Add assay data and explore the SAR tables
 - Add structures, Add ALogp98
 - Change units and sort the data
 - Understand the “gaps” in data
- Search – Inspect – Export
 - Save with structures or as .csv for other applications
 - (Structures saved as SMILES)
- Extras:
 - Finding compounds in ChEMBL and patent literature

Matched Pairs

Access to MCPairs Database Matched Molecular Pairs Service



RULEDESIGN™

Formerly Known as Compounds From Rules

Submit compound(s), enumerate products using Rules from the central database



SPOTDESIGN™

Submit a reference compound and explore your ideas seeking support from the central database



COMPOUND TO PAIRS

Search the central database of the matched pairs of a compound



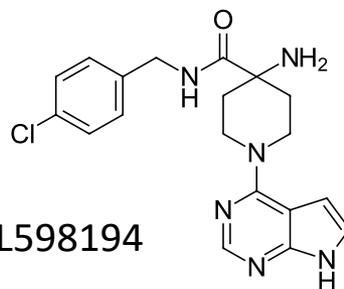
COMPOUND TO MEASUREMENTS

Search the central database for the current aggregated data of a compound



SAR worked example: AKT inhibitors

3
CHEMBL598194



MedChemica
CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

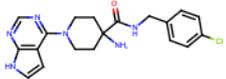
Compound to Pairs

Modules My Account

Input

Compound

Select Endpoint



Pairs Found:
77

Pairs with measurements:
50

Filter by Core

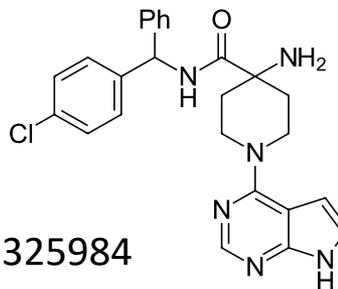
No Cores

Show ALogP98 Show Structures

Click to show structures in results

	A ?	B ↑↓	
▶	CHEMBL598194	CHEMBL2325988	🔍

SAR worked example: AKT inhibitors



37
CHEMBL2325984

MedChemica
CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

MedChemica

[Modules](#)
[My Account](#)

CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

Compound to Pairs

Input

Compound

Select Endpoint

Pairs Found:

55

Pairs with measurements:

55

Filter by Core

No Cores

Show
ALogP98

Show
Structures

	A ?	B	
▶	CHEMBL2325984	CHEMBL2325728	🔍

SAR worked example

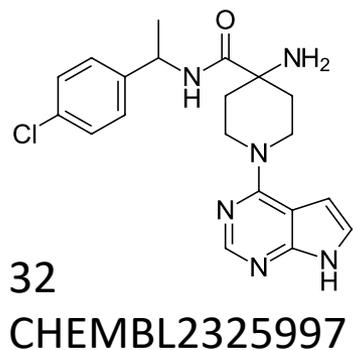


Table 2. Akt Enzyme and Cell Potency, Selectivity, and hERG Activity for α -Alkyl-Substituted Benzylamide Analogues

	R	IC ₅₀ (nM)					ROCK2 ^c	hERG ^d	log D	solubility ^e (μ M)
		Akt1 ^a	Akt2 ^a	Akt3 ^a	cell ^b					
3	H	13	66	57	328	66 [5]	5235	2.9	150	
32	Me	8	40	30	197	101 [13]	7200	2.7	180	

MedChemica
CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

Compound to Pairs

Modules My Account

Input

Compound: CHEMBL2325997

Select Endpoint: Choose Endpoint(s)

- Serine/threonine-protein_kinase_AKT3_Homo_sapien
- Serine/threonine-protein_kinase_AKT2_Homo_sapien
- Rho-associated_protein_kinase_2_Homo_sapiens_pIC
- Serine/threonine-protein_kinase_AKT_Homo_sapiens
- HERG_Homo_sapiens_pIC50_[CHEMBL240]

Pairs Found: 68

Pairs with measurements: 68

Filter by Core

No Cores

77
110
13
31
<1
<1
mit. ^bInhibition
1. ^dCHO cells,

Click to see data

AKT inhibitors

Input

Compound

Pairs Found:

68

Filter by Core

Pairs with measurements:

32

Click to show all structures

Or click to show just one pair

Show ALogP98 Show Structures

	A ?	B ↓		Serine/threonine-protein kinase_AKT...			HERG_Homo_sapiens_pIC50_[CHEMBL...		
				B ↓	Fold Change ↓		B ↓	Fold Change ↓	
				uM			uM		
▶	CHEMBL2325997	CHEMBL2325988	<input type="button" value="Q"/>	4.260	~ 140	↓	33.297	~ 4.6	↓
▶	CHEMBL2325997	CHEMBL2325728	<input type="button" value="Q"/>	0.018	~ 1.7	↑	21.028	~ 2.9	↓

Instant SAR table

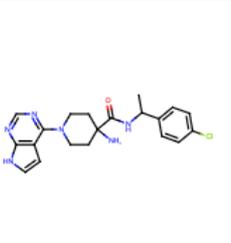
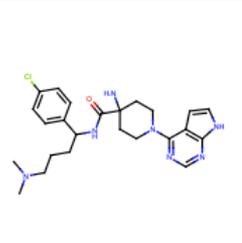
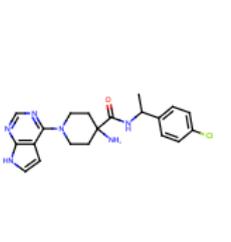
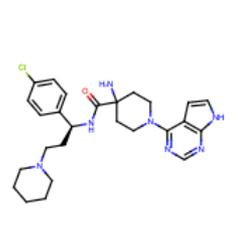
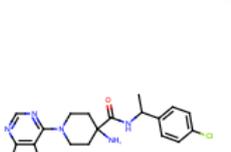
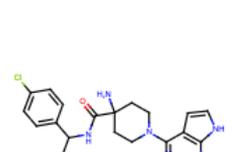
Click to add
lipophilicity
data

MedChemica
CREATING A STEP CHANGE IN MEDICINAL CHEMISTRY

Click to
change units

Show
ALogP98

Hide
Structures

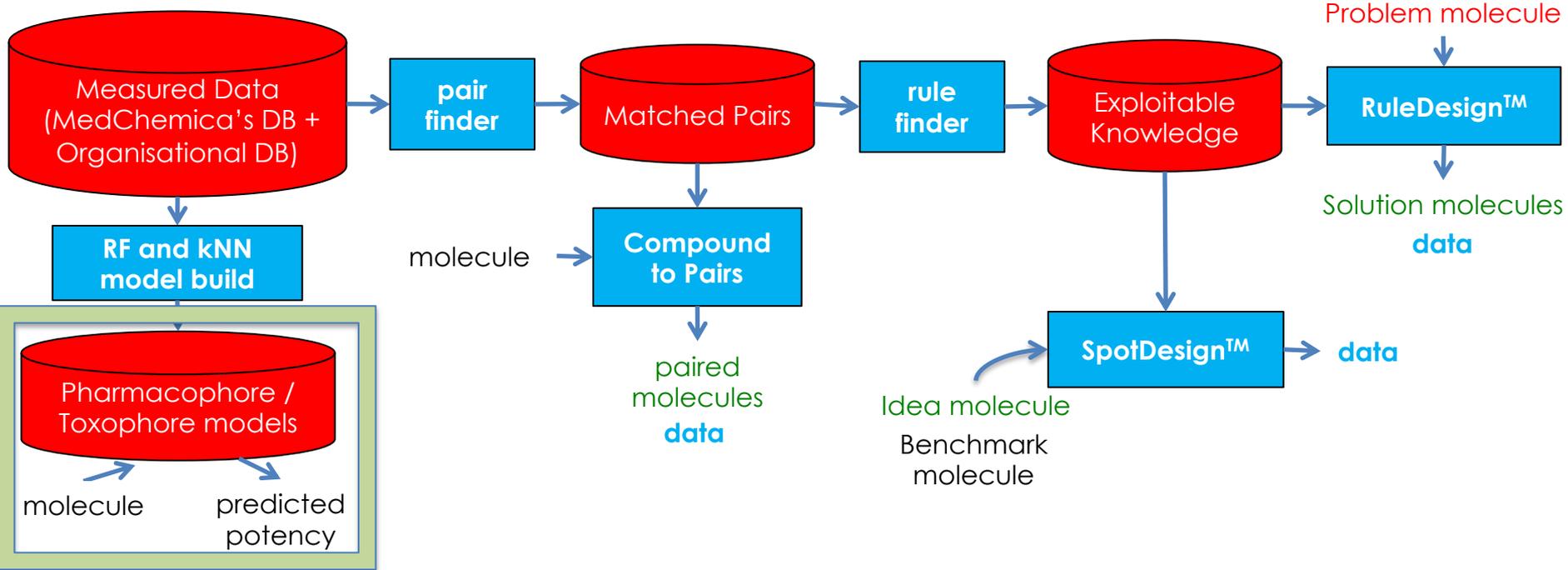
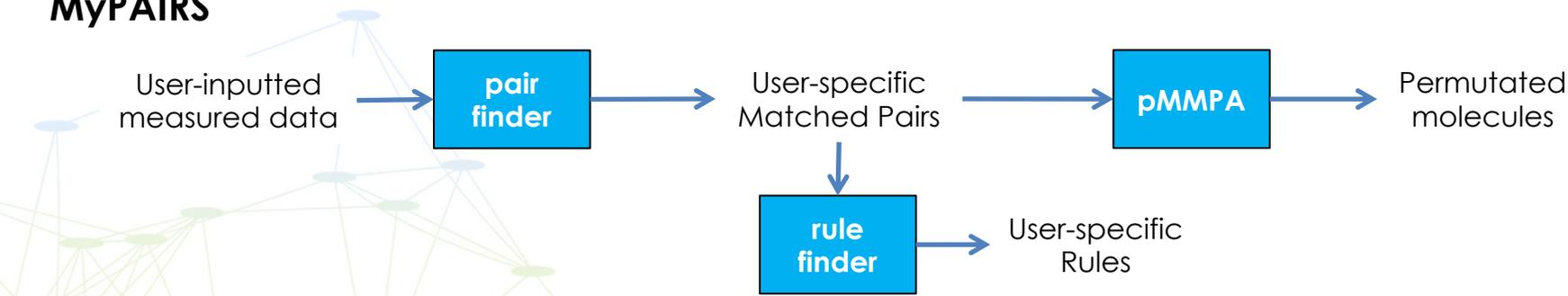
		Serine/threonine-protein_kinase_... uM			HERG_Homo_sapiens_pIC50_[CHEMBL...] uM	
		B ↑↓	Fold Change ↑↓		B ↑↓	Fold Change ↑↓
		0.004	~ 7.5 ↑		100.000	~ 14
		0.006	~ 5 ↑		29.370	~ 4.1
		0.007	~ 4.3 ↑		100.000	~ 14

Click to sort
pair with large
change to the
top

Questions on Compound-To-Pairs?

Please use Webinar tools if you are on-line

Exploiting data derived ADMET Knowledge

DATABASE TOOLS

MyPAIRS


Toxophore and Pharmacophore prediction models

- Models that utilise knowledge from the central database to predict potency
 - Potency at a toxicologically-linked Endpoint = Toxophore
 - Potency at a target Endpoint = Pharmacophore
- 2 types of models are used:
 - Regression Forest (RF)
 - Decision trees trained on central database data
 - Descriptors = pairs of features connected by shortest bond path
 - k Nearest Neighbour (kNN)
 - Compares query molecule with central database molecules
 - Descriptor = Morgan fingerprint
 - Distance metric = Tanimoto distance

Exercise – Assess potential toxophore activity of CHEMBL2325997

- Input compound name: CHEMBL2325997
- Select Toxicologically-linked Endpoints
- Run Toxophore prediction
- View results
 - Take a minute to understand the predictions using the help information provided (click ?)
- Export as ready-to-use PowerPoint presentation

Predict

Use pre-built database Machine Learning Models to make predictions



TOXOPHORES

Predict potential issues, using models from database data, for a compound



PHARMACOPHORES

Predict properties (potency), using models built from database data, for a compound

Chemistry Toolbox

A selection of chemistry tools



MyPAIRS

Take your data, find matched pairs, rules and permute new compounds



CLEAN, CALC and FILTER

Take your data, clean up the SMILES, calculate properties, and uses SMARTS to search or filter



DEPICT

A straight forward SMILES / SMARTS checker

Toxicity predictions for CHEMBL2325997

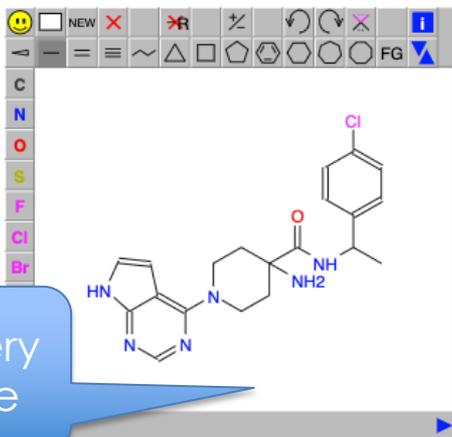
Input Molecule*

Single Multi

Compound Name:

CHEMBL2325997

SMILES*:

CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)NInput query
molecule

Assay Selection

Toxophore Assays

	Assay	Labels
<input checked="" type="checkbox"/>	Glutamate NMDA receptor; GRIN1/GRIN2B Homo sapiens pIC50 [CHEMBL1907603]	CNS
<input checked="" type="checkbox"/>	Voltage-gated L-type calcium channel alpha-1C subunit Oryctolagus cuniculus pIC50 [CHEMBL2830]	CVS
<input checked="" type="checkbox"/>	Neuronal acetylcholine receptor; alpha4/beta2 Homo sapiens pIC50 [CHEMBL1907589]	CNS
<input checked="" type="checkbox"/>	Serotonin 1b (5-HT1b) receptor Homo sapiens pIC50 [CHEMBL1898]	CNS
<input checked="" type="checkbox"/>	Phosphodiesterase 3A Homo sapiens pIC50 [CHEMBL241]	CVS
<input checked="" type="checkbox"/>	Serotonin 1a (5-HT1a) receptor Homo sapiens pIC50 [CHEMBL214]	CNS
<input checked="" type="checkbox"/>	Dopamine D1 receptor Homo sapiens pIC50 [CHEMBL2056]	CNS
<input checked="" type="checkbox"/>	Muscarinic acetylcholine receptor M2 Homo sapiens pIC50 [CHEMBL211]	CNS
<input type="checkbox"/>	Sodium channel protein type V alpha	

Select toxicity
Endpoint(s)

Go!

Submit

Toxicity predictions for CHEMBL2325997

Red = predicted potency at toxic Endpoint

Scroll through results

Green = predicted inactivity at toxic Endpoint

the highest ordered biosystem label assigned.

CNS

Acetylcholinesterase *Electrophorus electricus* pIC50 [CHEMBL4078]

Cannabinoid CB1 receptor *Homo sapiens* pIC50 [CHEMBL2325997]

Delta opioid receptor *Homo sapiens* pIC50 [CHEMBL236]

Dopamine D1 receptor *Homo sapiens* pIC50 [CHEMBL2056]

Dopamine D2 receptor *Homo sapiens* pIC50 [CHEMBL217]

Dopamine transporter *Homo sapiens* pIC50 [CHEMBL238]

GABA-A receptor; anion channel *Rattus norvegicus* pIC50 [CHEMBL1907607]

Glutamate NMDA receptor; GRIN1/GRIN2B *Homo sapiens* pIC50 [CHEMBL1907603]

Kappa opioid receptor *Homo sapiens* pIC50 [CHEMBL237]

Monoamine oxidase A *Homo sapiens* pIC50 [CHEMBL1951]

Mu opioid receptor *Homo sapiens* pIC50 [CHEMBL233]

Muscarinic acetylcholine receptor M1 *Homo sapiens* pIC50 [CHEMBL216]

Muscarinic acetylcholine receptor M2 *Homo sapiens* pIC50 [CHEMBL211]

Neuronal acetylcholine receptor; alpha4/beta2 *Homo sapiens* pIC50 [CHEMBL1907589]

Norepinephrine transporter *Homo sapiens* pIC50 [CHEMBL2325997]

Phosphodiesterase 4 *Homo sapiens* pIC50 [CHEMBL2093863]

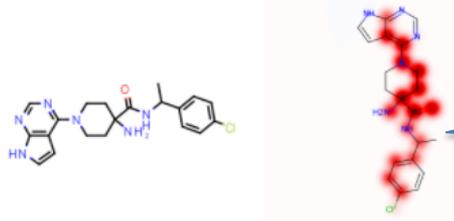
Click on Endpoint for breakdown

Tip: click ? for more info

Tip: click to design less toxic compounds

Kappa opioid receptor Homo sapiens pIC50 [CHEMBL237] CC(c1ccc(cc1)Cl)NC(=O)C2(CCN(CC2)c3c4cc[nH]c4ncn3)N

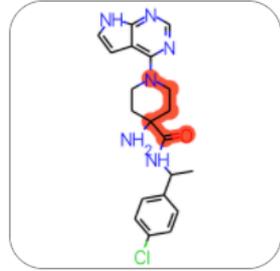
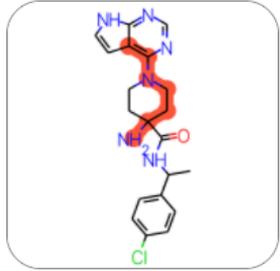
Unit
To Rule Design



Overall heat map of atom contributions

RF prediction (based on Cohen's d and R²)

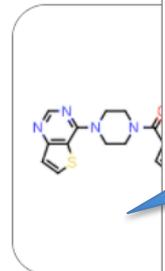
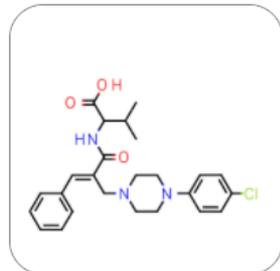
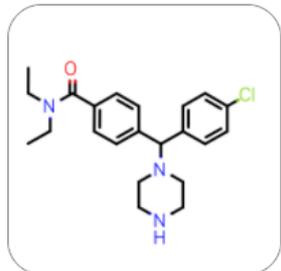
2 Pharmacophore(s)
Active
In domain
pIC50: 6.229
Range: 5.346 - 7.112



Most important descriptors from RF

kNN prediction (based on activity of NN and Tanimoto distance to NN)

Nearest Neighbour(s)
Active
Out of domain
pIC50: 6.114
Range: 5.945 - 6.283



Nearest neighbours in database

CHEMBL128006
pIC50: 5.000
Tan. distance: 0.730

CHEMBL4067441
pIC50: 6.032
Tan. distance: 0.732

CHEMBL3
pIC50:
Tan. distance:

Predict

Use pre-built database Machine Learning Models to make predictions



TOXOPHORES

Predict potential issues, using models from database data, for a compound



PHARMACOPHORES

Predict properties (potency), using models built from database data, for a compound

Chemistry Toolbox

A selection of chemistry tools



MyPAIRS

Take your data, find matched pairs, rules and permute new compounds



CLEAN, CALC and FILTER

Take your data, clean up the SMILES, calculate properties, and uses SMARTS to search or filter



DEPICT

A straight forward SMILES / SMARTS checker

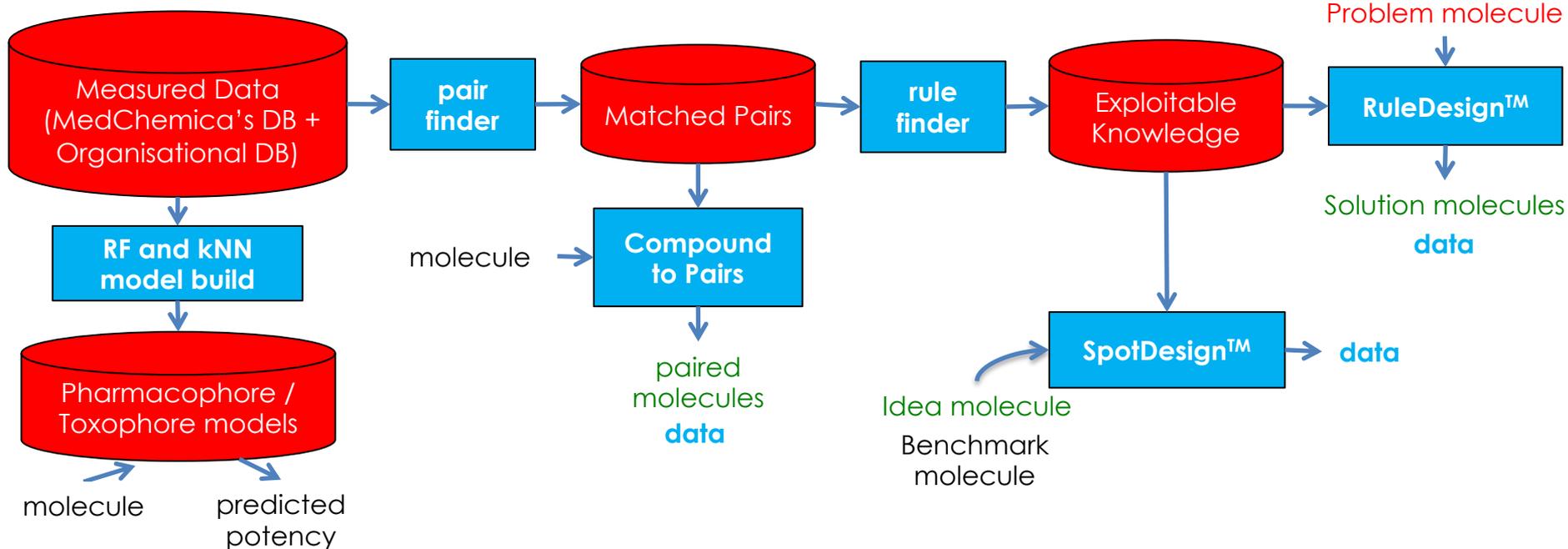
Same methods as Toxophores but
green = active
red = inactive

Questions on Toxophores / Pharmacophores?

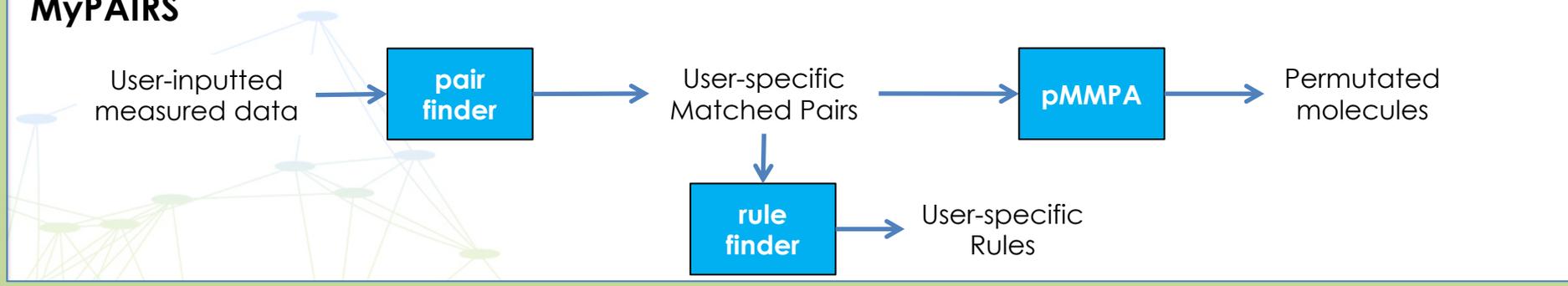
Please use Webinar tools if you are on-line

Exploiting data derived ADMET Knowledge

DATABASE TOOLS



MyPAIRS



These tools do not interact with the central database of knowledge

Chemistry Toolbox

A selection of chemistry tools



MyPAIRS

Take your data, find matched pairs, rules and permute new compounds



CLEAN, CALC and FILTER

Take your data, clean up the SMILES, calculate properties, and uses SMARTS to search or filter



DEPICT

A straight forward SMILES / SMARTS checker

MyPairs

- Applies MCPairs methods to user-inputted data
 - Does not interact with central database
- Available methods:
 - Pair Finding
 - Rule Finding
 - Permutative MMPA (molecule suggestions based on application of transformations)
- Example use on a live project:
 - COVID moonshot
 - Open-source antiviral discovery project
 - Targeting SARS-COV-2 main protease
 - Input data = protease binding data dated 14-12-2020
 - Does pMMPA used on historical project data suggest our current lead compound (3-02-2021)?

1 Matched Molecular Pairs Finder - MyPairs

Select a file:

Input data file

Current Selection: **Fl_agg_data_14_12_2020_23_26_47.txt**
 Available file types: Compound Data file or SMILES file (Direct Pairs only)

Pair Finding Method ⓘ

Choose pair finding method

Environment Size Cut-Off ⓘ

- Ignore Chirality
 Show Hydrogens
 Disengage RDKit Calculations

+

+

2 Calculate Rules

 Add Rules Calculated from Matched Pairs

3 Permutative Matched Molecular Pair Analysis (pMMPA)

 Apply pMMPA

Click here to edit pair finding settings

#

Job ID

Timestamp

Progress

Save

No Jobs Available

Hover over to read information on each parameter

FI Settings

Max Batch Size ⓘ

Fragmentation SMARTS Pattern ⓘ

Fragments are generated by disconnecting single rotatable bonds (see SMARTS pattern). We require one fragment to be smaller than the other. The smaller fragment is a 'modification' and the larger a 'common core'. Define the maximum size of the modification fragment (8 to 16 Ha is typical).

Max Heavy Atoms in Fragment ⓘ

Ratio, by heavy atoms, for Fragmentation ⓘ

Apply fragment size cutoff using ratio to size of whole molecule

Apply largest frag algorithm - selects the best fragmentation pattern (called Filargest on the command line)

MCSS Settings

MCSS Heavy Atom Overlap ⓘ

Rapid method?

Click here to enable rule finding

1 Matched Molecular Pairs Finder - MyPairs

Select a file:

Current Selection: **Fl_agg_data_14_12_2020_23_26_47.txt**

Available file types: [Compound Data file](#) or [SMILES file \(Direct Pairs only\)](#)

Pair Finding Method ⓘ

Environment Size Cut-Off ⓘ

- Ignore Chirality
 Show Hydrogens
 Disengage RDKit Calculations

+

+

2 Calculate Rules

 Add Rules Calculated from Matched Pairs

3 Permutative Matched Molecular Pair Analysis (pMMPA)

 Apply pMMPA

Click here to enable permutation of molecules

#

Job ID

Timestamp

Progress

Save

No Jobs Available

1 Matched Molecular Pairs Finder - MyPairs

Select a file:

Current Selection: **FI_agg_data_14_12_2020_23_26_47.txt**

Available file types: [Compound Data file](#) or [SMILES file \(Direct Pairs only\)](#)

Pair Finding Method ⁱ

Environment Size Cut-Off ⁱ

Ignore Chirality
 Show Hydrogens
 Disengage RDKit Calculations

FI Settings

MCSS Settings

2 Calculate Rules

Add Rules Calculated from Matched Pairs
 Minimum Pairs required for a Rule

3 Permutative Matched Molecular Pair Analysis (pMMPA)

Apply pMMPA

Number of Generations Input measurement data is on a logarithmic scale e.g. pIC50. A cutoff can be applied to only yield new molecules with a likely prediction above this value.

Prediction Threshold ⁱ

Available file types: [Compound Data file](#)

Available file types: [MCPairs Rules file](#) or [pMMPA Rules input file](#)

Available file types: [Compound Data file](#) or [SMILES file](#)

Physical Property Filters
Filter out compounds outside of the selected bounds

Lipophilicity
 PSA
 HBA
 HBD
 Molecular Weight
 Charge

Hover over for description of each parameter

Default is to apply 1 generation of permutations

Option to add seed molecules (subset of input data you wish transformations to be applied to)

Option to add additional transformations for permutation

Option to filter results by physical property

Show Hydrogens
 Disengage RDKit Calculations

FI Settings +

MCSS Settings +

minimise

Available file types: Compound Data file

Available file types: MCPairs Rules file or pMMPA Rules input file

Available file types: Compound Data file or SMILES file

Physical Property Filters
Filter out compounds outside of the selected bounds

Lipophilicity
 PSA
 HBA
 HBD
 Molecular Weight
 Charge

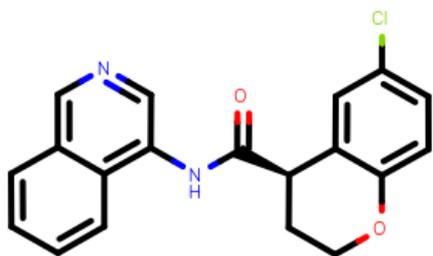
#	Job ID	Timestamp	Progress	Save
0	BQZIA-U0FBK_2021-02-03_10_39_54_609375	Feb 3, 2021, 10:55:42 AM	Finished	<input type="button" value="Save"/>

Export zip file
of results

MyPairs – COVID results

- Use excel to apply project-relevant filters e.g:
 - Sort suggestions by predicted median binding measurement
 - Filter for CLogP < 4
 - And we find our current lead compound as a suggestion:

	A	B	C	I	J	K	L	M	N	O	P	Q	R	T
1	id	structure	generation	precursor	precursor_mcs	pair_dat	CLogP	PSA	RMM	HB _A	HB _B	measurement	qualifier	prediction_median
1405	prod_gen_1_3930	Cc1cncc1NC(=O)CN2CCN(CC2)C(=O)CCl	1	Cc1cncc1NC(=O)CN2CCN(CC2)C(=O)C	Cc1cncc1NC(=O)CN2CCN(CC2)		-0.43	65.54	310.78	6	1		~	7.5
1409	prod_gen_1_284	[H][C@H]1([C@H](c2c(ccc(c2C)Cl)OC1)C(=O)Nc3c4c(cnc3)cccc4)[H	1	C[C@H]1COc2ccc(cc2[C@H]1C(=O)Nc3cncc4c3cccc4)Cl	[H][C@H]1COc2ccc(cc2C@H		3.55	51.22	352.81	4	1		~	7.3
1432	prod_gen_1_3667	Cc1cccc1NC(=O)CC2C(=O)N(C(=O)S2)c3cccc(c3)C(F)(F)F	1	Cc1cccc1NC(=O)CC2C(=O)N(C(=O)S2)c3cccc3	Cc1cccc1NC(=O)CC2C(=O)N(C		3.89	66.48	408.39	5	1		~	7.2
1435	prod_gen_1_1755	c1ccc2c(c1)cncc2NC(=O)[C@H]3CCNc4c3cc(cc4)Cl	1	c1ccc2c(c1)cncc2NC(=O)[C@H]3CCOc4c3cc(cc4)Cl	CC[C@H](c1cccc(c1)Cl)C(=O)Nc		3.12	54.02	337.8	4	2		~	7.1
1436	prod_gen_1_1791	c1ccc2c(c1)cncc2NC(=O)C3CCNc4c3cc(cc4)Br	1	c1ccc2c(c1)cncc2NC(=O)C3CCOc4c3cc(cc4)Br	CCC(c1cccc(c1)Br)C(=O)Nc2cnc		3.31	54.02	382.25	4	2		~	7.1
1438	prod_gen_1_2384	c1cc(c2c(c1)c(c[nH]2)CCNC(=O)CC)CN3CCN(CC3)C(=O)CCl	1	CC(=O)NCCc1c[nH]c2c1cccc2CN3CCN(CC3)C(=O)CCl	CC(=O)NCCc1c[nH]c2c1cccc2C		1.34	68.44	411.33	6	2		~	7.1
1439	prod_gen_1_2717	c1ccc2c(c1)cncc2NC(=O)[C@H]3CCS(=O)(=O)c4c3cccc4	1	c1ccc2c(c1)cncc2NC(=O)[C@H]3CCOc4c3cc(cc4)Cl	CC[C@H](c1cccc1)C(=O)Nc2cr		1.89	76.13	352.41	5	1		~	7.1

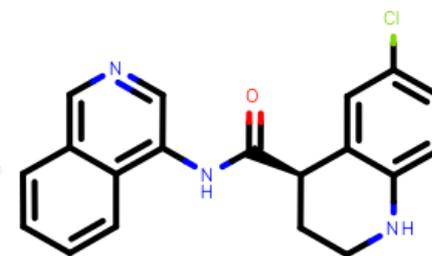


Lead Compound 14-12-2020

pIC₅₀ = 6.7

Moderate antiviral cell activity

High microsomal clearance



Lead Compound 3-2-2021

pIC₅₀ = 6.6

Good antiviral cell activity

Moderate microsomal clearance

– Project benefit:

- Alternative lead compound found with better antiviral and ADMET properties

MyPairs – COVID results

- pmmpa_evidence_gen_1.csv
 - Provides matched pair evidence for suggestions

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	product_id	product_structure	precursor_id	precursor_structure	idsmirks	smirks	environment	compound_name_A	compound_structure_A	measurement_A	qualifier_A	compound_name_B	compound_structure_B	measurement_B	qualifier_B
2	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	MAT-POS-8a69d52e-3 C[C@H]1C[C@H])(c2cc(4.3 =		CHEMBL94657	C[C@]1(CCOC2c1cc(cc		4.7 =
3	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	MAT-POS-8a69d52e-3 C[C@H]1C[C@H])(c2cc(4.3 =		CHEMBL2182047	C[C@]1(CCOC2c1cc(cc2		6.5 =
4	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	MAT-POS-8a69d52e-2 C[C@H]1C[C@H])(c2c(6.2 =		CHEMBL94657	C[C@]1(CCOC2c1cc(cc		4.7 =
5	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	MAT-POS-8a69d52e-2 C[C@H]1C[C@H])(c2c(6.2 =		CHEMBL2182047	C[C@]1(CCOC2c1cc(cc2		6.5 =
6	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	EDJ-MED-e4b030d8-3 C[C@H]1C[C@H])(c2cc(cc		4.7 =		CHEMBL94657	C[C@]1(CCOC2c1cc(cc		4.7 =
7	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	EDJ-MED-e4b030d8-3 C[C@H]1C[C@H])(c2cc(cc		4.7 =		CHEMBL2182047	C[C@]1(CCOC2c1cc(cc2		6.5 =
8	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	EDJ-MED-e4b030d8-2 C[C@H]1C[C@H])(c2cc(6.6 =		CHEMBL94657	C[C@]1(CCOC2c1cc(cc		4.7 =
9	prod_gen_1_236	[H][C@]1(Oc2c(cc(c EDJ-MED-e4bC C[C@H]1C[C@H])(c2c MAT-POS [H][C:1]1((4	EDJ-MED-e4b030d8-2 C[C@H]1C[C@H])(c2cc(6.6 =		CHEMBL2182047	C[C@]1(CCOC2c1cc(cc2		6.5 =
10	prod_gen_1_237	c1cc2c(cc1F)cncc2N MAT-POS-b3e c1ccc2c(c1)cncc2NC(MAT-POS [c:1][c:2]1					4	MAT-POS-0c8fa4a7-3 c1ccc2c(c1)c(nn2C(=O)C		4 <		CHEMBL259181	c1cc(cc(c1)C)C(=O)n2		4 <
11	prod_gen_1_237	c1cc2c(cc1F)cncc2N MAT-POS-b3e c1ccc2c(c1)cncc2NC(MAT-POS [c:1][c:2]1					4	EDJ-MED-c8e7a002-1: c1ccc2c(c1)c(n[nH]2)NC(4 <		CHEMBL1667924	c1cc(cc(c1)C)C(=O)Nc		4 <
12	prod_gen_1_237	c1cc2c(cc1F)cncc2N MAT-POS-b3e c1ccc2c(c1)cncc2NC(MAT-POS [c:1][c:2]1					4	VLA-UCB-1dbca3b4-1: c1ccc2c(c1)cncc2NC(=O)N		6.4 =		CHEMBL3407863	c1cc2cncc(c2cc1F)NC(=		6.1 =
13	prod_gen_1_237	c1cc2c(cc1F)cncc2N MAT-POS-b3e c1ccc2c(c1)cncc2NC(MAT-POS [c:1][c:2]1					4	BEN-DND-7e92b6ca-7 c1ccc(cc1)S(=O)(=O)N2C		4 <		CHEMBL48412	c1ccc(cc1)S(=O)(=O)N2		4 <
14	prod_gen_1_238	c1cc2c(c1)cncc2NC ALP-POS-477d c1ccc2c(c1)cncc2NC(JAG-UCB- [C:1]([H])c					4	JAG-UCB-52b62a6f-11 Cc1ccn2cnnc2c1(C(=O)C		4 <		CHEMBL1667898	Cc1ccn2cnnc2c1NC(=O)N		4 <
15	prod_gen_1_239	[H][C@]1(Oc2c(cc(c MAT-POS-8a6 C[C@H]1COC2ccc(MAT-POS [H][C:1]1((4	MAT-POS-8a69d52e-5 C[C@H]1COC2ccc(4 <		CHEMBL133016	C[C@H]1C[C@H])(c2cc		6.6 =

Matched pair structures and data listed for each suggestion

Questions on MyPairs?

Please use Webinar tools if you are on-line

And Finally...

- Problems, suggestions, ideas for interface / calculation engine
 - contact@medchemica.com