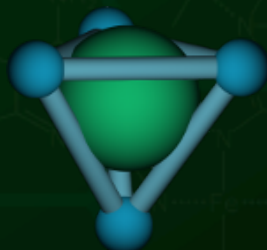


Cheminformatics workflows using the mobile + cloud platform

Dr. Alex M. Clark

October 2013

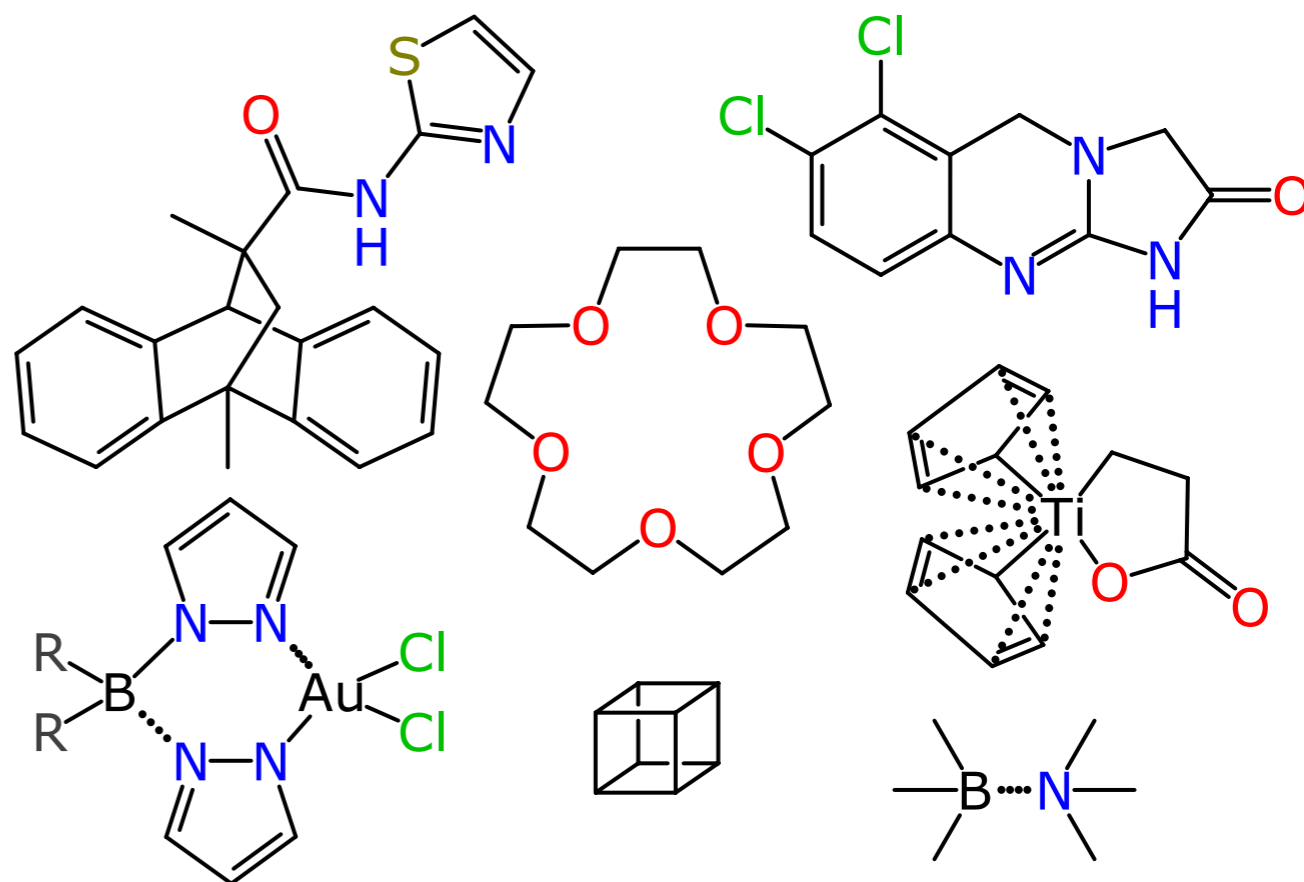


© 2013 Molecular Materials Informatics, Inc.

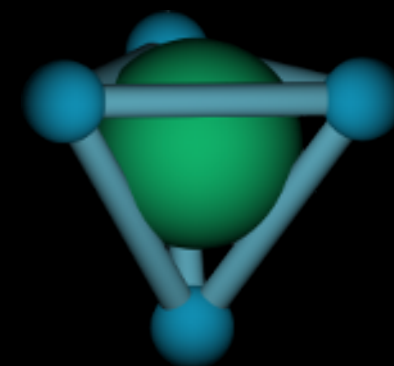
<http://molmatinf.com>

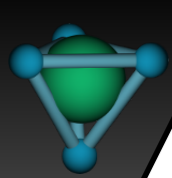
Overview

- Chemical structures
- Mobile apps
- Cloud computing



**MOLECULAR
MATERIALS
INFORMATICS**





3rd Revolution



mainframes
minicomputers

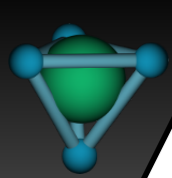


personal computers
portable laptops



mobile tablets
smartphones

- The mobile platform is revolutionary: **a clean break**
 - entirely new user interface
 - no backward compatibility
 - highly constrained resources
 - applicable to entirely new situations



Two platform stacks



desktop/
laptop



phone/
tablet



file/database
server



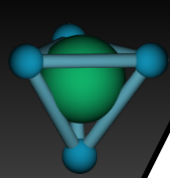
web
services



compute
cluster



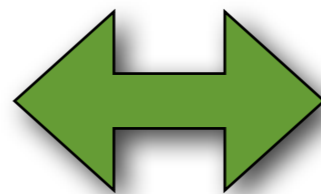
cloud
resources



Delivery

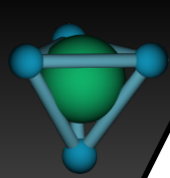
Apps

- User experience
- Cached data
- Light calculations
- Sharing



Cloud

- Web API
- Heavy calculations
- Large data
- Centralised storage



Challenges

- Mobile user interfaces are **completely different** to desktop products
- Expectations: functionality, extreme simplicity
- Complex interfaces much harder: drawing chemical structures requires redesign
- Visualisation is more difficult: screen size
- Storage limitations: big data is a problem
- Calculations split *mobile vs. cloud*

App Catalog



**Mobile Molecular
DataSheet (MMDS)**



MolSync



SAR Table



MolPrime+



MolPrime



Green Solvents



Lab Solvents



Approved Drugs



Reaction101



Yield101



Living Molecules



TB Mobile



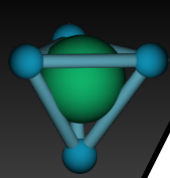
ChemSpider



SPRESImobile

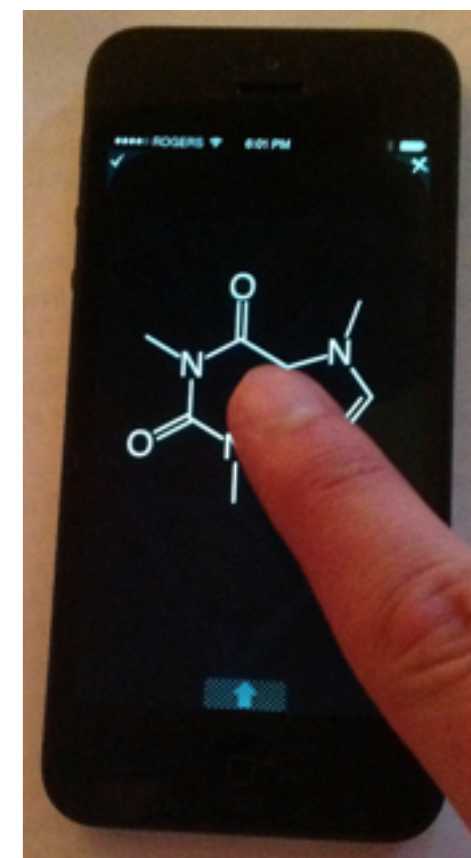


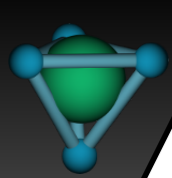
**Open Drug
Discovery Teams**



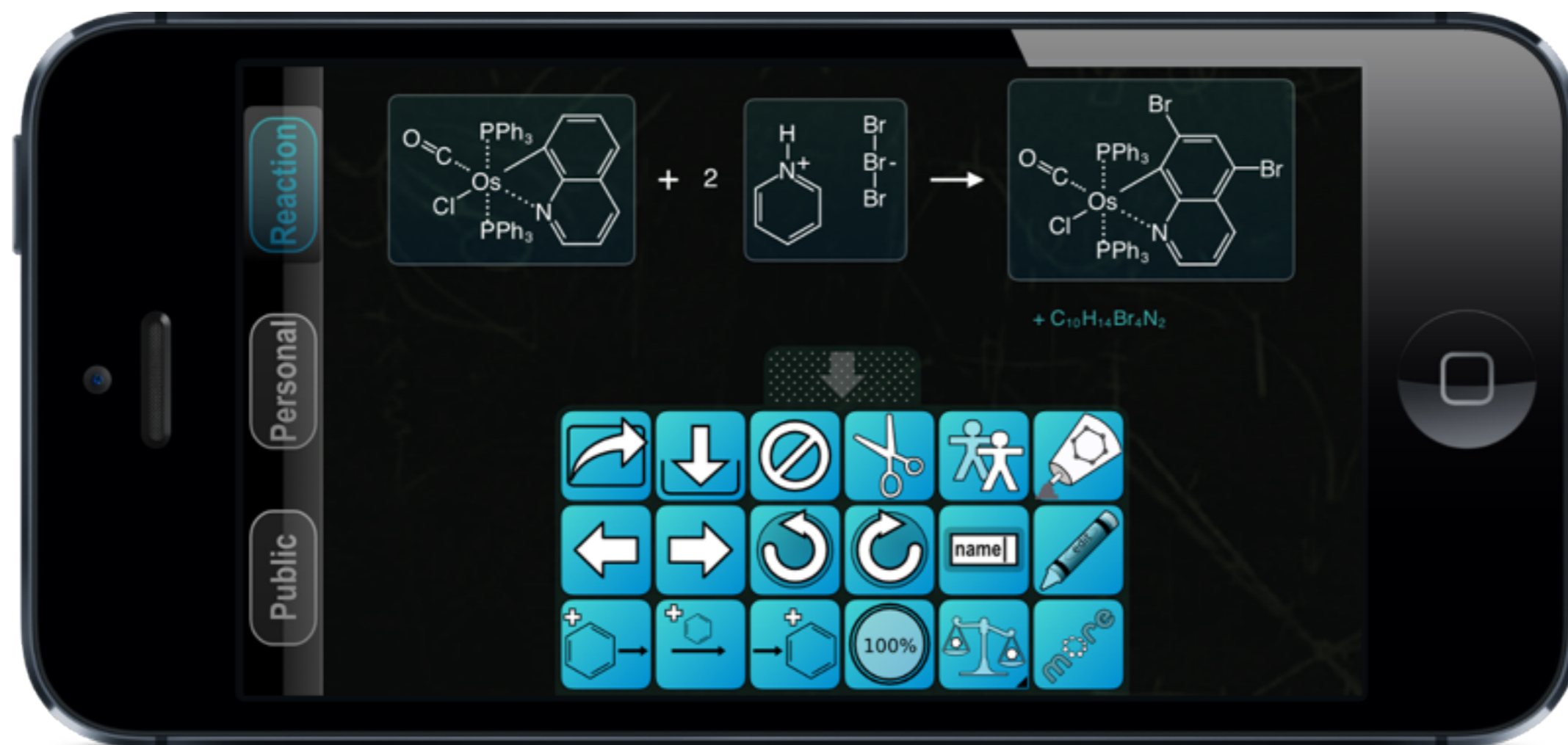
Drawing Structures

- Cheminformatics requires a user interface for drawing structures
- Interface needs to be:
 - **capable** for complex structures
 - **publication** quality
 - **fast** to use
 - **phone**-size form factor
- Traditional interface paradigm on touchscreen: unusable...





Drawing Reactions



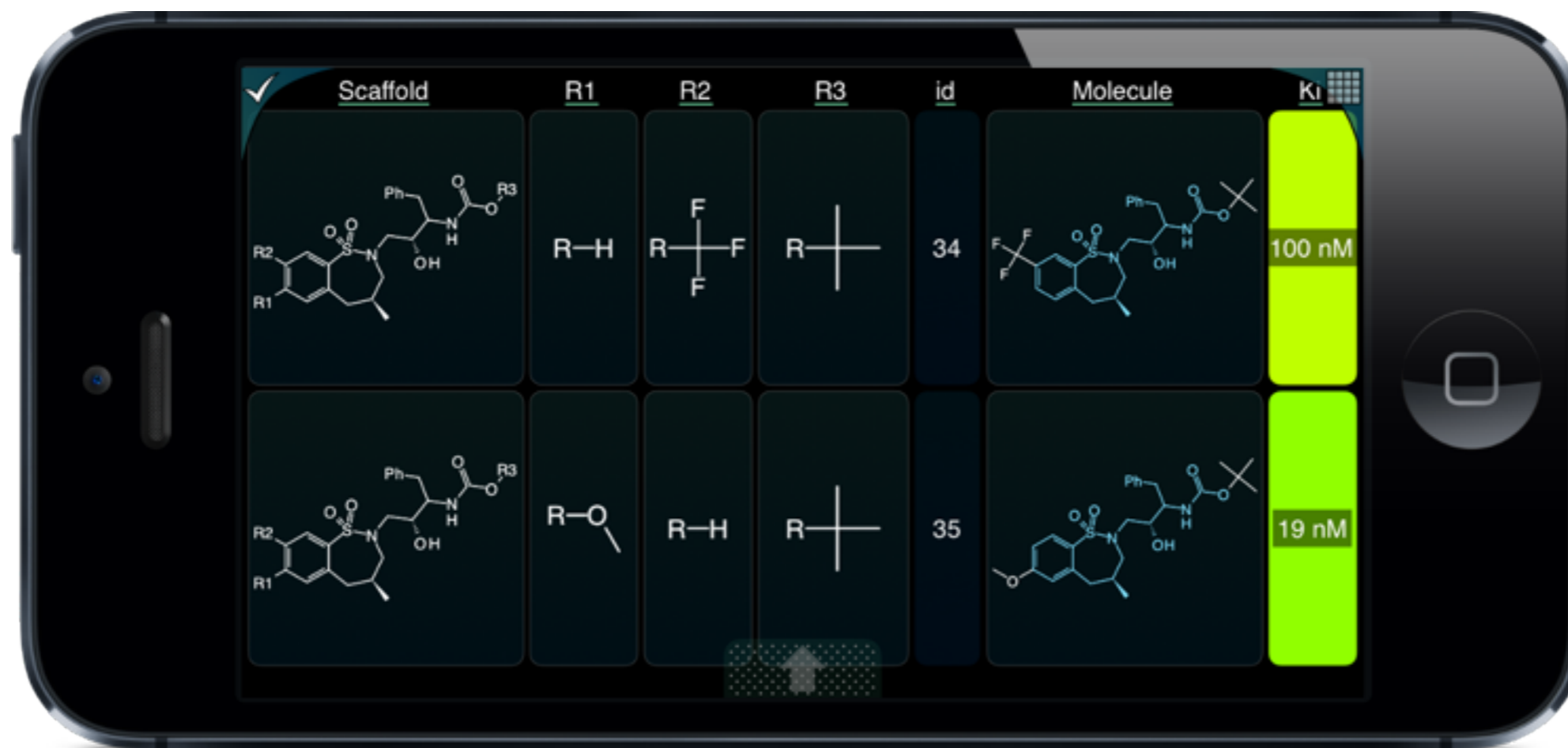
- Components: reuse sketcher, iPhone friendly
- Higher level of markup useful for lab notebooks
- Used by **MMDS**, **Reaction101**, **Yield101**

Assembling DataSheets

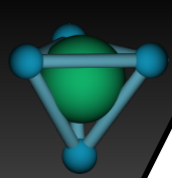


- **MMDS**: *molecular datasheet* unit
 - molecular structures
 - reaction schemes
 - scalar data (text, numbers)
- Table-like structure, roughly equivalent to SDfile/RDfile
- Operate on individual rows or whole datasheets

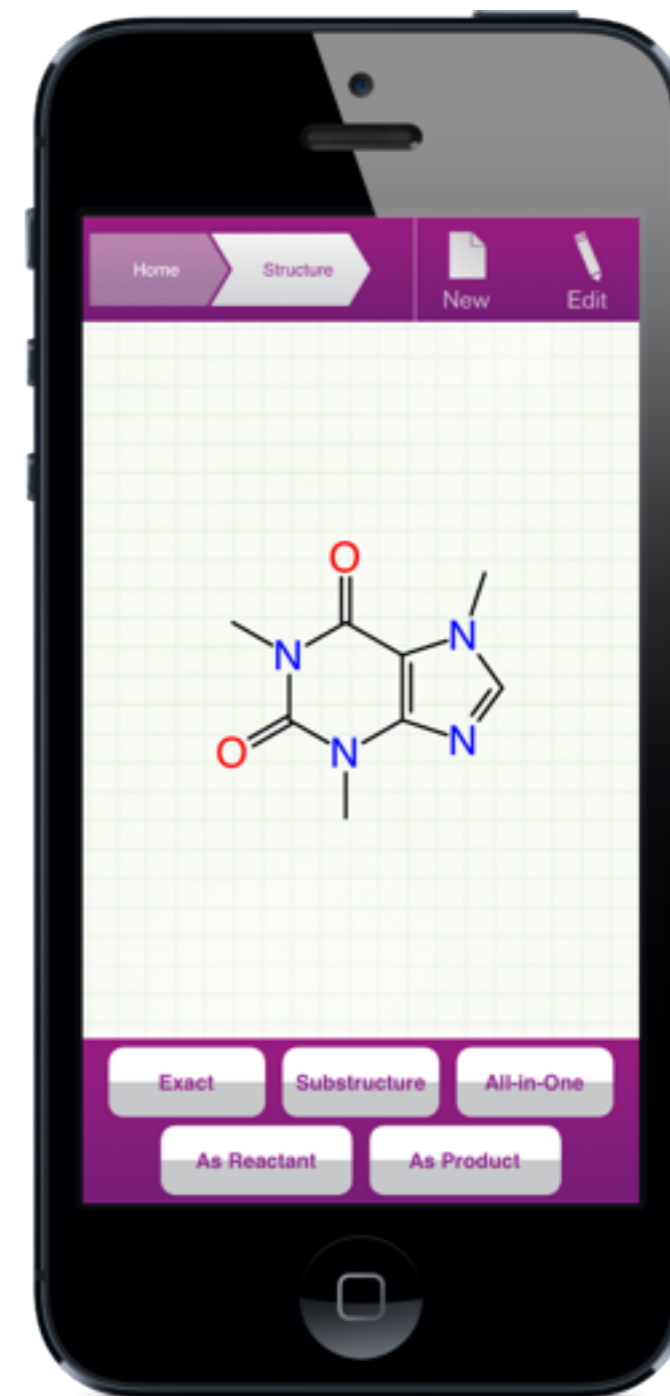
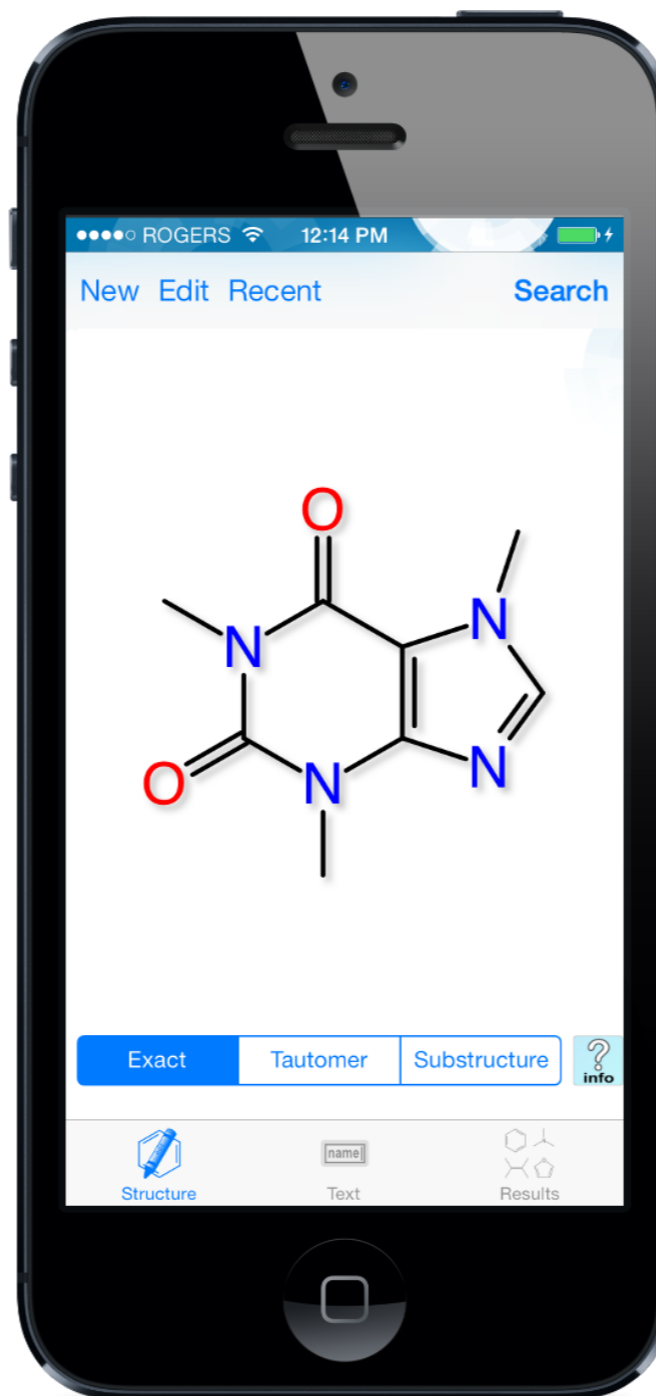
Scaffolds & Substituents



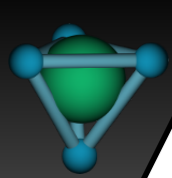
- Scaffold-centric model useful to drug discovery
- **SAR Table** app: optimised for data entry
- And: analysis, visualisation and calculation



Searching



- App interfaces for structure searching



Calculation



Molecule

The molecule column will be used to calculate the requested numeric properties, for each row:

Molecule

Host: <http://molsync.com/MolSync>

Properties

- MW
- MF
- Melting
- logP
- MR
- TPSA

Cancel Calculate

- Structures used to calculate properties
- Single **molecules**, or whole **datasheets**
- Simple calculations: done by app
- Difficult calculations: use webservice

Importing

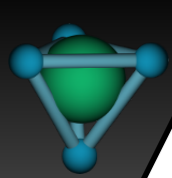
- Receive recognised filetypes
 - clipboard
 - email
 - web
 - other apps (IPC)
 - webservices
 - remote filesystems
- Documented formats, e.g. SketchEl, MDL, ChemDraw, CML



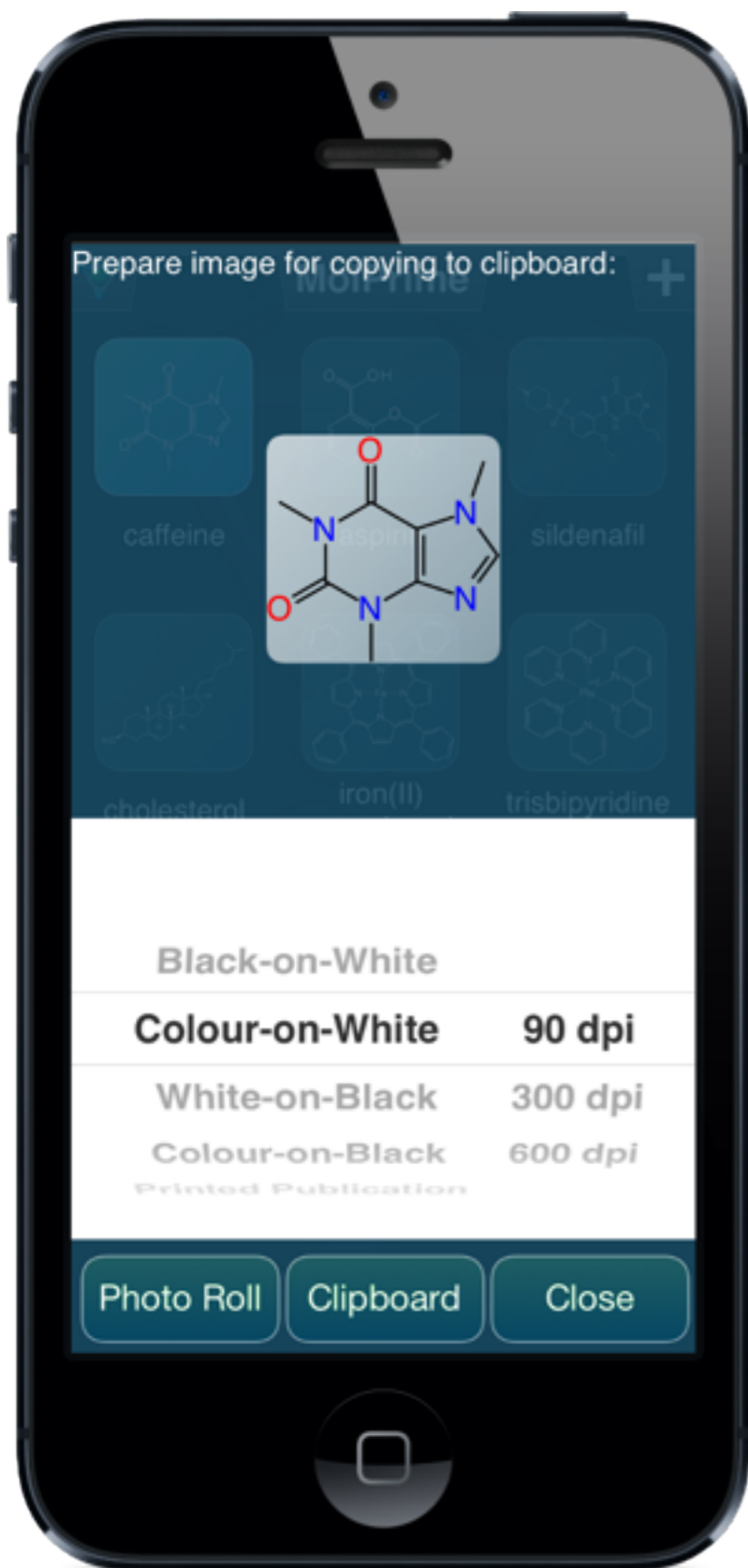
Exporting



- Apps are only as good as their ability to *use* the data
- Export actual data using cheminformatics formats
 - clipboard
 - email
 - web sharing
 - twitter
 - webservice
 - remote filesystems

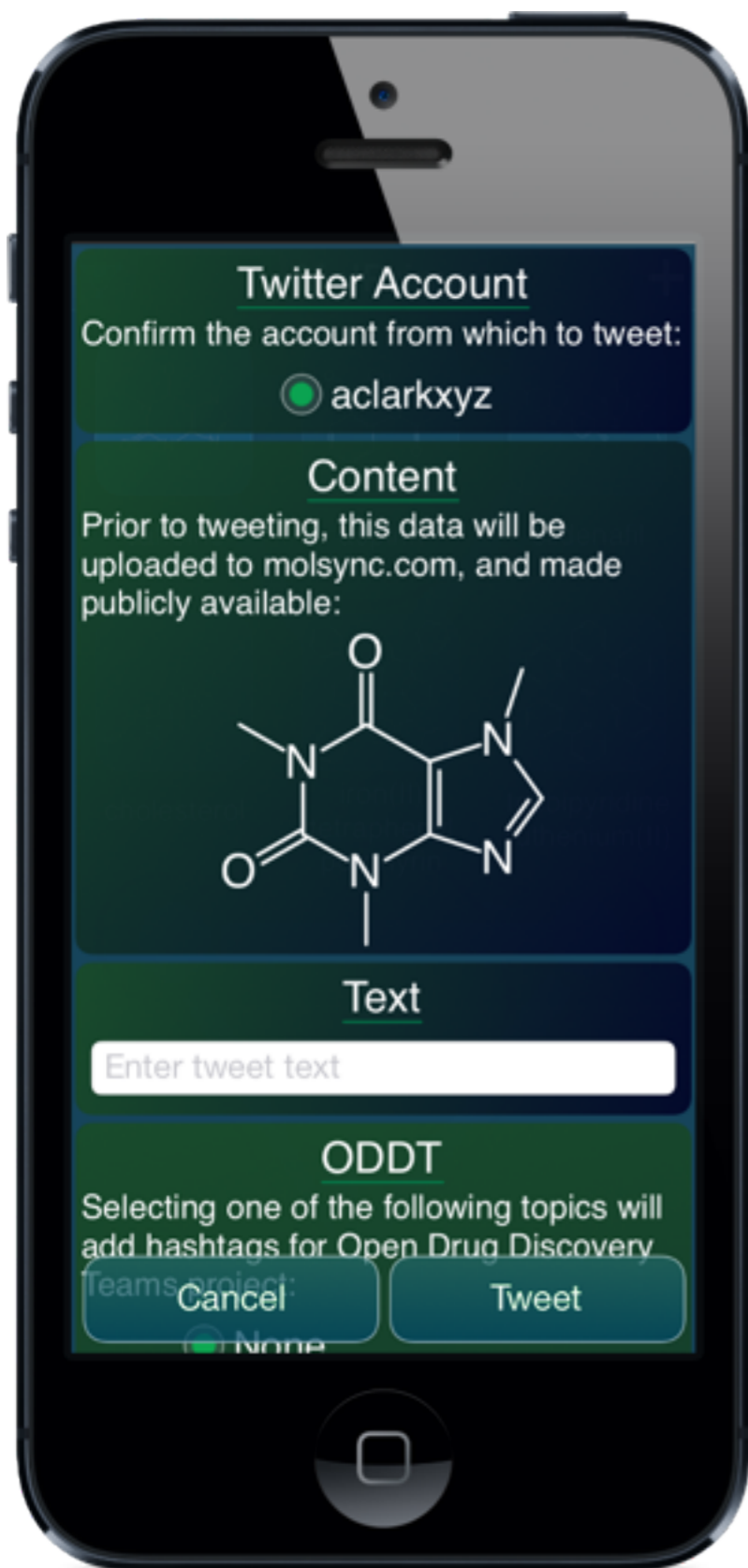


Graphics



- Prepare bitmapped graphics for:
 - clipboard
 - photo roll
 - email
- Can incorporate into other apps on the device (e.g. *Keynote*)
- Vector graphics: SVG, EPS, Microsoft Word & Excel

Sharing



- Share data publicly or privately
- Upload to a webservice
 - store raw data
 - dynamically render
 - hosted by *molsync.com*
- Tweet directly from the app: share the link

Real world workflow: searching for new Tuberculosis drugs

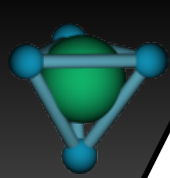
Dr. Alex M. Clark

October 2013



© 2013 Molecular Materials Informatics, Inc.

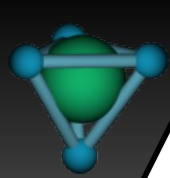
<http://molmatinf.com>



Outline

- Workflow is representative of lead discovery
- Searching for new *Mycobacterium tuberculosis* drugs
- Open data, open science, collaborating with experimentalists
- Workflow involves:
 - mobile apps
 - cloud-based webservices
 - traditional software

<http://molmatinf.com/venice.html>

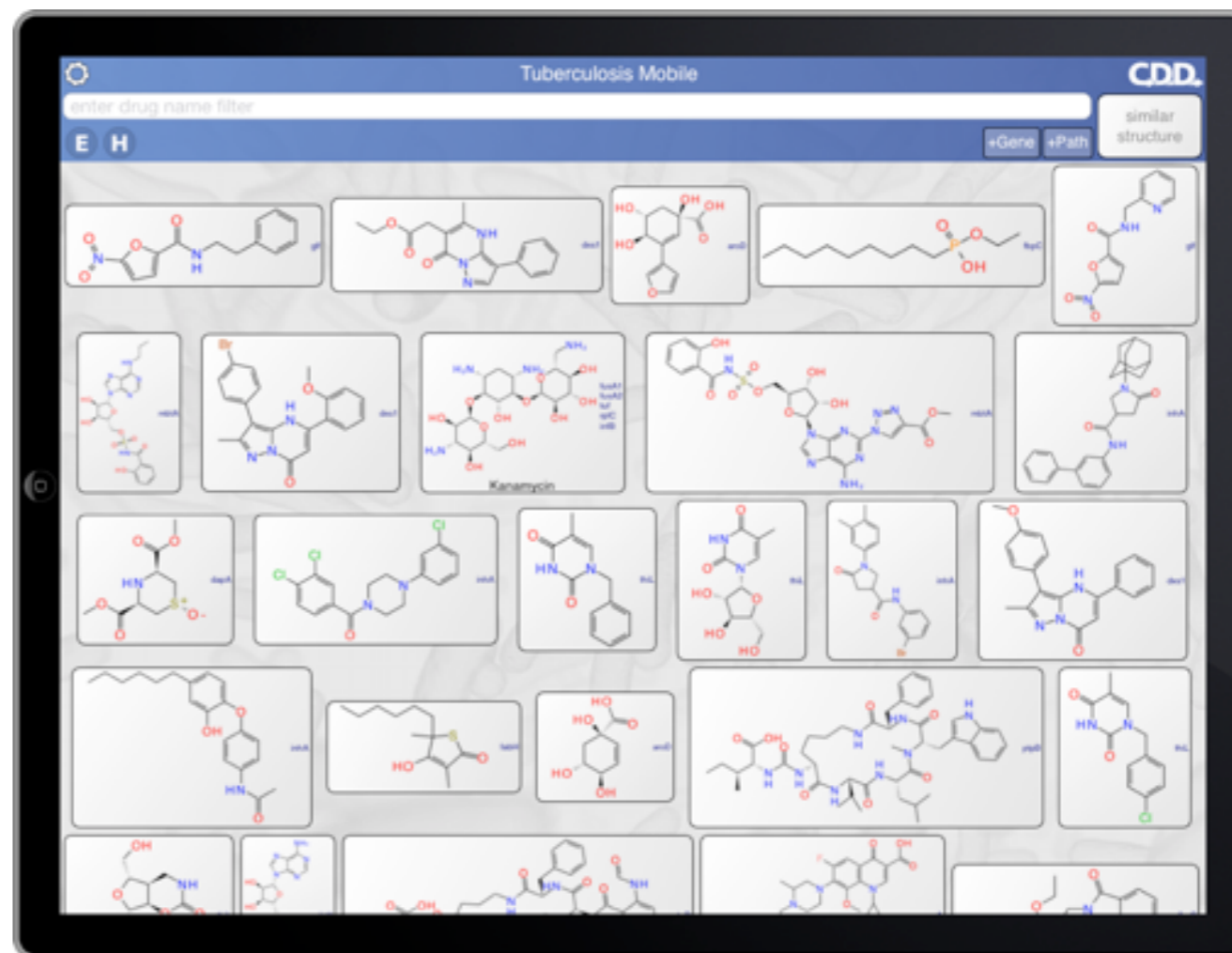


TB Mobile

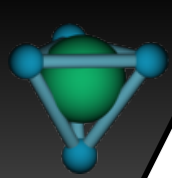
- Curated data for **tuberculosis** drugs

- iOS & Android

- CDD, NIAID grant, Sean Ekins

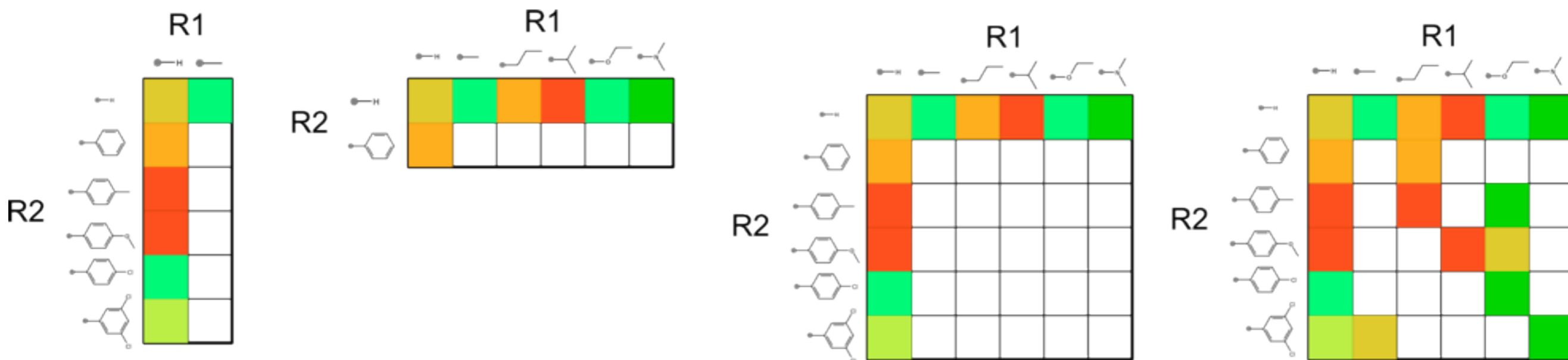
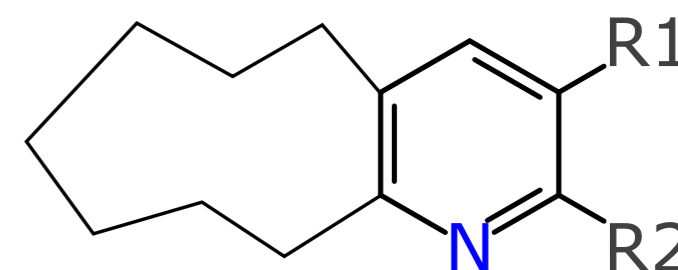


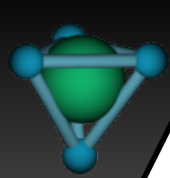
- App presents a browseable list of structures
- Can filter by target, sort by structural similarity
- Visual data exportable....



A Good Scaffold

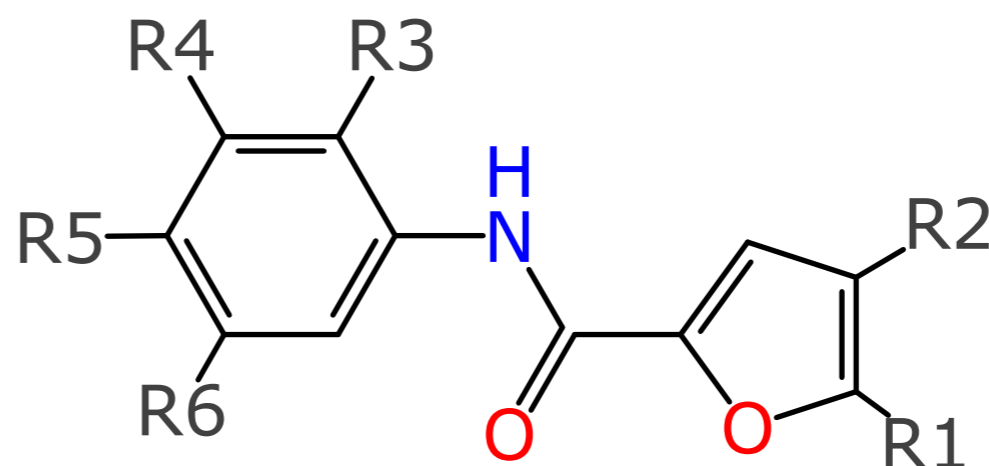
- Objective: use the data from **TB Mobile** to find a scaffold, then use to infer gene target from a larger collection of SAR data
- What makes a good scaffold?



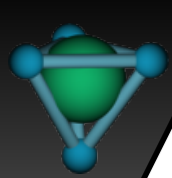


Source Materials

- Scaffold:



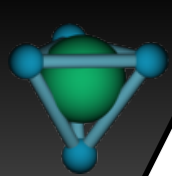
- Scaffold origin: inhibitor of **Glf** target
- 87 molecules with *in vitro* activity (yes/no)
- Scaffold seems to elicit an activity pattern
- Next step: load it into the app ecosystem...



Importing



- Email attachment to self: open in **SAR Table**

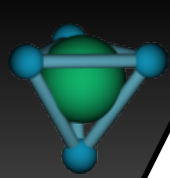


Raw DataSheet

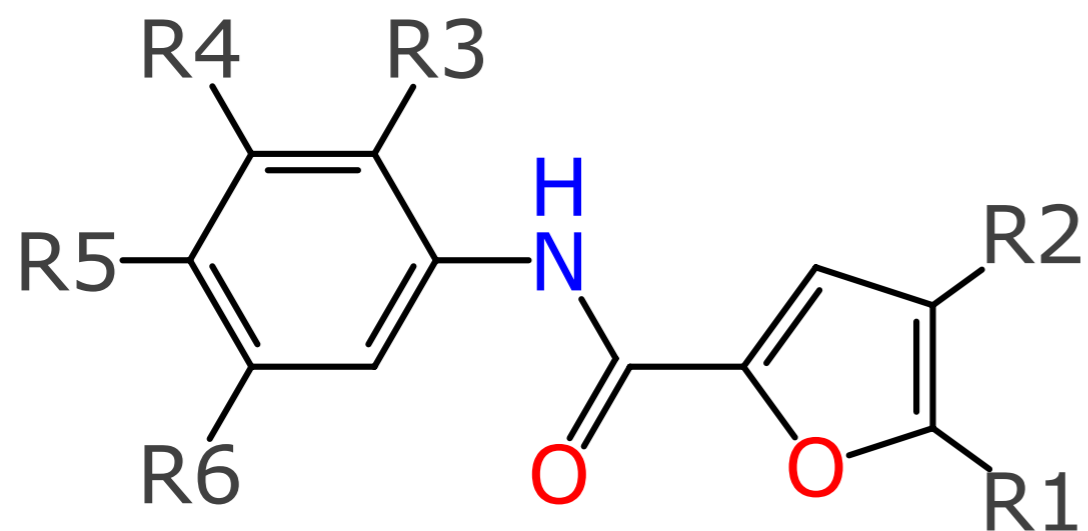


catfold	Molecule	Activity
		0
		0
		0
		0

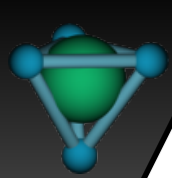
- Just **Molecule** and **Activity**...



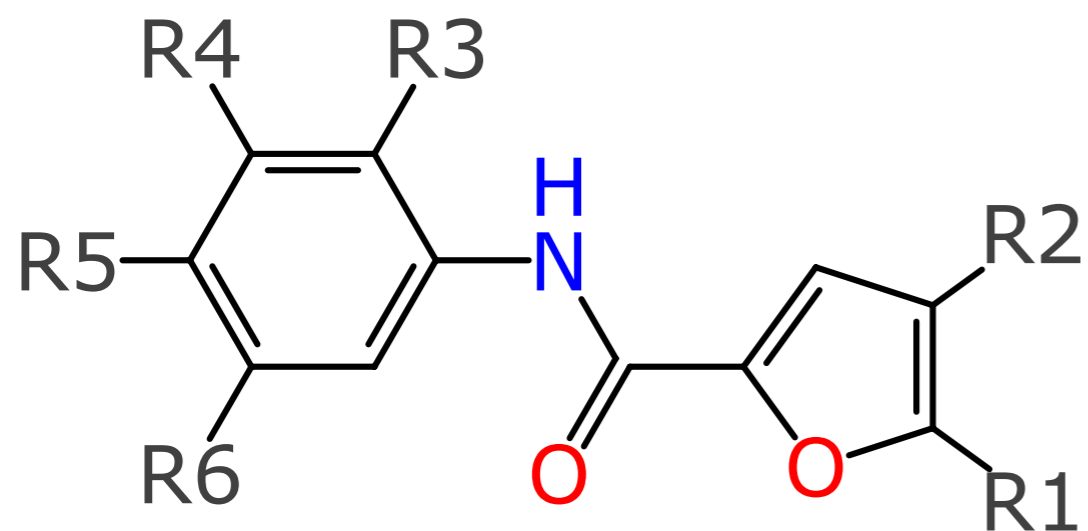
Draw Scaffold



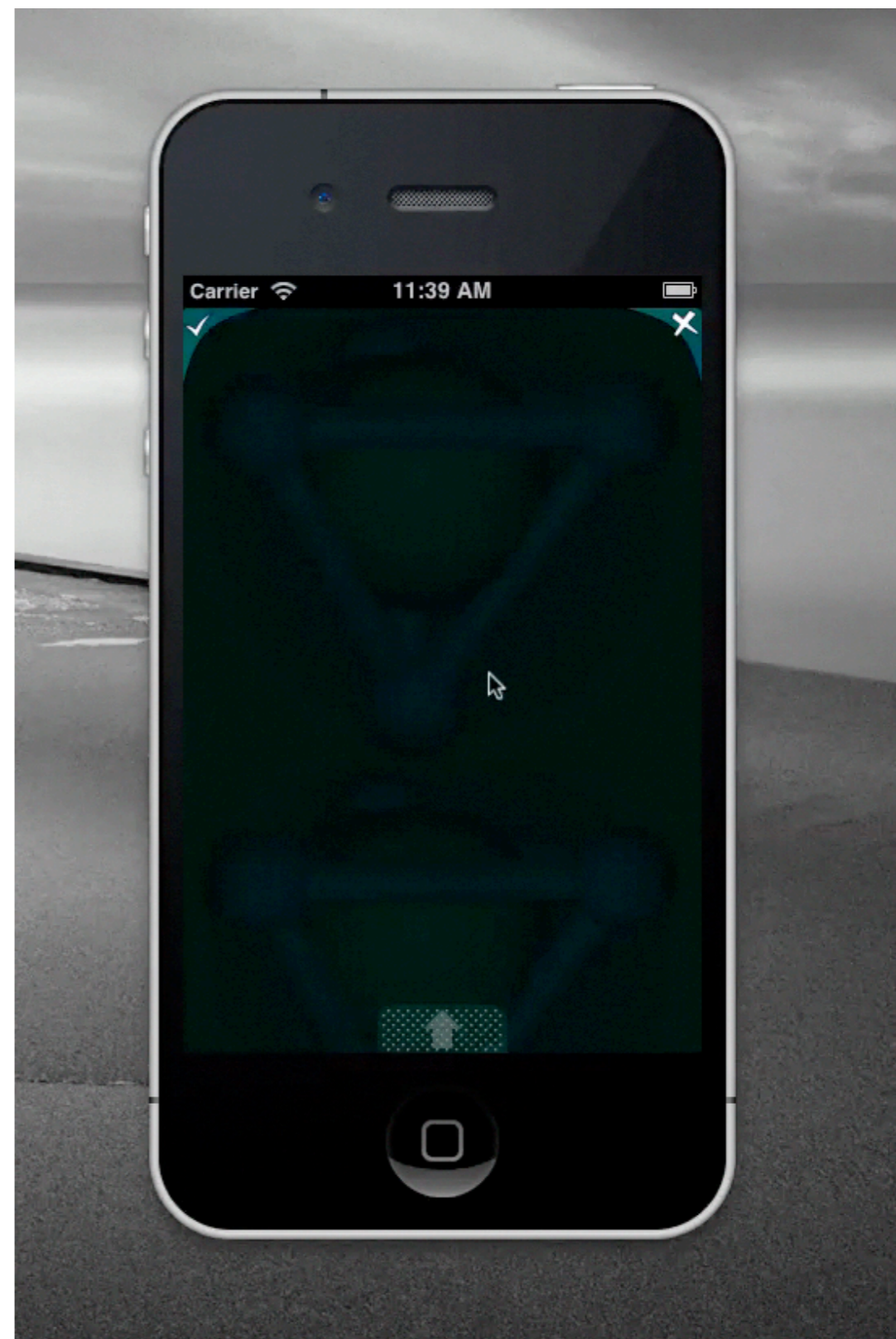
- Select first row
- Use built in sketcher to draw scaffold



Draw Scaffold

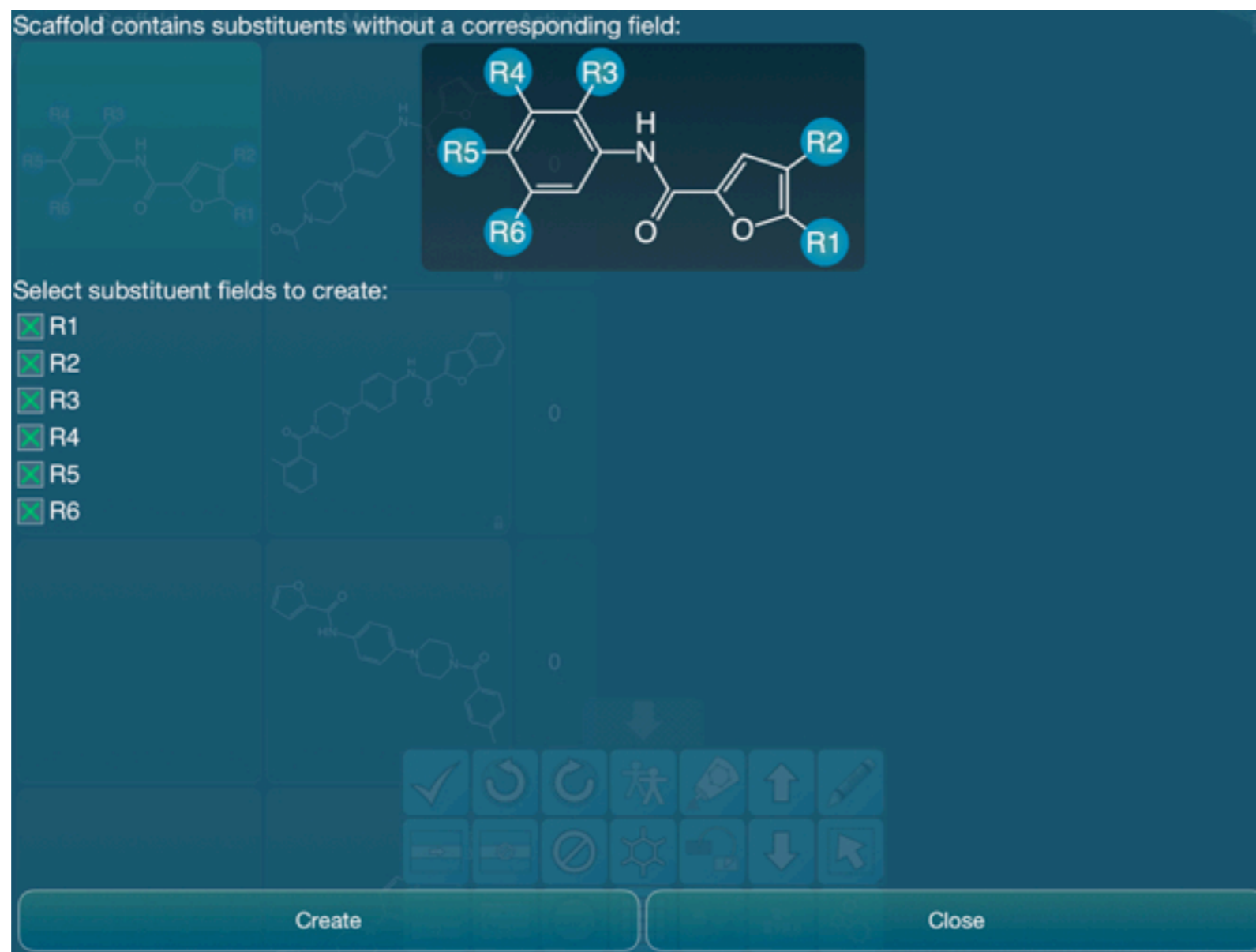


- Select first row
- Use built in sketcher to draw scaffold



Substituent Fields

Scaffold contains substituents without a corresponding field:



Select substituent fields to create:

- R1
- R2
- R3
- R4
- R5
- R6

Create Close

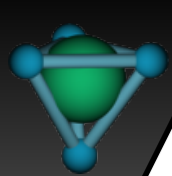
- R-group labels: offer to create fields

Scaffold & Molecule

The screenshot displays a software interface for molecular informatics. It features a table with columns for Scaffold, R1, R2, R3, R4, R5, R6, Molecule, and Activity. The first row is highlighted in green and shows a scaffold with six substituent positions (R1-R6) and a corresponding molecule with a bromine atom. The second and third rows show 'n/a' for the scaffold and '0' for the activity. The fourth row shows a molecule with a different substituent. A toolbar with various icons is visible at the bottom of the interface.

Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Activity
	?	?	?	?	?	?		0
n/a	n/a	n/a	n/a	n/a	n/a	n/a		0
n/a	n/a	n/a	n/a	n/a	n/a	n/a		0
n/a	n/a							0

- Need to assign substituents: automatically

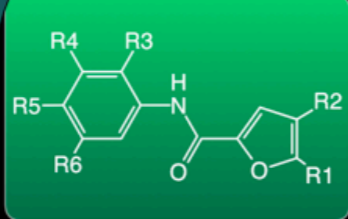

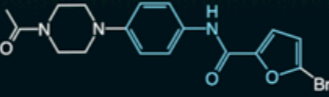
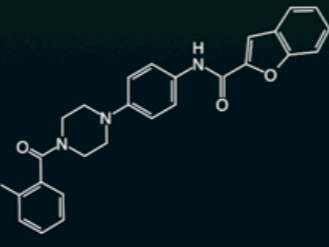
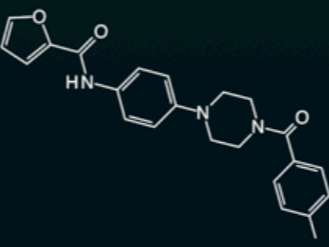
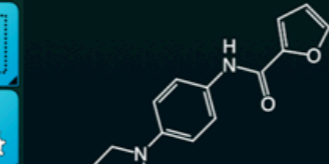



Scaffold Matching



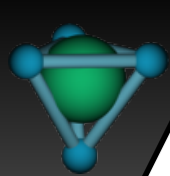
- Perform single **scaffold:molecule** match

Scaffold Matching

Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Activity
	R-Br	R-H	R-H	R-H		R-H		0
	n/a	n/a	n/a	n/a	n/a	n/a		0
	n/a	n/a	n/a	n/a	n/a	n/a		0
	n/a	n/a	n/a	n/a	n/a	n/a		0



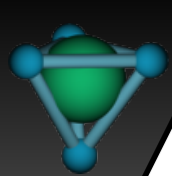
- Perform single **scaffold:molecule** match



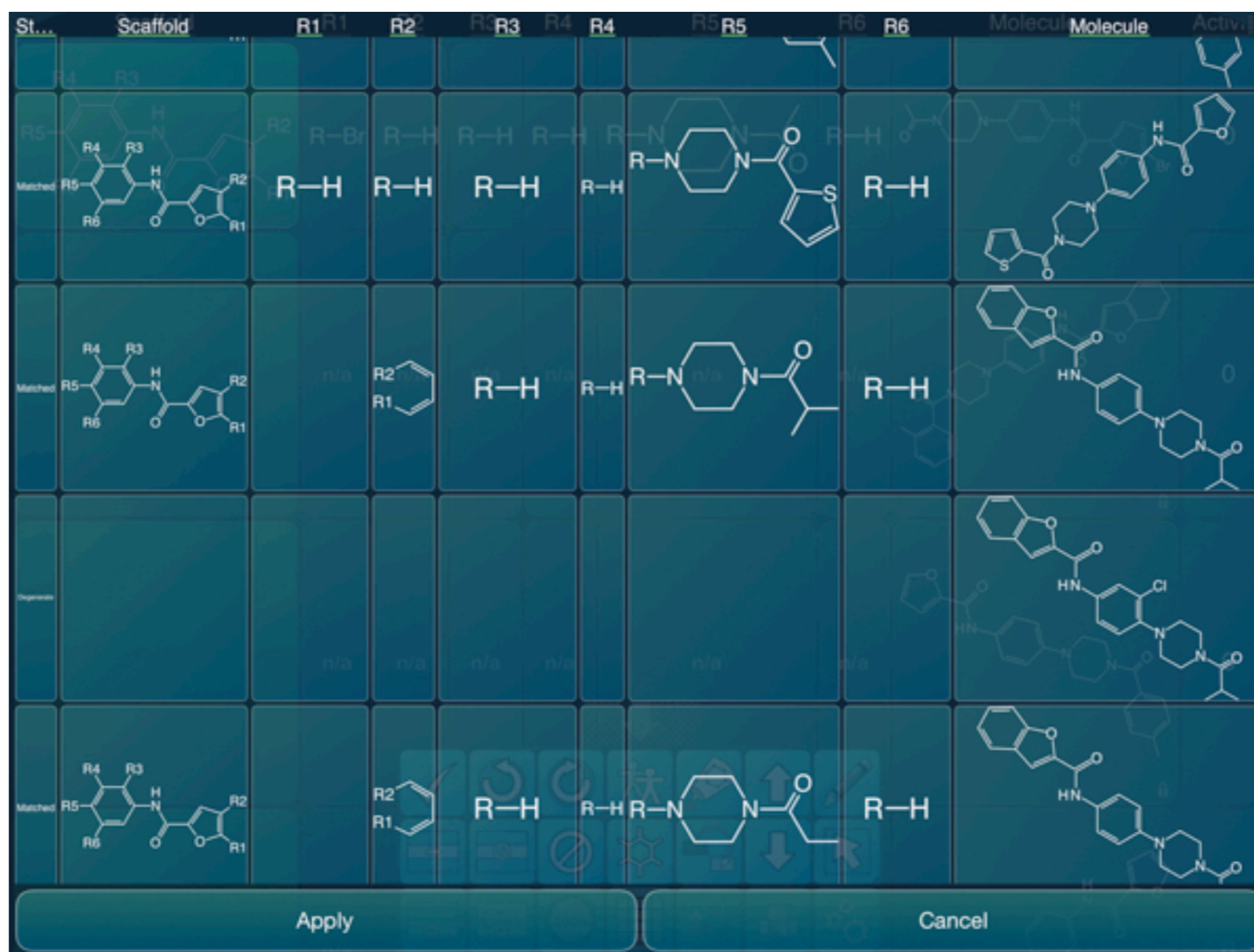
Bulk Matching

The screenshot displays a software interface for bulk matching scaffolds. At the top, a table titled "Match scaffolds:" has columns for Scaffold, R1, R2, R3, R6, Molecule, and Activity. The first row shows a scaffold with R1-R6 labels, R1=R-Br, R2=R-H, R3=R-H, R6=R-H, a molecule structure, and an activity of 0. Below this, a section labeled "to constructed molecules:" shows a grid of molecule structures. At the bottom, there are "Submit" and "Cancel" buttons, and a "Host: molsync.com" watermark.

- Duplicate scaffold, match to all molecules

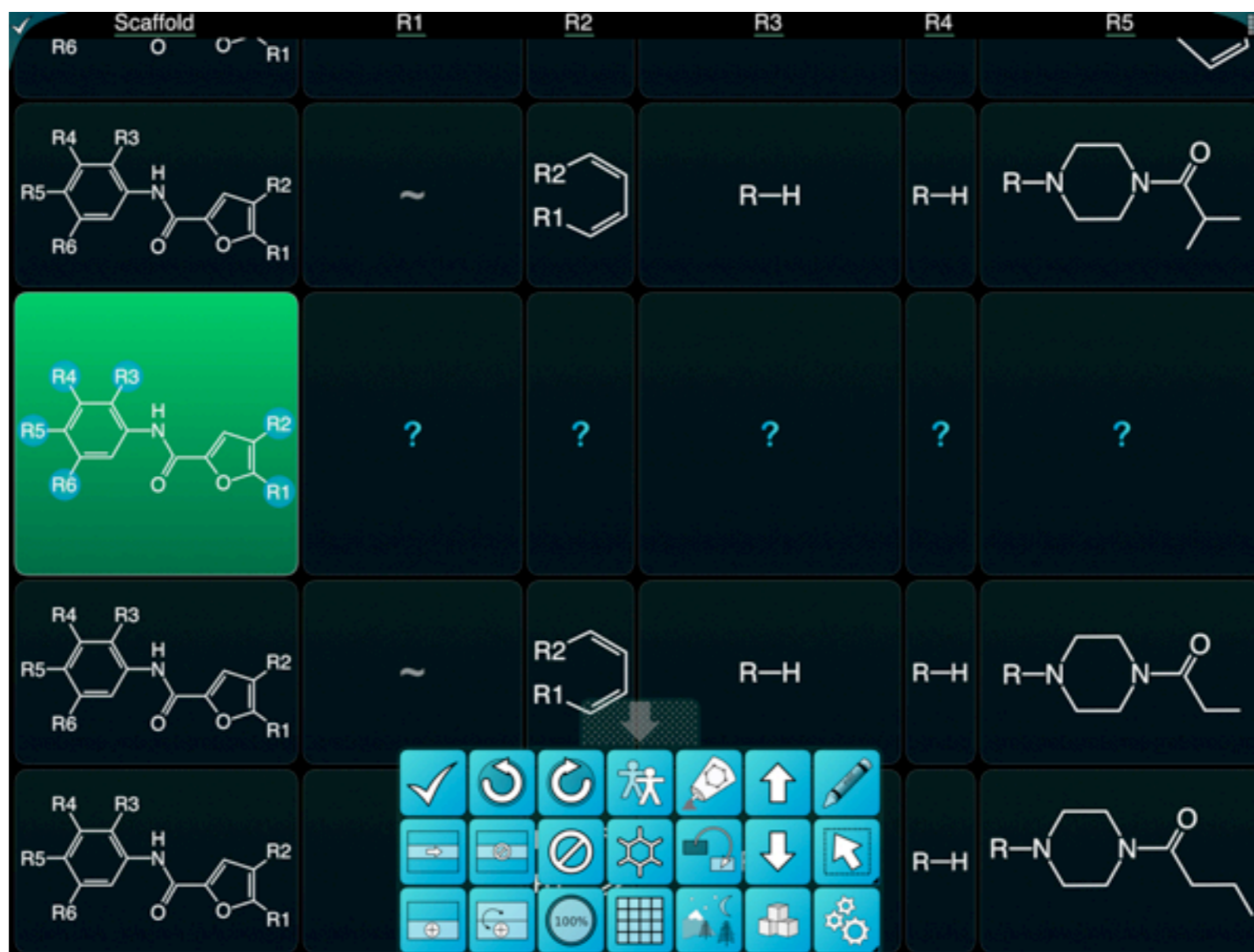


Bulk Results



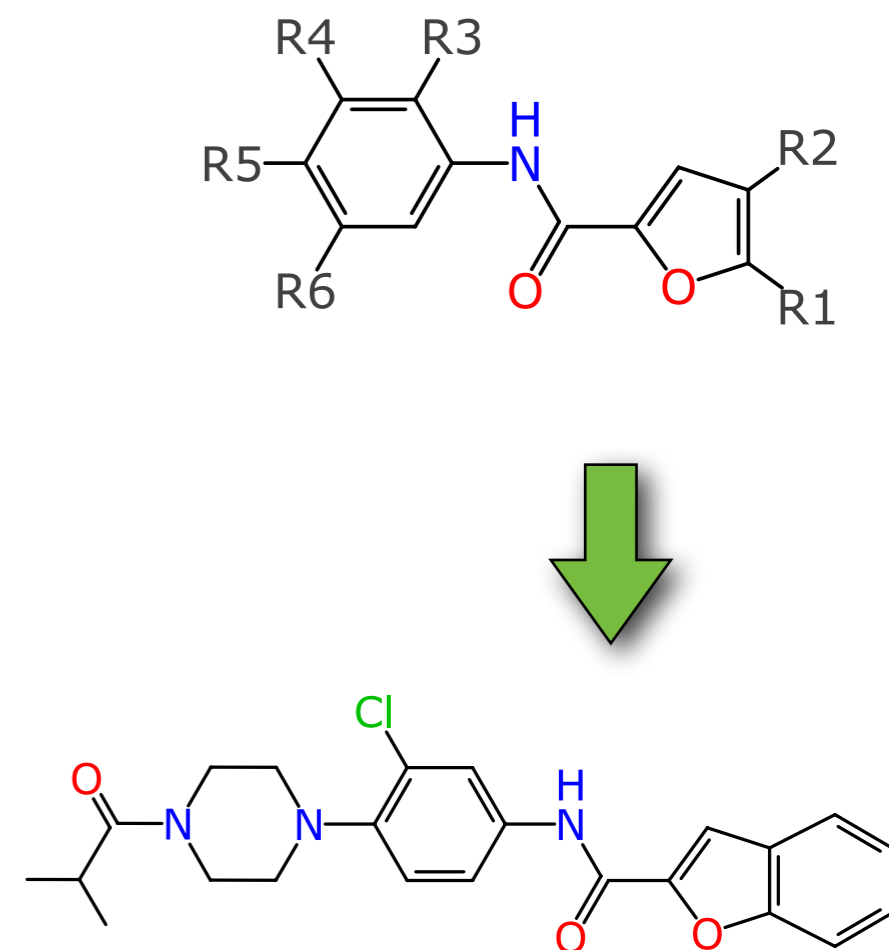
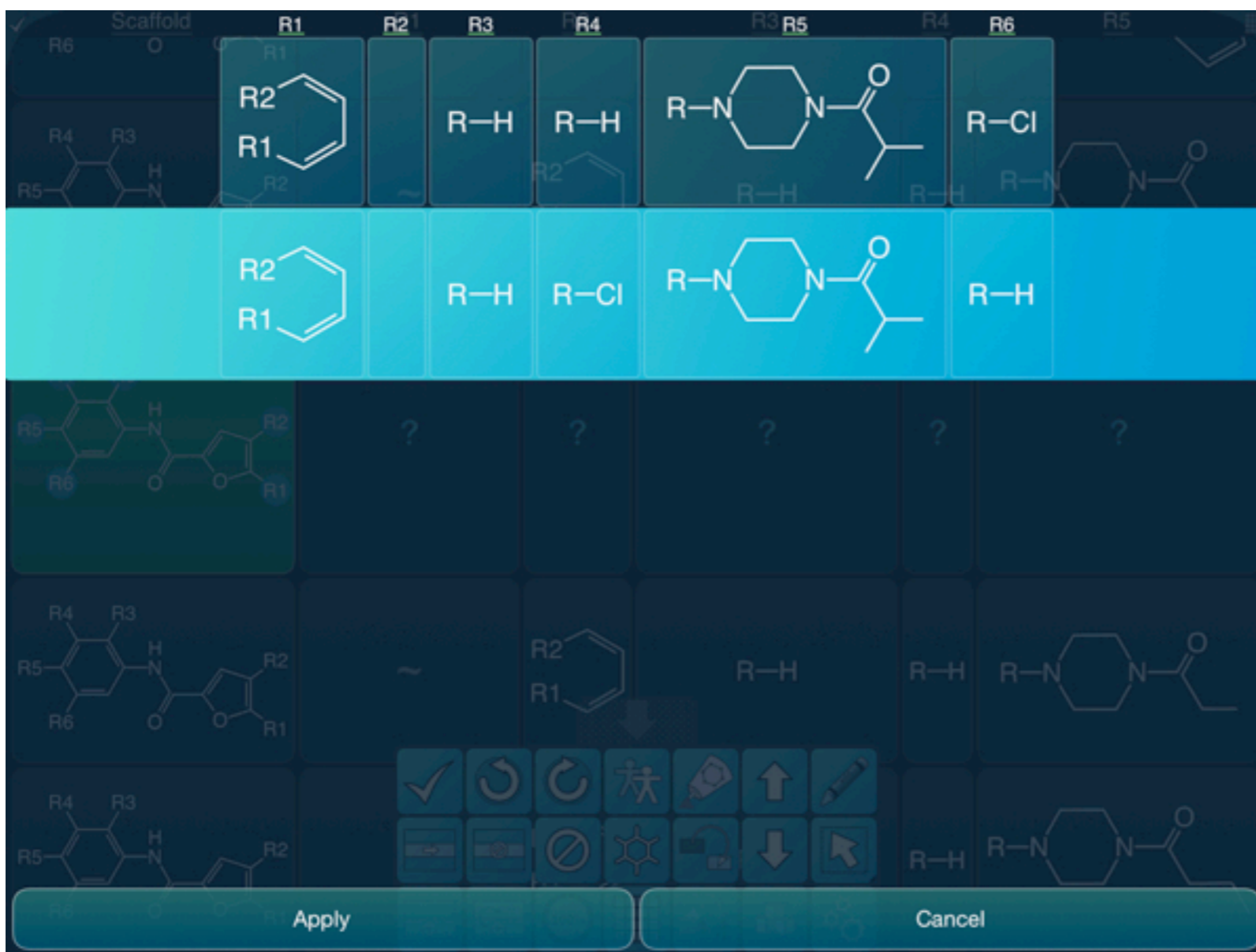
- Only unambiguous results are provided

Bulk Results



- Only unambiguous results are provided

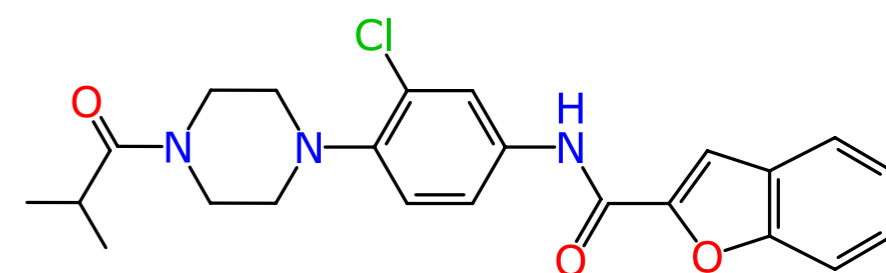
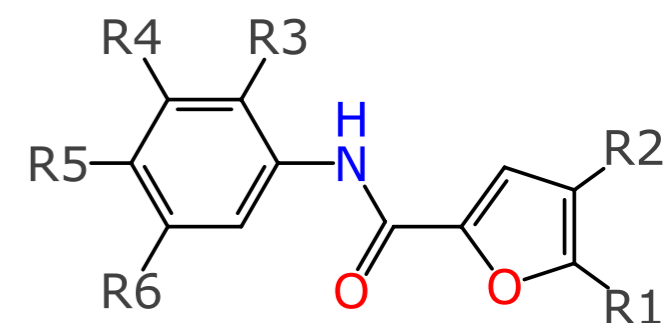
Symmetry/Degeneracy



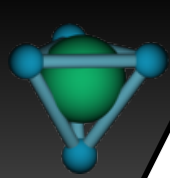
- Two possible results: note degeneracy **R4, R6**

Symmetry/Degeneracy

Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Activ
	R-H	R-H	R-H	R-H		R-H		0
	~		R-H	R-H		R-H		0
	~		R-H	R-Cl		R-H		0
	~		R-H	R-H		R-H		0
	~		R-H	R-H		R-H		0
		R-H	R-H	R-H		R-H		0
	n/a	n/a	n/a		n/a	n/a		0



- Two possible results: note degeneracy **R4, R6**



Finishing Assignment

- Single & bulk scaffold matching: semi-automated assignment, human tiebreakers
- Have 87 compounds, 21 active against TB
- Based on scaffold from a binder of **Glf** enzyme (essential for mycobacterial growth)
- What next?
- Look for other compounds: same scaffold

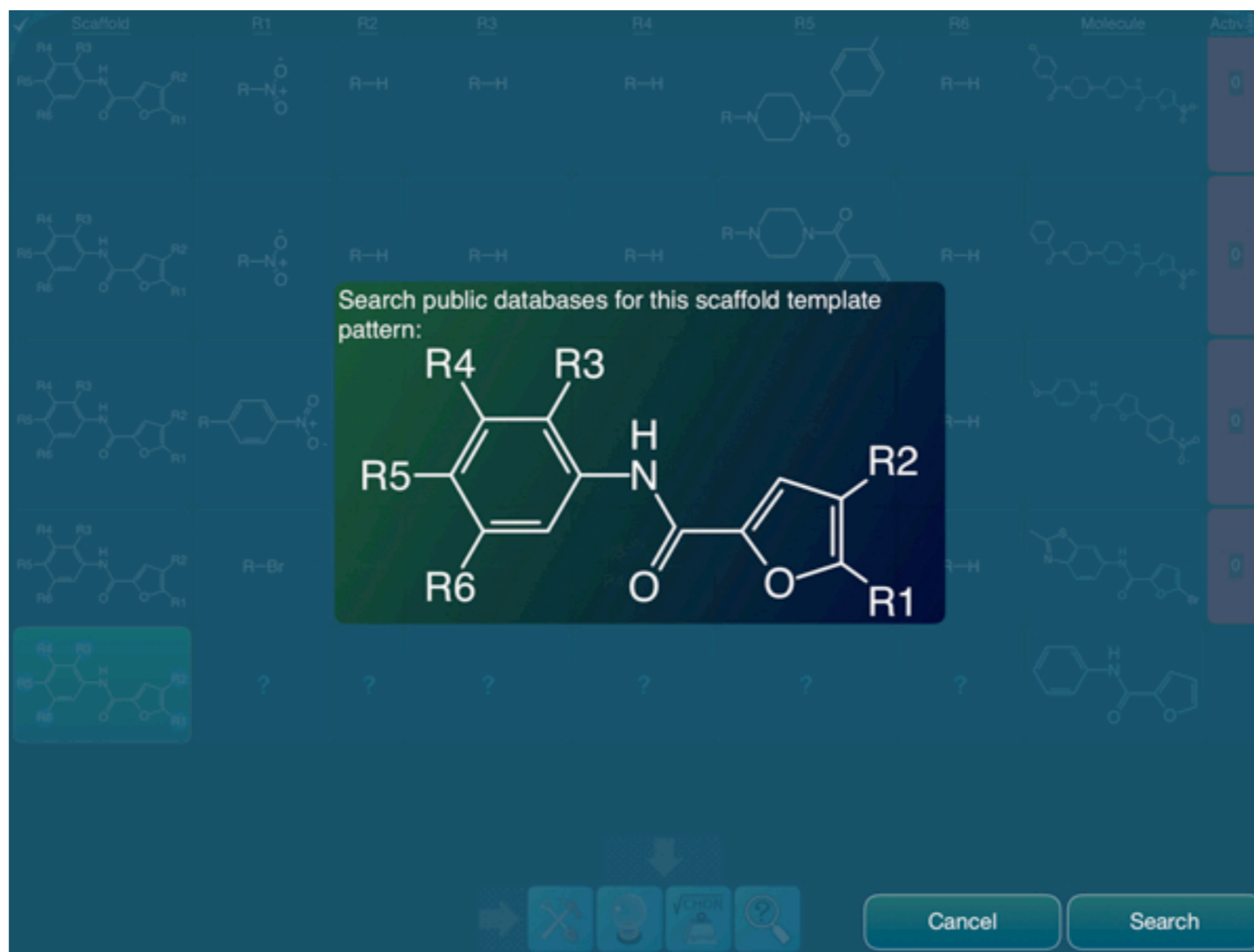
Searching

The screenshot displays a molecular search interface with a grid of templates. The grid has columns labeled Scaffold, R1, R2, R3, R4, R5, R6, Molecule, and Activ. The Scaffold column shows a complex heterocyclic structure with six substituent positions (R1-R6). The R1-R6 columns show various chemical groups or symbols. The Molecule column shows the resulting molecule for each template. The Activ column shows a red bar with a '0' in a white box. A toolbar at the bottom contains various icons for navigation and editing.

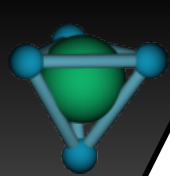
Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Activ
		R-H	R-H	R-H		R-H		0
		R-H	R-H	R-H		R-H		0
		R-H	R-H	R-H	R-O	R-H		0
	R-Br	R-H	R-H		~	R-H		0
	?	?	?	?	?	?		

- Search public databases using template

Searching



- Search public databases using template



Search Results

The screenshot displays a list of search results for furancarboxamide derivatives. Each result includes a chemical structure, a name, and a price icon. The results are as follows:

Chemical Structure	Name	Price Icon
	N-(3-chlorophenyl)-2-furancarboxamide	\$
	N-(3-chlorophenyl)-2-furancarboxamide	\$
	N-(5-chloro-2-hydroxyphenyl)-2-furancarboxamide	\$
	N-(2-methoxyphenyl)-2-furancarboxamide	\$
	N-(4-methoxyphenyl)-2-furancarboxamide	\$
	5-Nitro-furan-2-carboxylic acid (3-bromo-phenyl)-amide	\$

At the bottom of the screenshot, there are buttons for "Close" and "Import".

- Results come back with scaffolds assigned

Sources

The screenshot displays a chemical information page for N-(3-chlorophenyl)-2-furancarboxamide. At the top, the chemical structure is shown. Below it, the name and synonyms are listed. The source is identified as PubChem with ID 282652. A section titled 'Vendors' lists several suppliers with their respective IDs and contact information. At the bottom, there is a 'Close' button and an 'Import' button.

Name: N-(3-chlorophenyl)-2-furancarboxamide
Synonyms: N-(3-chlorophenyl)furan-2-carboxamide
N-(3-chlorophenyl)-2-furamide

Sourced from PubChem
PubChem: [282652](#)

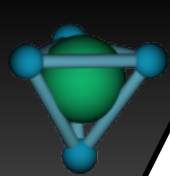
Vendors

ZINC	ZINC00102771
ChemExper Chemical Directory	JoCBPCShQBSJwwKWVI@QMMB@@
ChemBridge	5237258
Vitas-M Laboratory	STK873810
MolPort	MolPort-000-417-896
AKos Consulting & Solutions	AKOS001569463
TimTec	ST011030

5-Nitro-furan-2-carboxylic acid (3-bromo-phenyl)-amide

Close Import

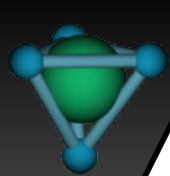
- Data from **ChemSpider**, **PubChem**, **ChEBI**
- Includes database links and vendor information



Importing

Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Act
		R-H	R-H	R-H		R-H		0
		R-H	R-H	R-H		R-H		0
	R-Br	R-H	R-H		~	R-H		0
	?	?	?	?	?	?		
	R-H	R-H	R-H	R-H	R-H	R-Cl		
	R-H	R-H	R-H	R-Cl	R-H	R-H		
	R-H	R-H	R-OH	R-H	R-H	R-Cl		
	R-H	R-H			R-H	R-H		

- Search results imported in marked-up form

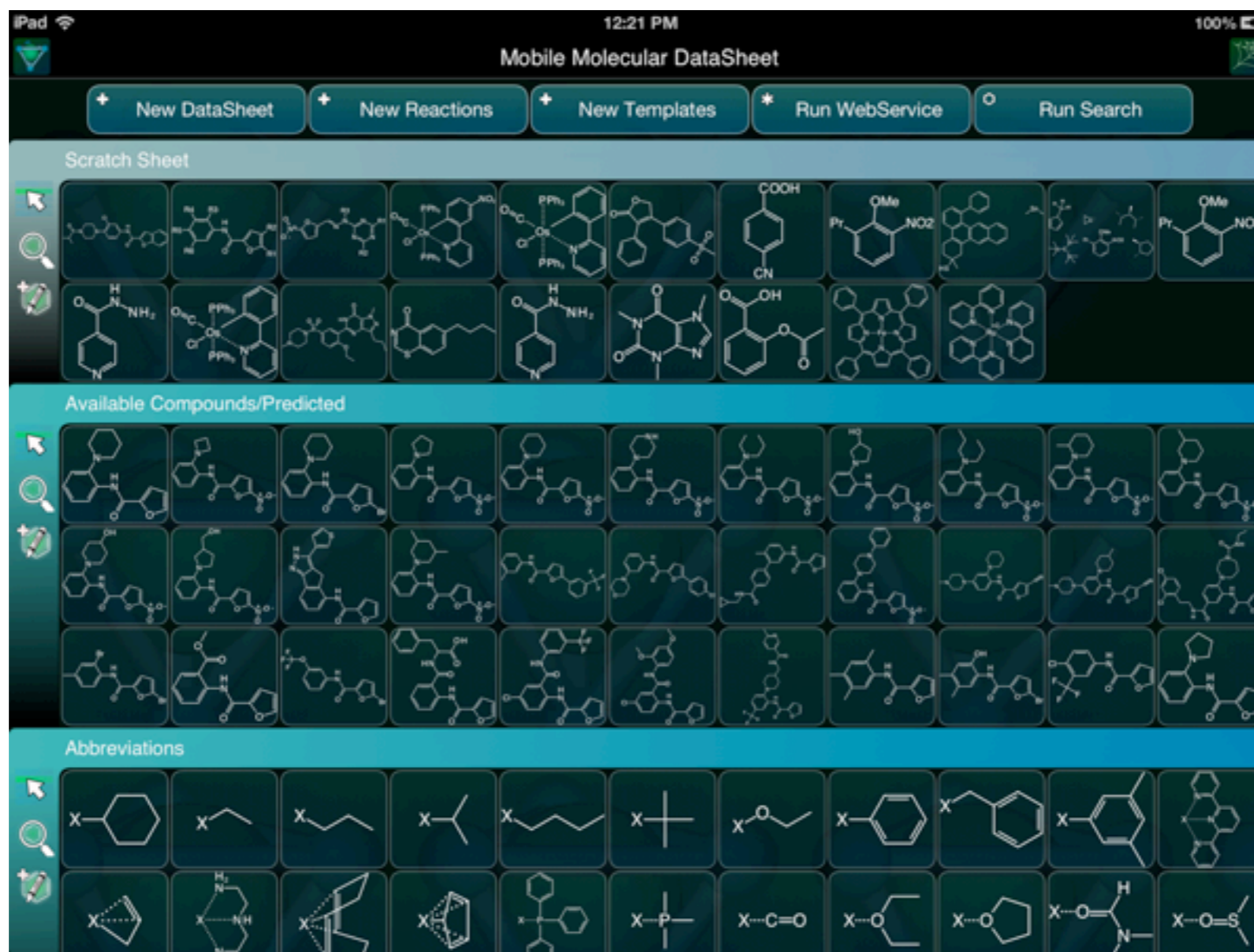


Model Application

Scaffold	Molecule	R1	R2	R3	R4	R5	R6	Active	Activity	CASRN	PubChemID	CHEBIID	CSID	Vendors
			R-H	R-H	R-H	R-O	R-H	False	0					
		R-Br	R-H	R-H		~	R-H	False	0					
		R-H	R-H	R-H	R-H	R-H	R-Cl	True	0.402	282652				ZINC, ChemExper, Chemical Directory, ChemBridge, Vitas-M Laboratory, MolPort, AKos Consulting & Solutions, TimTec
		R-H	R-H	R-O	R-H	R-H	R-H	True	0.402	303387				ZINC, ChemBridge, ASINEX, Vitas-M Laboratory, MolPort, AKos Consulting & Solutions, TimTec
		R-H	R-H	R-H	R-H	R-O	R-H	True	0.387	303388				ZINC, ChemBridge, ASINEX, Vitas-M Laboratory, MolPort, AKos Consulting & Solutions, ChemRoad Products, TimTec, ChemFrag
		R-H	R-H	R-H	R-H	R-	R-Cl	True	0.387	303389				ZINC, ASINEX, Specs, Vitas-M Laboratory, MolPort, AKos Consulting & Solutions, Enamine
			R-H	R-H	R-Br	R-H	R-H	True	0.407			CHEBI:413225		
			R-H	R-H	R-H	R-H	R-F	True	0.407			CHEBI:413226		
			R-H	R-H	R-H	R-H	R-Cl	True	0.407			CHEBI:413625		
			R-H	R-H	R-H	R-H	R-OH	True	0.337			CHEBI:413654		
		R-H	R-H					True	0.388			CHEBI:517764		

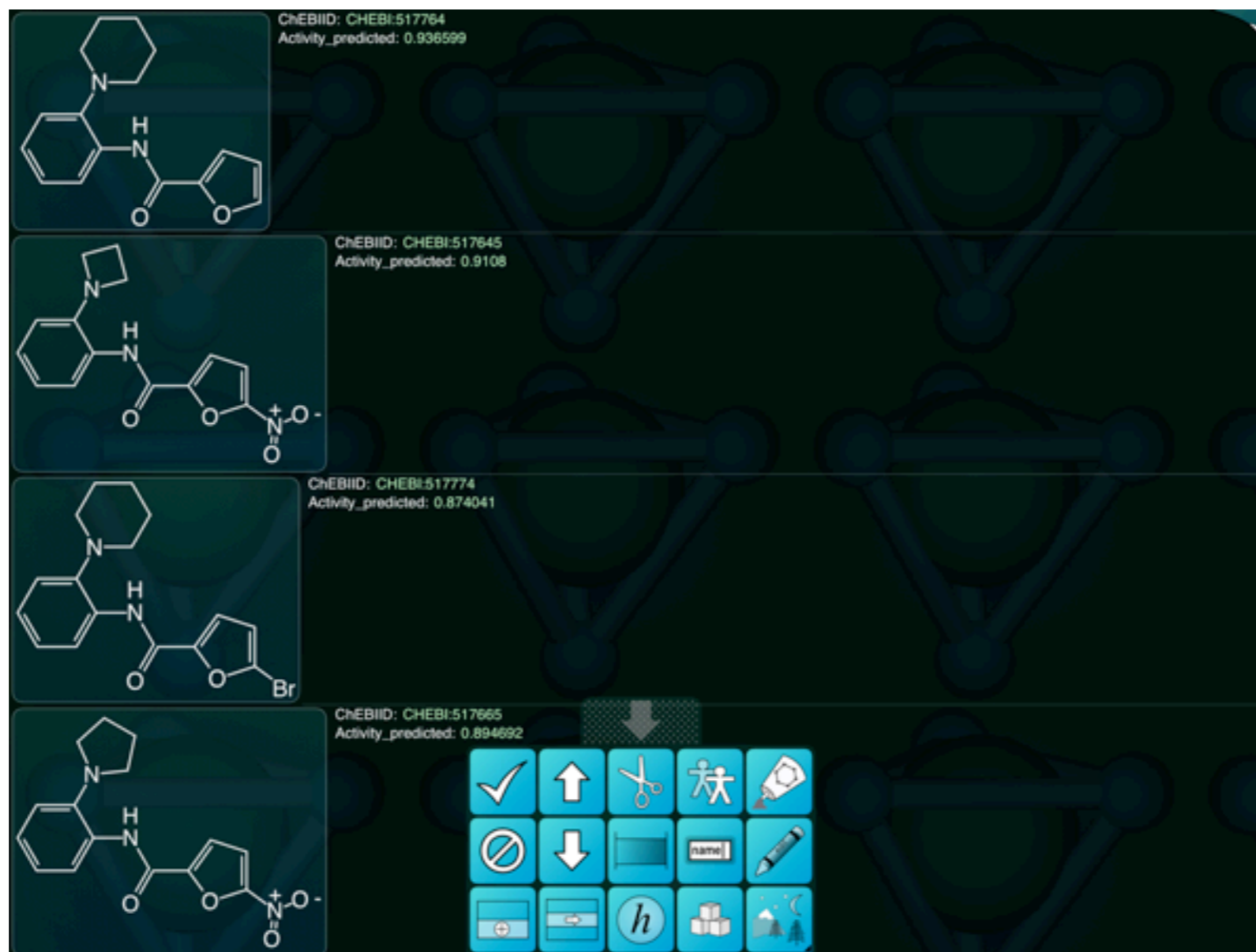
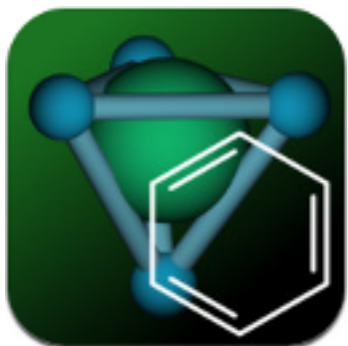
- Predicted activities shown as coloured wedges

Selected Candidates

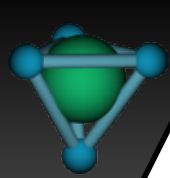


- Compile a list: predicted activity > 0.8
- Open in **Mobile Molecular DataSheet** (MMDS) app

Selected Candidates



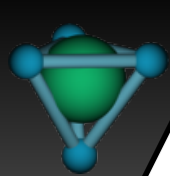
- Compile a list: predicted activity > 0.8
- Open in **Mobile Molecular DataSheet** (MMDS) app



Sharing by Email

The screenshot displays the 'Mobile Molecular DataSheet' application interface. At the top, there are navigation buttons: '+ New DataSheet', '+ New Reactions', '+ New Templates', '* Run Webservice', and 'o Run Search'. Below these is a 'Scratch Sheet' containing a grid of various chemical structures. Underneath is a section labeled 'Available Compounds/Predicted' with another grid of structures. An 'Export' menu is overlaid on the left side of the 'Available Compounds/Predicted' section, listing options: 'Export DataSheet', 'Email DataSheet', 'Open With', 'Tweet This', and 'Share on Web'. The 'Email DataSheet' option is highlighted. At the bottom, there is a third grid of chemical structures, including various functional groups and ring systems.

- Very easy to share using email attachments
- Recipient can open on any device, including mobile

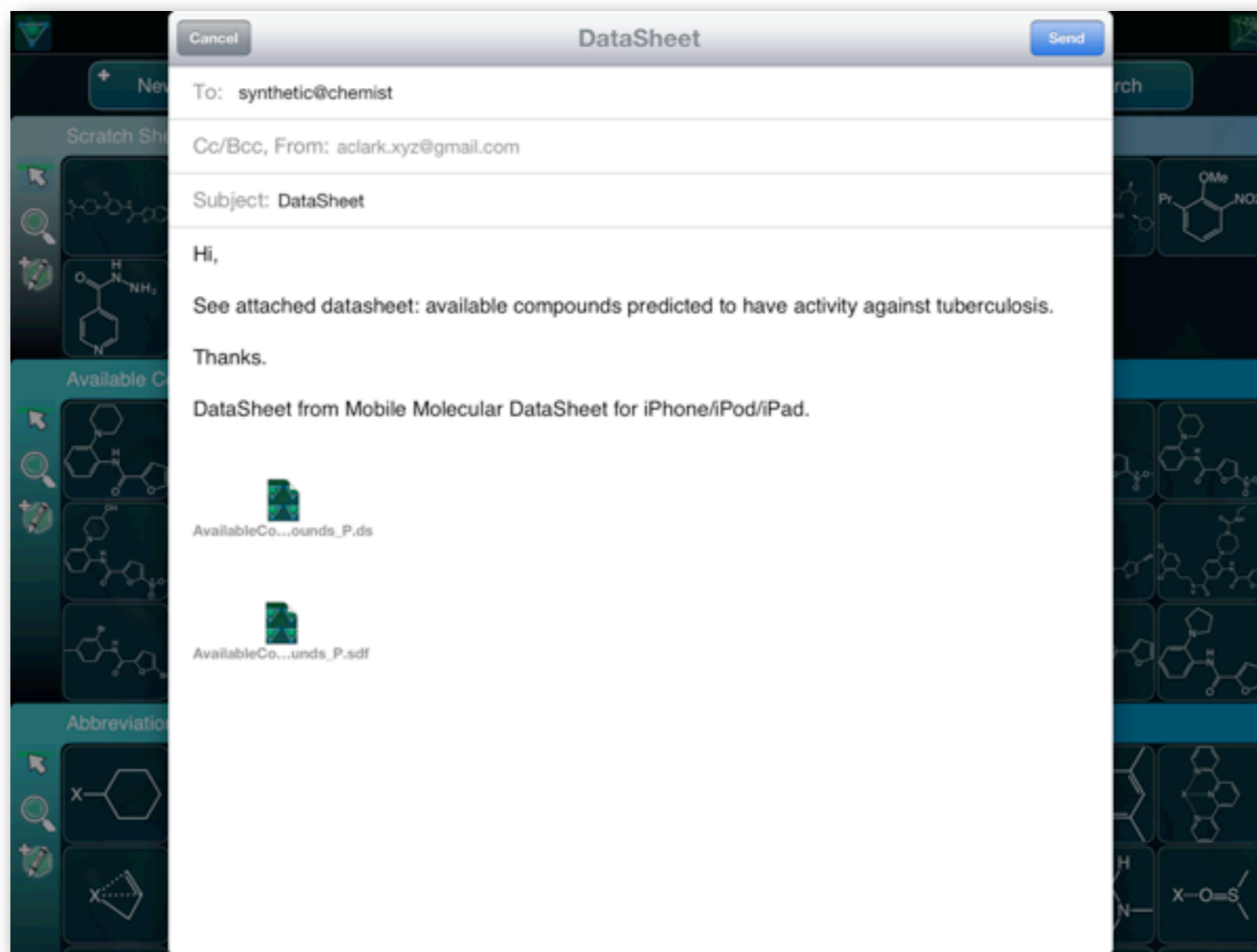


Sharing by Email

The screenshot displays the 'MobEmail DataSheet' interface. At the top, there are buttons for 'New Reactions', 'New Templates', 'Run Webservice', and 'Run Search'. Below these, the 'Default attachments' section lists '- DataSheet XML' and '- MDL SDF'. The 'Optional attachments' section includes checkboxes for 'Extended MDL SDF', 'Chemical Markup Language', 'SVG vector graphics (zip)', 'HTML with embedded SVG', 'Microsoft Word', and 'Microsoft Excel'. A 'Graphics Rendering' section offers radio button options: 'Black-on-White', 'White-on-Black', 'Colour-on-White' (which is selected), 'Colour-on-Black', and 'Printed Publication'. At the bottom, there are 'Compose' and 'Cancel' buttons. The background features a grid of chemical structures.

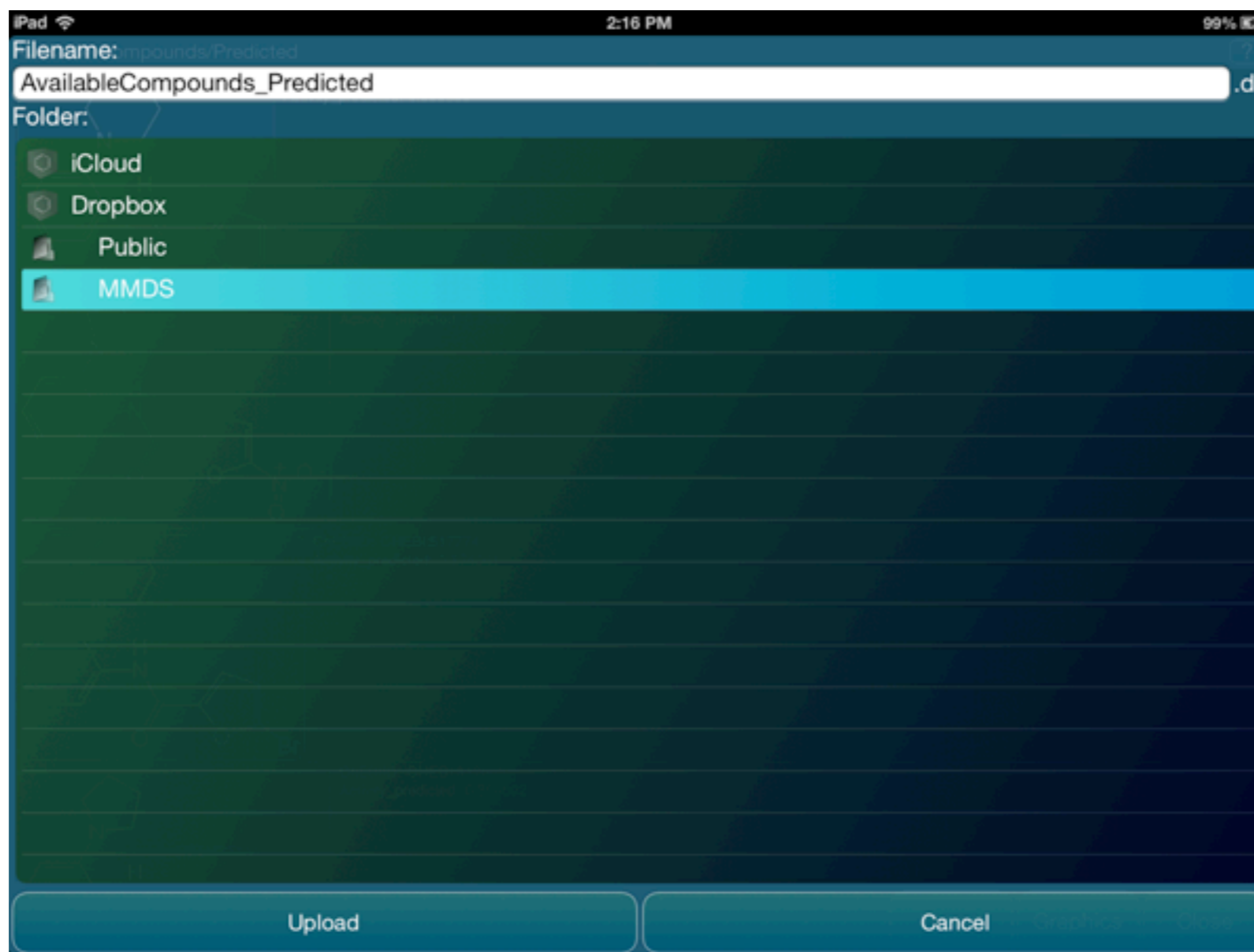
- Very easy to share using email attachments
- Recipient can open on any device, including mobile

Sharing by Email

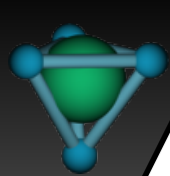


- Very easy to share using email attachments
- Recipient can open on any device, including mobile

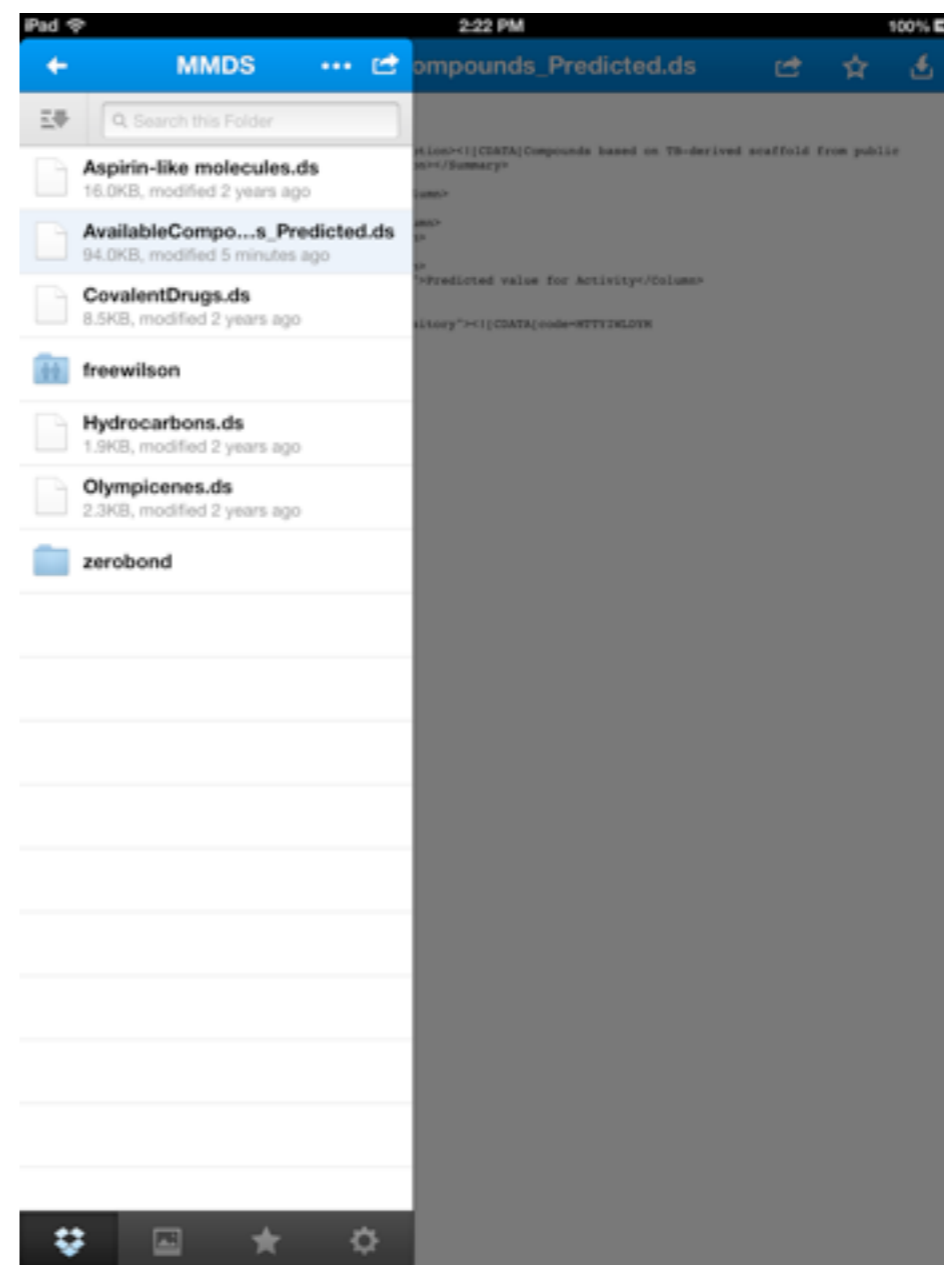
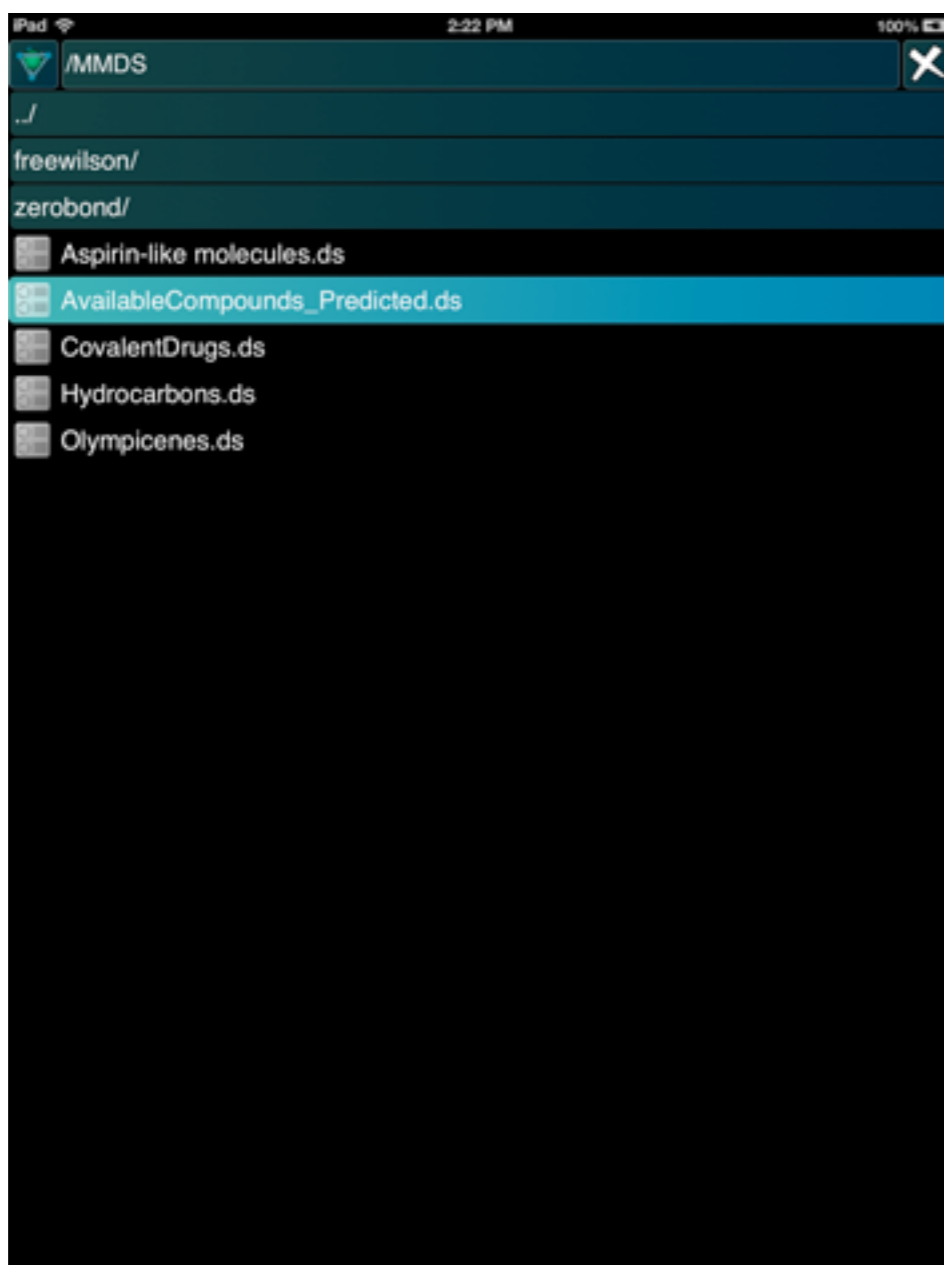
Sharing via Dropbox



- **MMDS** can interact with **MolSync** app
- DataSheets are uploaded & synchronised in **Dropbox**

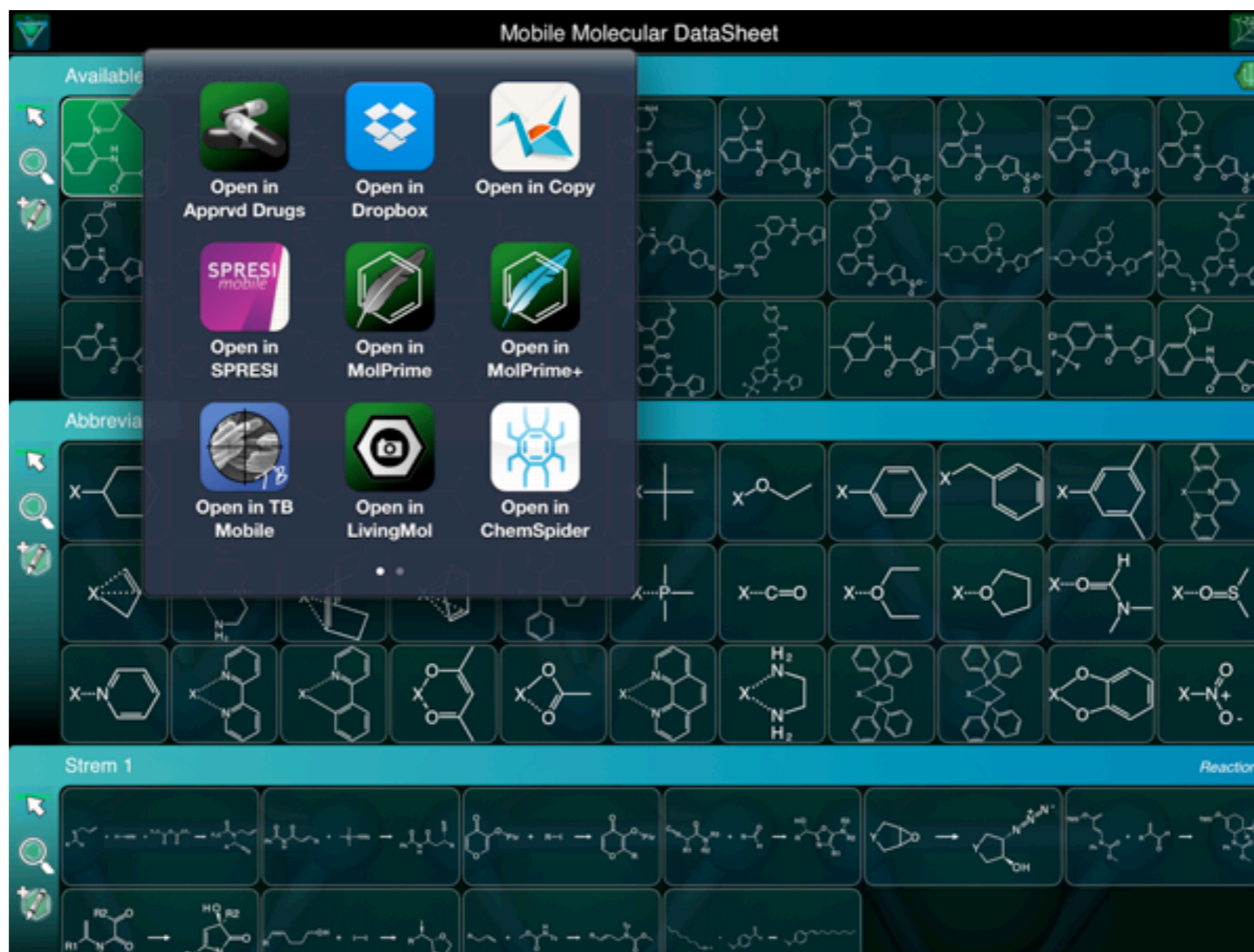


MolSync & Dropbox



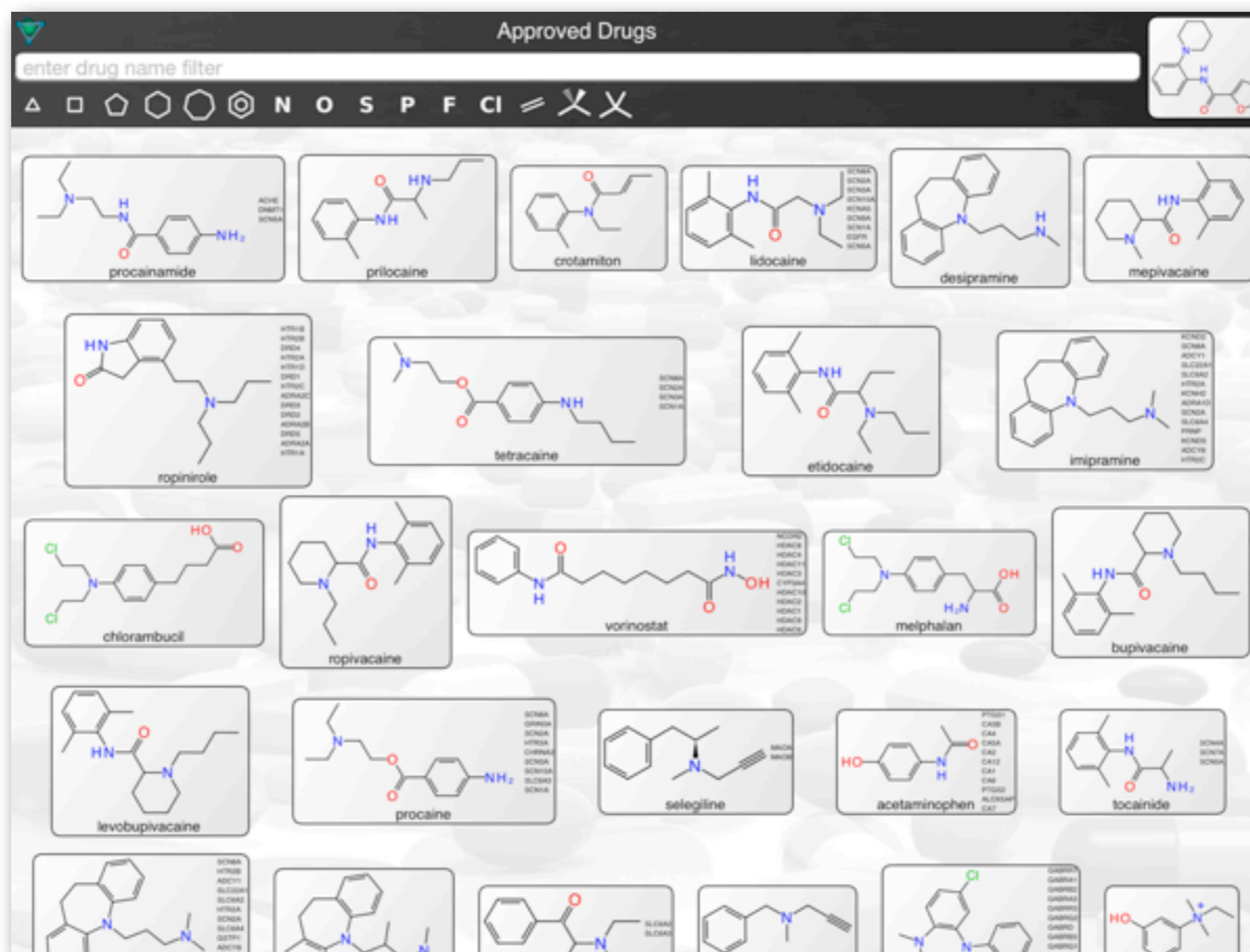
- **MolSync** app is a chemistry-aware file browser
- Can make use of **Dropbox**'s collaboration features

Lookup FDA Reference



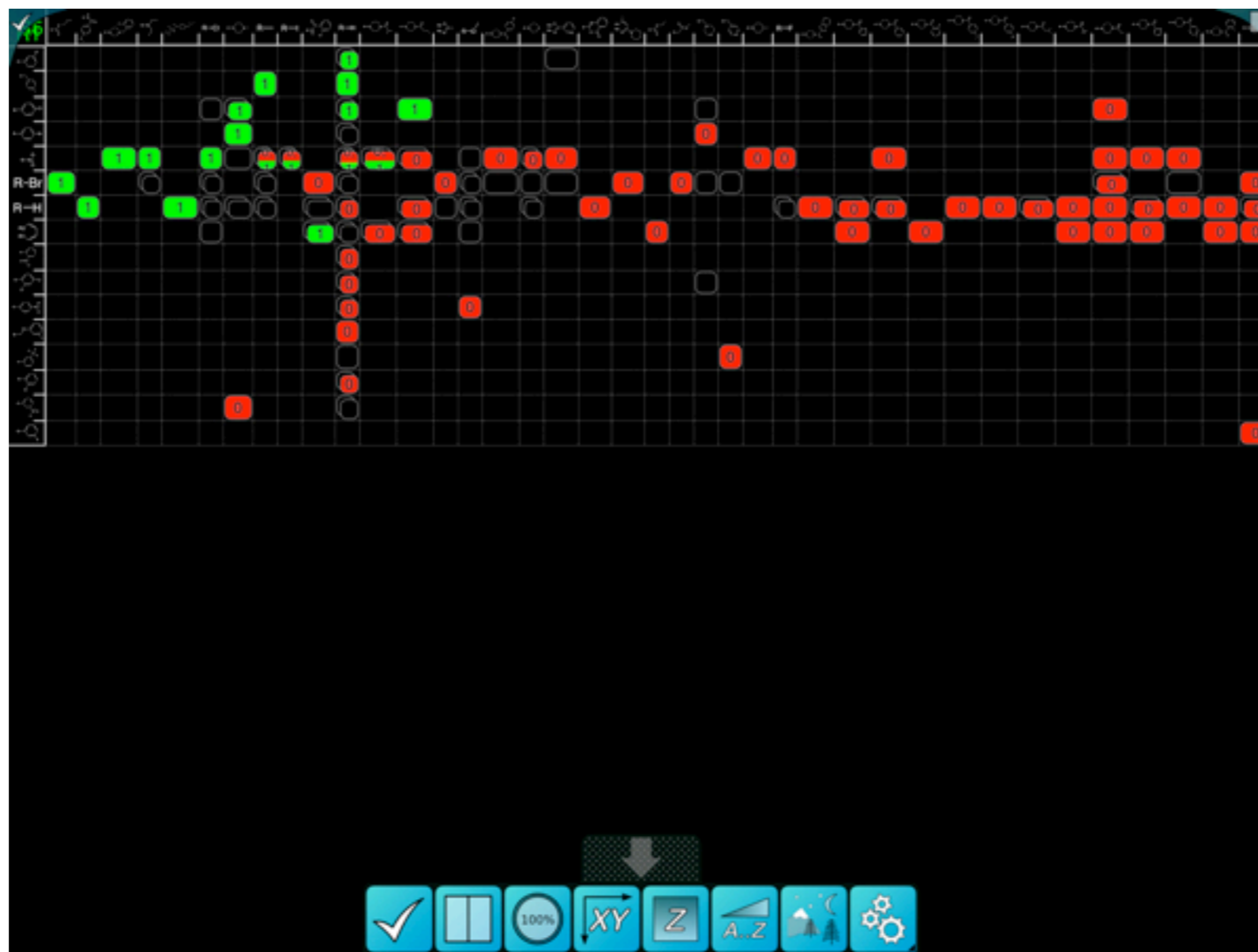
- *Open With* the **Approved Drugs** app

Approved Drugs

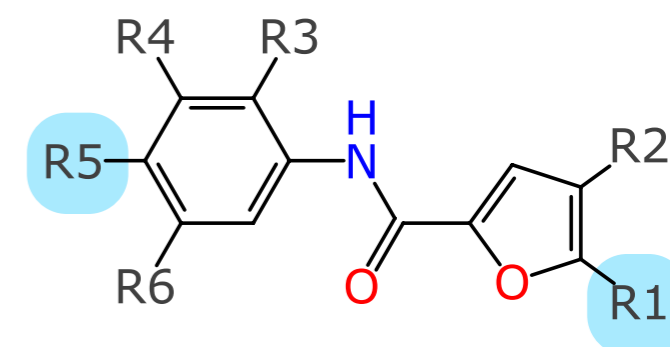


- 1300 structures for FDA-approved drugs
- Reference search by structural similarity

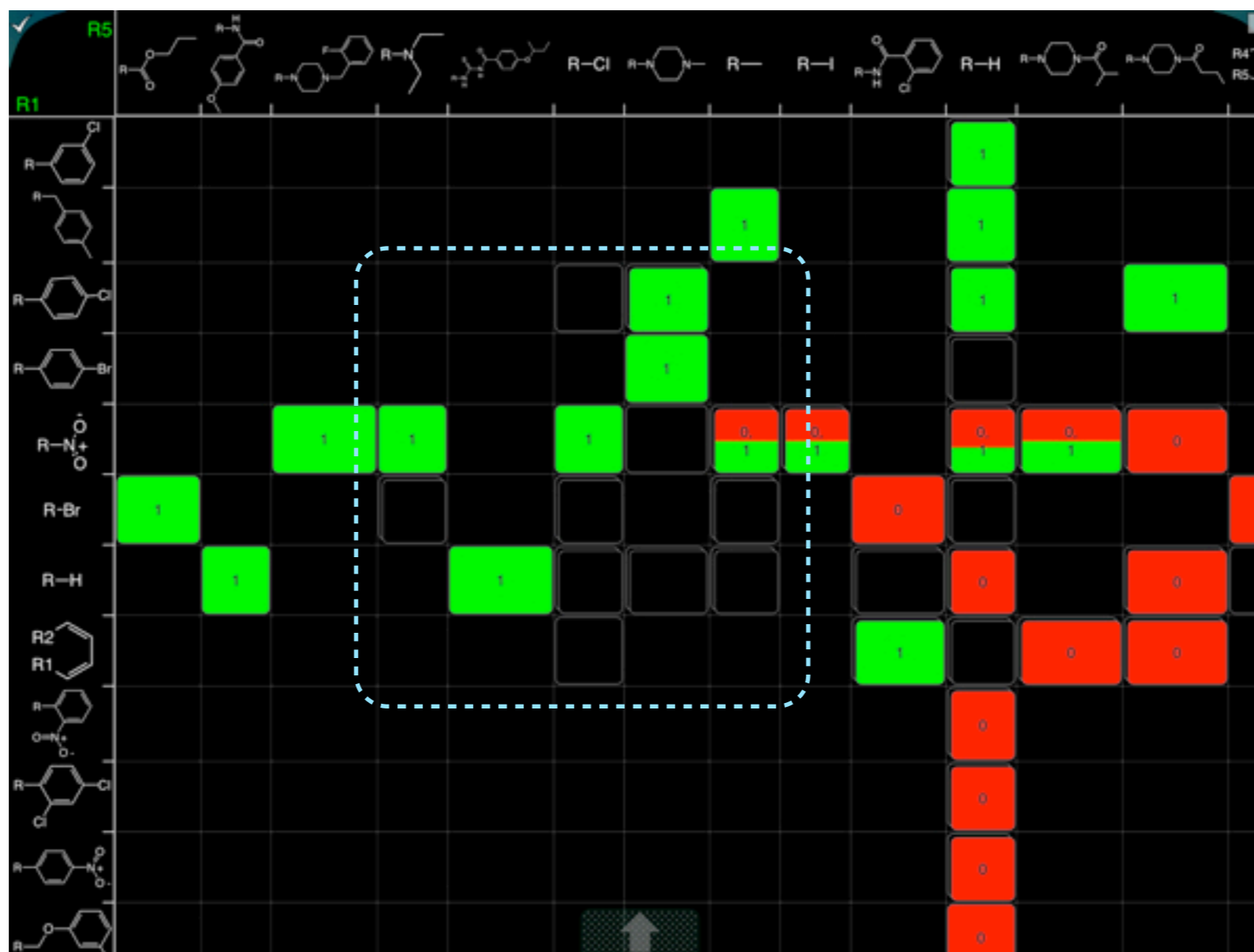
Matrix View



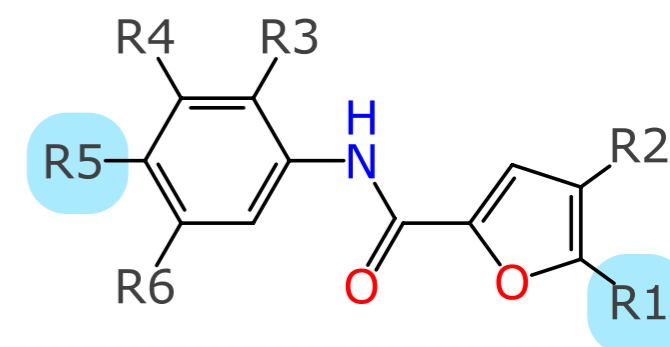
- Plot structures in a grid: **R1** vs **R5**

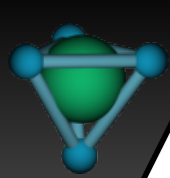


Matrix View

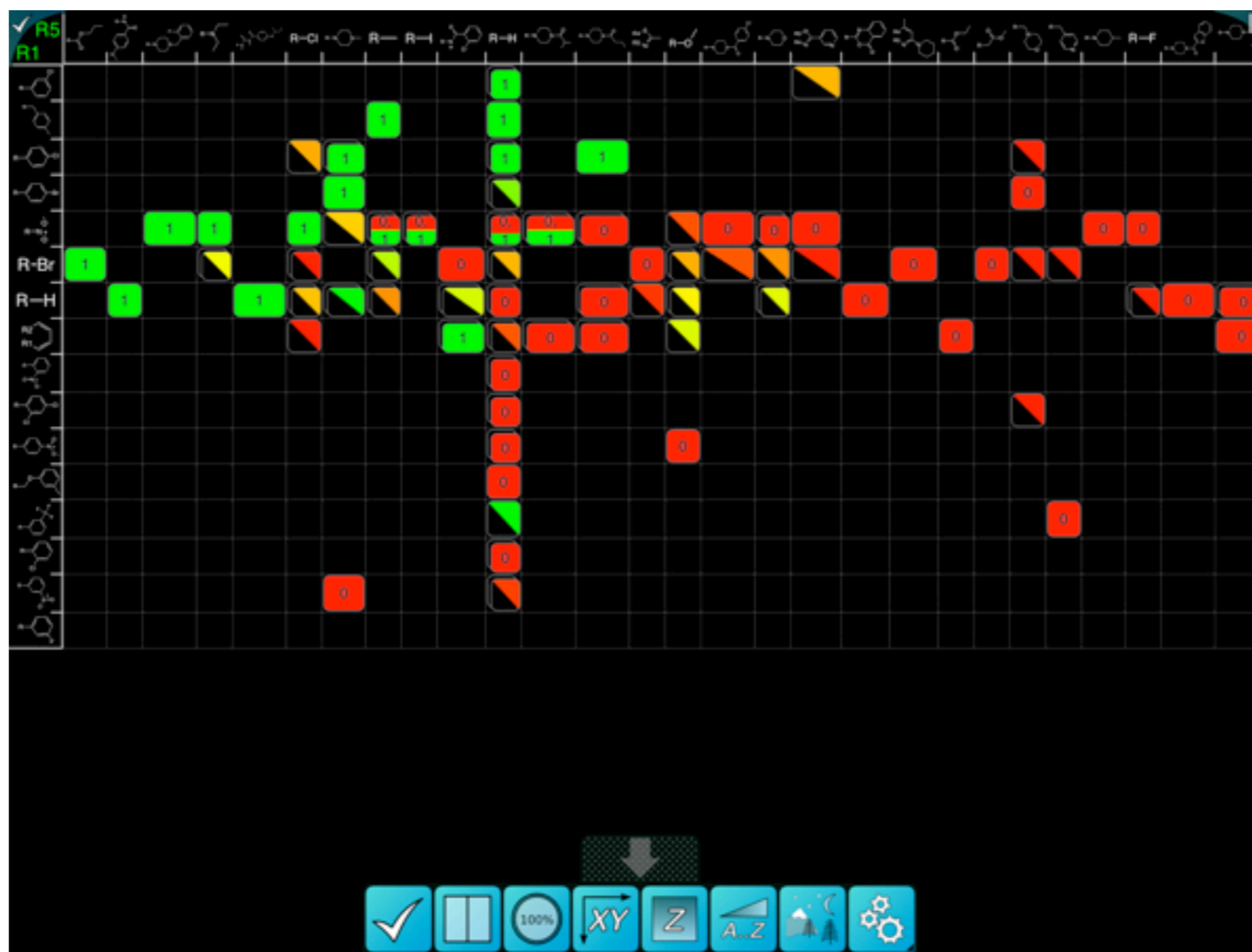


- Plot structures in a grid: **R1** vs **R5**



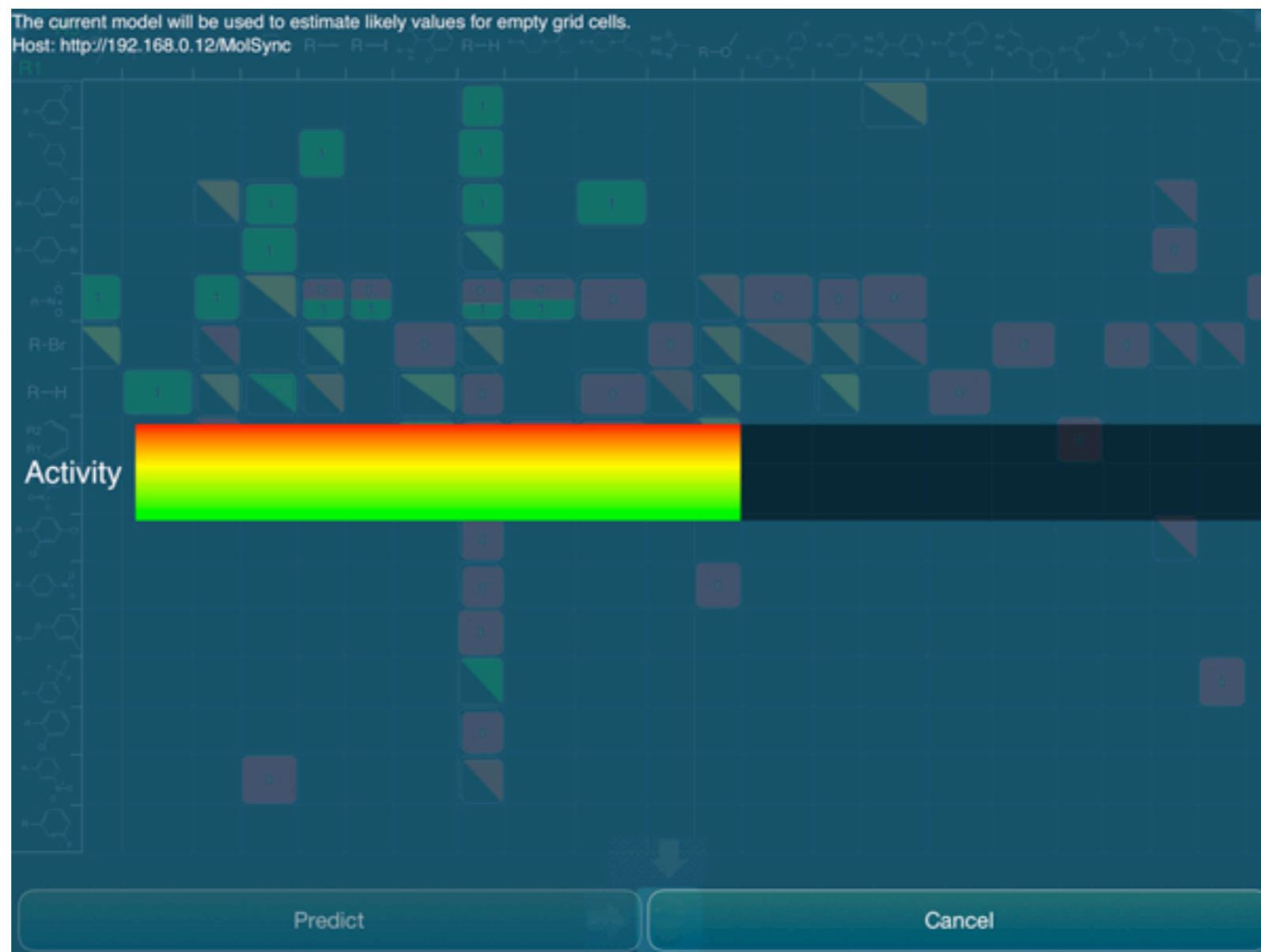


Matrix + Predictions



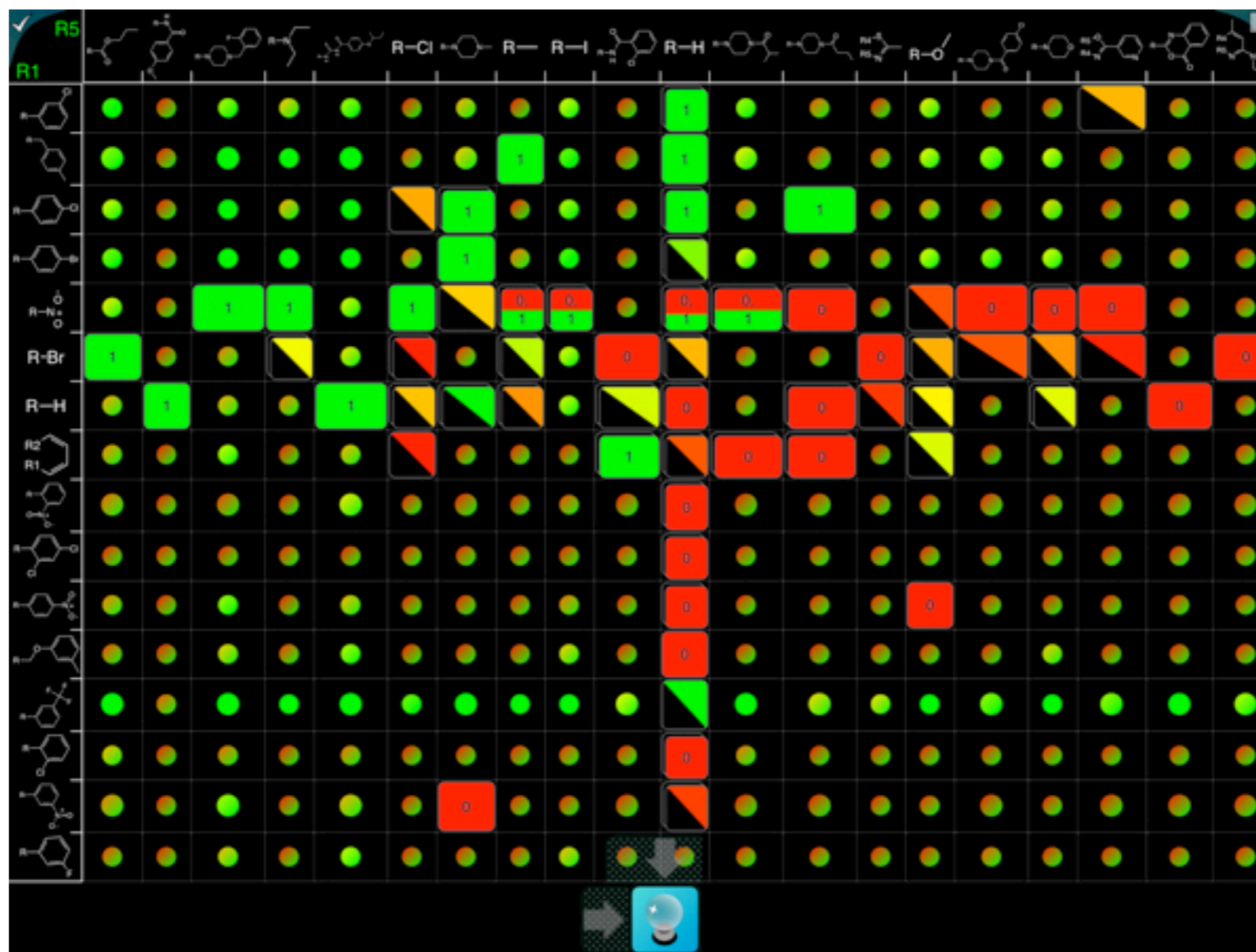
- Compounds with unknown activity predicted shown using wedge style

Matrix + Hypotheticals



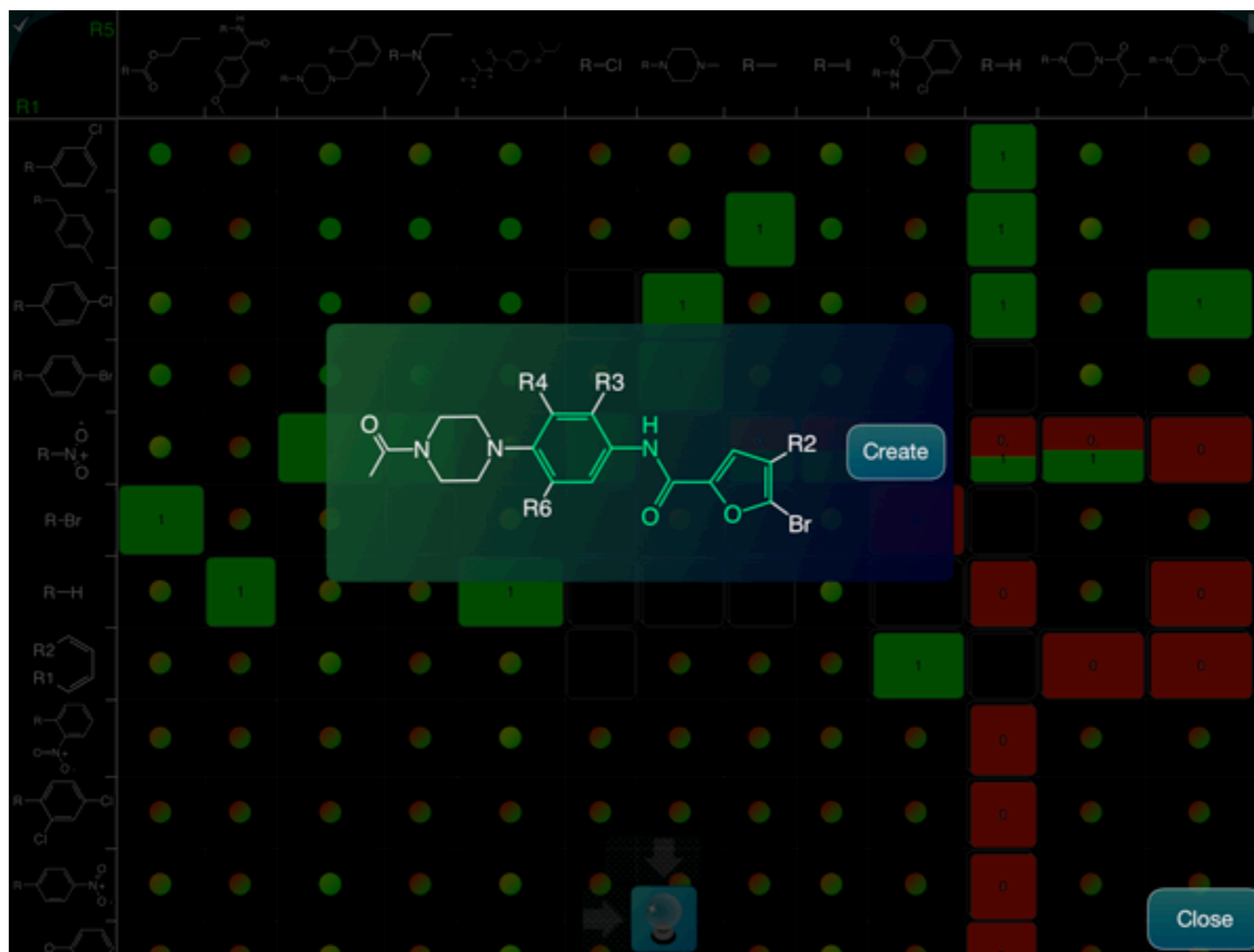
- Empty squares: propose compounds & predict

Matrix + Hypotheticals



- Empty squares: propose compounds & predict

Proposing Compounds



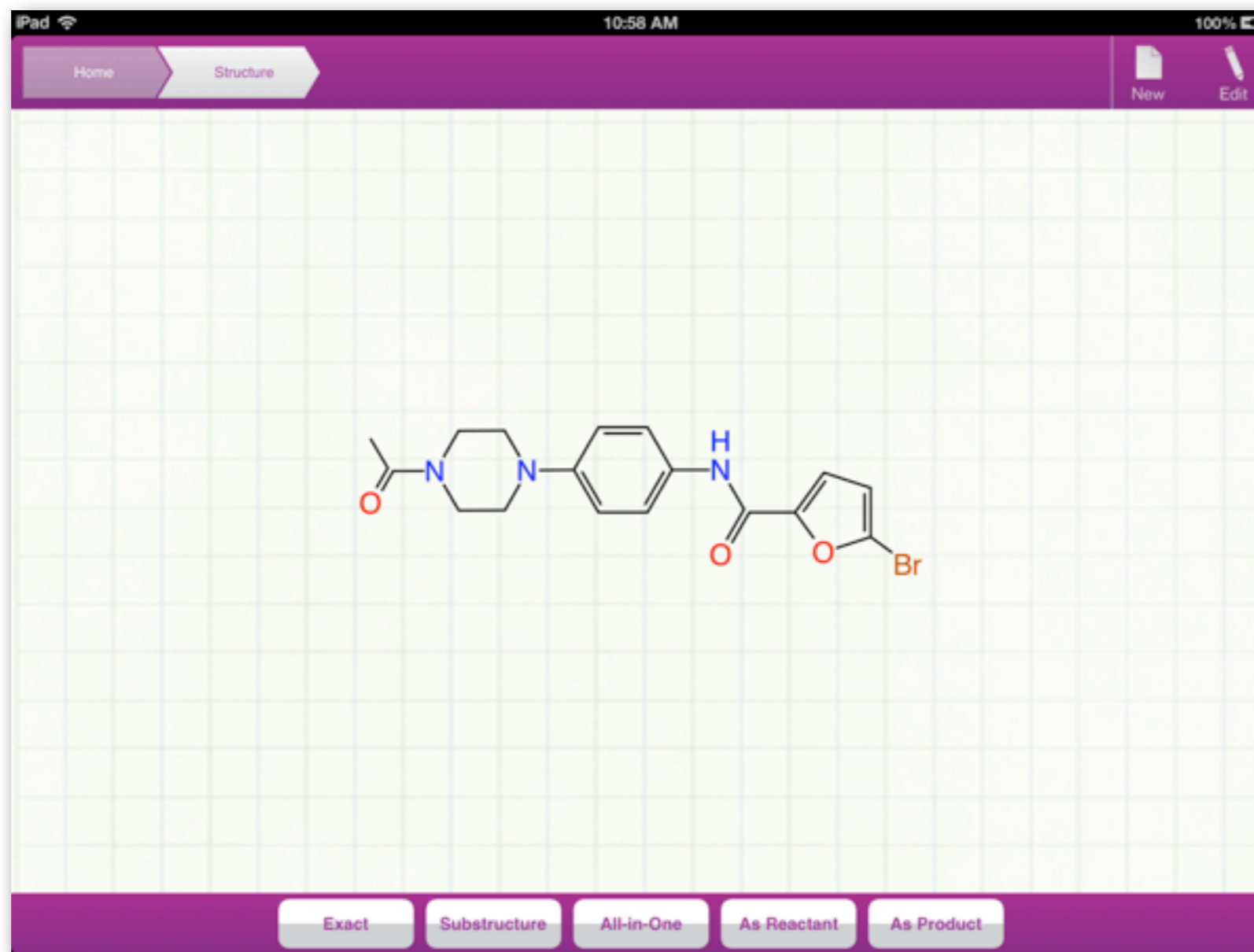
- Tap on a square: prompt to create partially defined entry, defining **R1** and **R5**

Hypothetical Compound

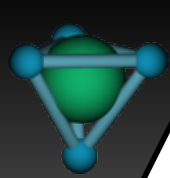
Scaffold	Molecule	R1	R2	R3	R4	R5	R6	Activity	CASRN
		R-Br	R-H		R-H	R-H		Green	(1)
		R-Br	R-H		R-H	R-H		Green	(1)
		R-H	R-H		R-H	R-H		Red	(0.0947 557)
		R-H	R-H		R-H	R-H		Yellow	(0.5129 41)
		R-N ⁺ ₃ ⁻	R-H	R-H	R-H	R-H	R-H	Red	(0)
		R-Br	?	?	?		?		

- New compound added: partial definition
- **R2**, **R3**, **R4** and **R6** unspecified

Lookup in SPRESI



- Use **SPRESImobile** app to find similar compounds: looking for a synthesis...



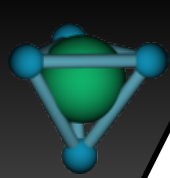
Similar Backbone

The screenshot displays a mobile application interface for molecular informatics. The top navigation bar includes 'Home', 'Structure', and 'Molecules'. The main content area shows a chemical structure of 5-bromo-N-(2-methylbenzyl)furan-2-carboxamide. The structure consists of a furan ring with a bromine atom at the 5-position, a carbonyl group at the 2-position, and an N-(2-methylbenzyl)amido group. The interface includes several informational panels:

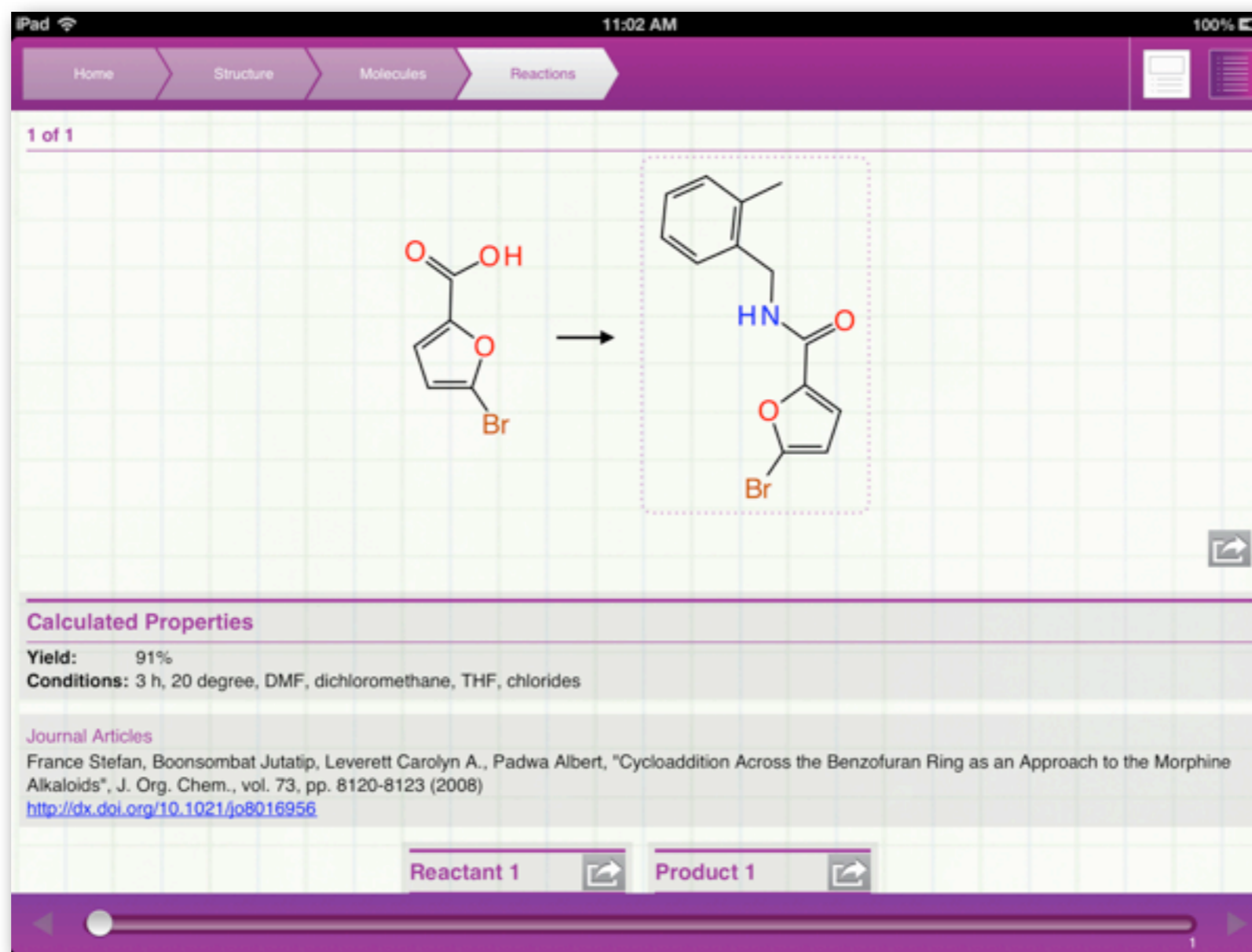
- Calculated Properties:**
 - Calculated Log P: 3.082
 - Rotatable Bonds: 5
 - H-Acceptors: 2
 - H-Donors: 1
- Name & Synonyms (2):**
 - 5-Bromofuran-2-carbonic acid, N-(2-methylbenzyl)amide
 - 5-Bromofuran-2-carbonsaeure, N-(2-methylbenzyl)amid
- Journal Articles (1):**
 - France Stefan, Boonsombat Jutapit, Leverett Carolyn A., Padwa Albert, "Cycloaddition Across the Benzofuran Ring as an Approach to the Morphine Alkaloids", J. Org. Chem., vol. 73, pp. 8120-8123 (2008) <http://dx.doi.org/10.1021/jo8016956>
- Information:**
 - Registration No.: 4199887-000
 - Molecular Formula: C₁₃H₁₂BrNO₂
 - Molecular Weight: 293.01

Buttons for 'As Reactant (1)' and 'As Product (1)' are visible below the structure. The bottom of the screen shows a navigation bar with a back arrow, a home button, and a forward arrow, along with a page number '100'.

- Found a compound with similar core, possible synthesis template: consult literature

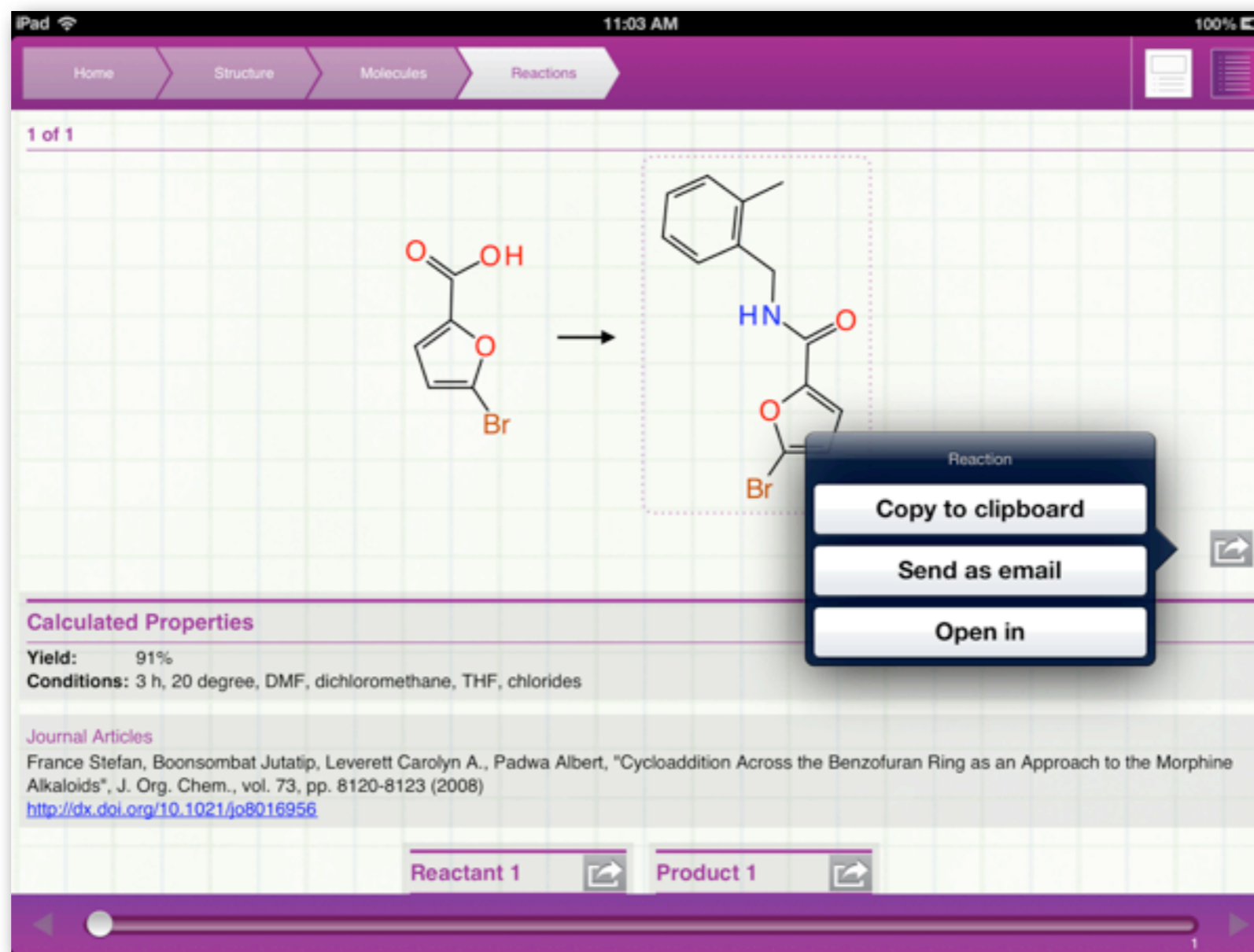


Find Reactions



- Lookup synthetic routes: amide condensation
- Open the reaction data in **Yield101**

Find Reactions



- Lookup synthetic routes: amide condensation
- Open the reaction data in **Yield101**

Find Reactions

1 of 1

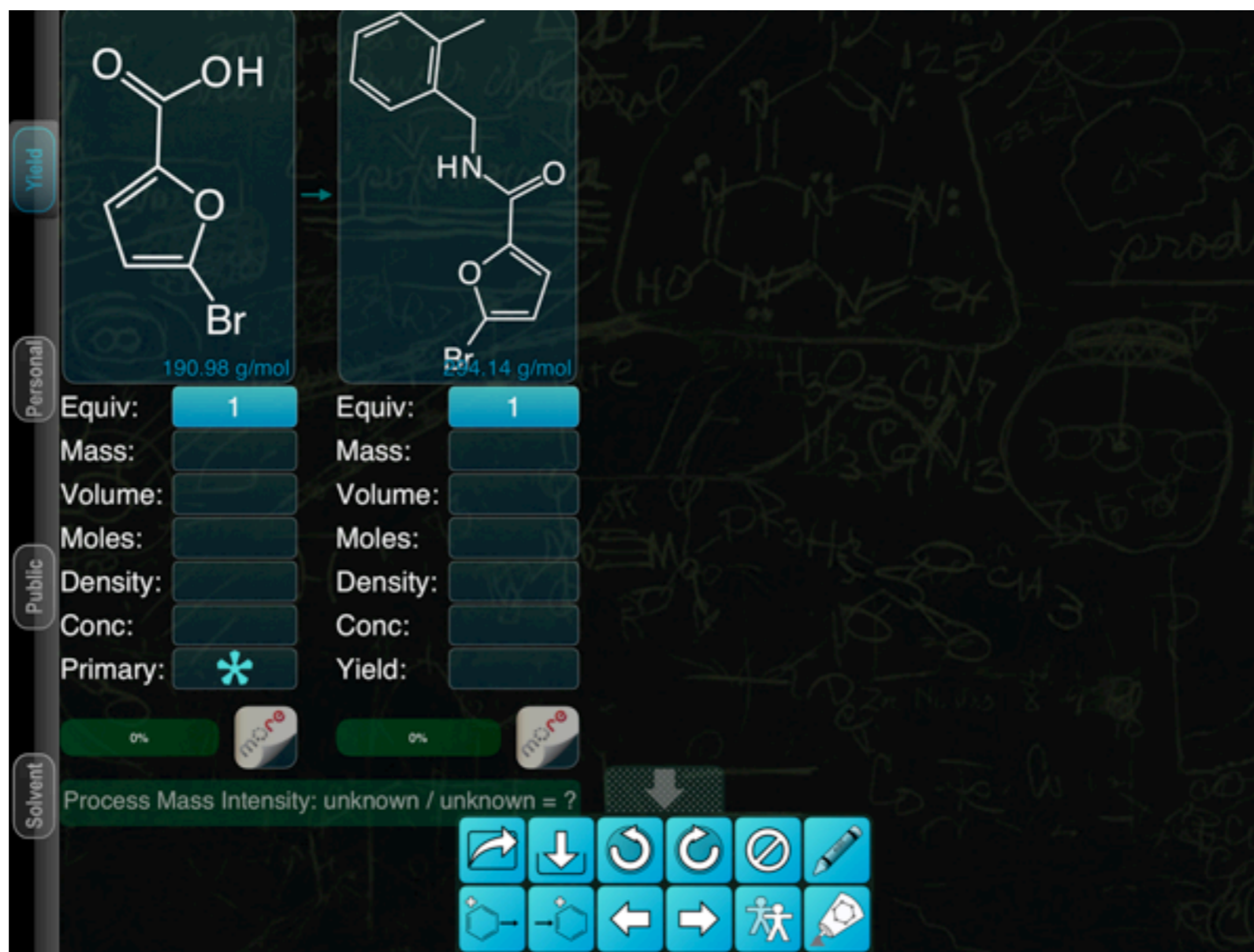
Yield: 91%
Conditions: 3 h, 20 degree, DMF, dichloromethane, THF, chlorides

Journal Articles
France Stefan, Boonsombat Jutatip, Leverett Carolyn A., Padwa Albert, "Cyclic Alkaloids", J. Org. Chem., vol. 73, pp. 8120-8123 (2008)
<http://dx.doi.org/10.1021/jo8016956>

Reactant 1 Product 1

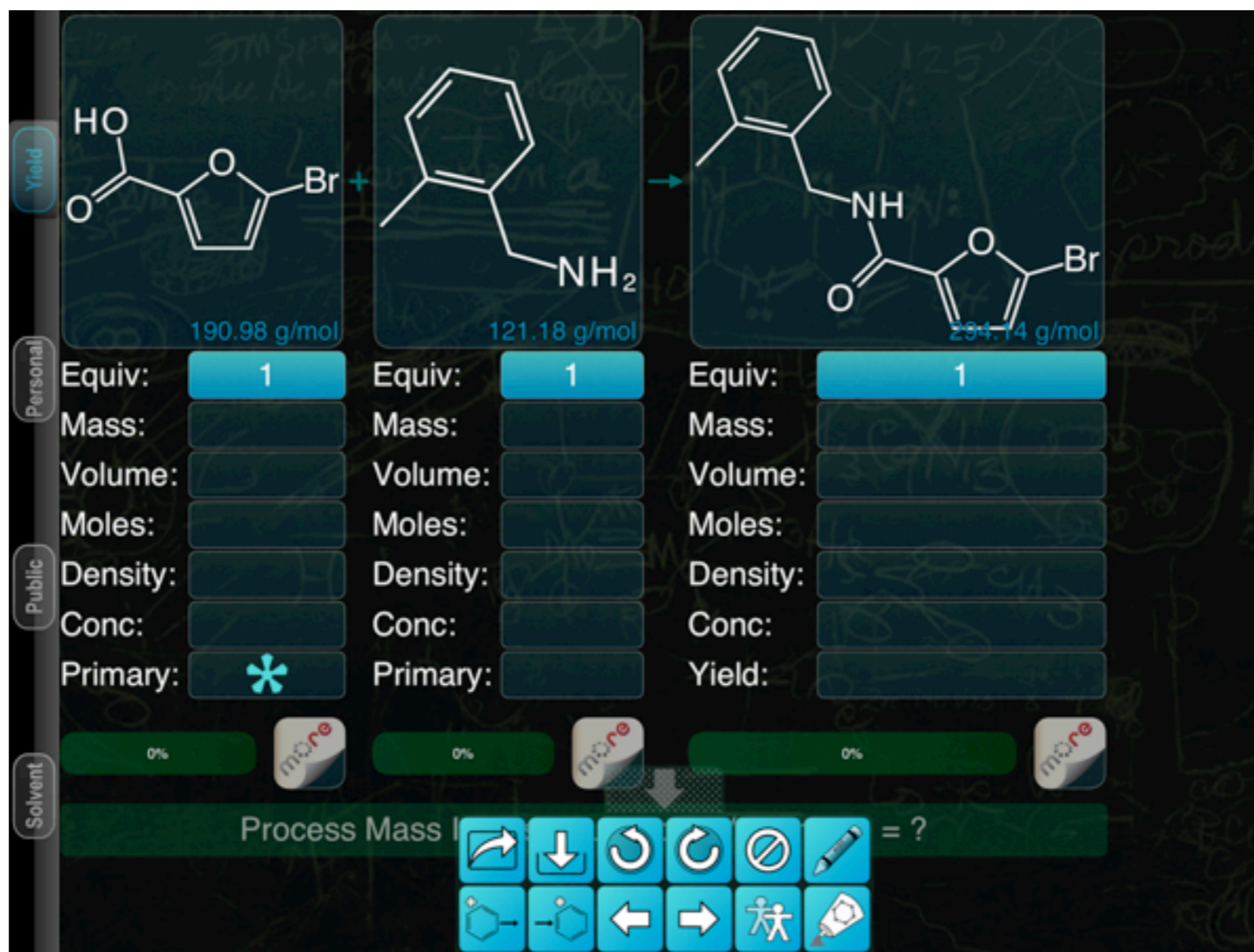
- Lookup synthetic routes: amide condensation
- Open the reaction data in **Yield101**

Yield101



- Prototype synthetic chemistry lab notebook
- Reaction components imported

Reaction Scheme



- Tidyup structures, add in reagent

Lookup Starting Material



The screenshot shows the Sigma-Aldrich website on an iPad. The page is for the product "5-Bromo-2-furoic acid 99%". The chemical structure is displayed as O=C(O)c1cc(Br)oc1. The page includes a navigation bar with "SIGMA-ALDRICH" and "200,000+ PRODUCTS", "SERVICES", and "SUPPORT". Below the product name, there are tabs for "Purchase", "Safety & Documentation", and "Protocols & Papers" (with a count of 2). The "Properties" section includes a table with the following data:

Property	Value
Related Categories	Building Blocks, C4 to C7, Chemical Synthesis, Furans, Halogenated Heterocycles, More...
assay	99%
mp	188-190 °C(lit.)

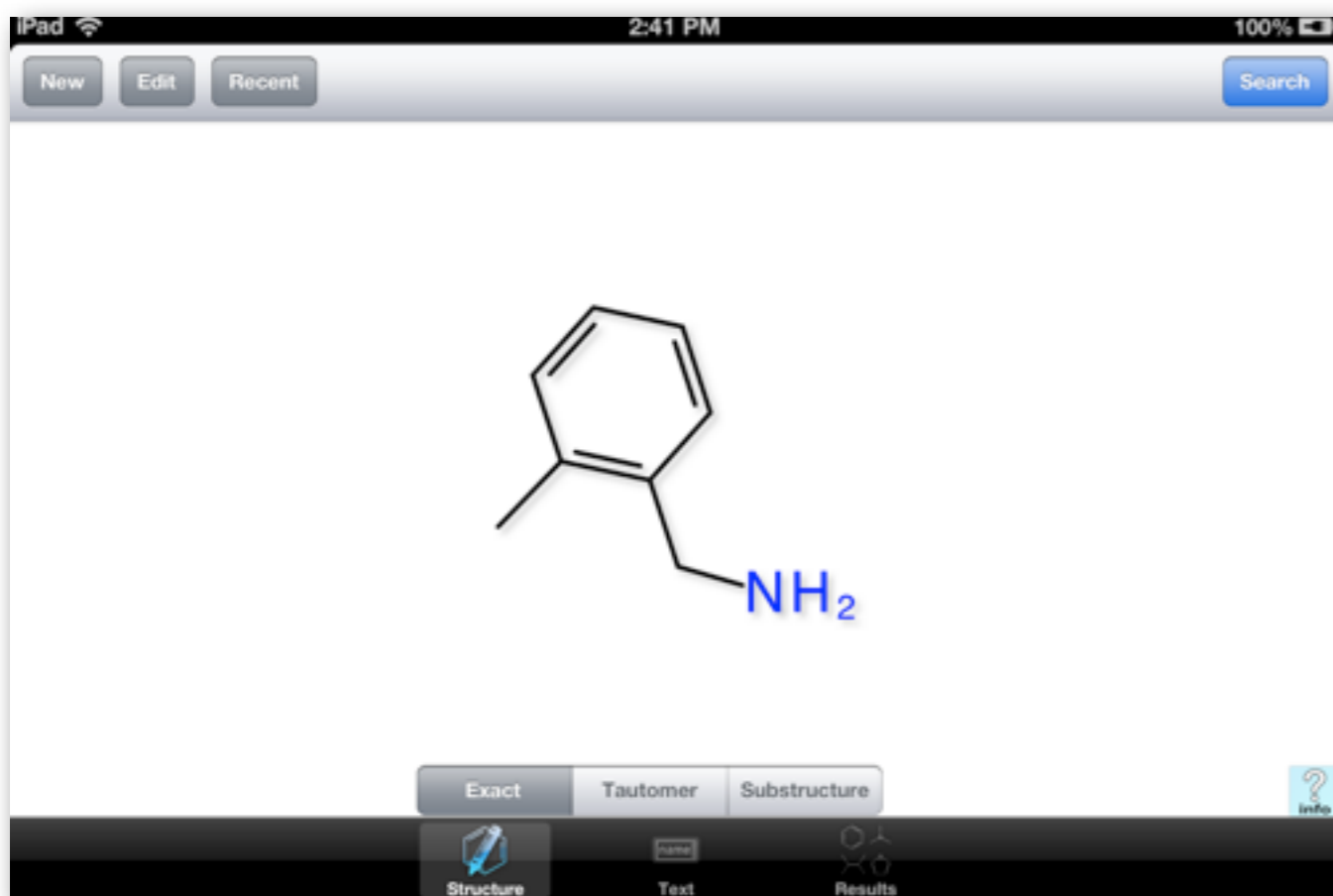
The "Price and Availability" section shows two options:

SKU-Pack Size	Availability	Price (CAD)	Quantity
B67406-25G	Ships on 09/23/13 - FROM	130.50	0
B67406-100G	Ships on 09/23/13 - FROM	473.00	0

At the bottom right, there is an "ADD TO CART" button.

- Use **Mobile Reagents** integration to find commercial source for starting material

Lookup in ChemSpider



www.chemspider.com/Chemical-Structure.6727.htm

o-Xylamine | C₉H₁₁N | ChemSpider

ChemSpider
Search and share chemistry

Search term: 6727 (Found by CSID)

o-Xylamine
ChemSpider ID: 6727
Molecular Formula: C₉H₁₁N
Average mass: 121.179604 Da
Monoisotopic mass: 121.089149 Da

Systematic name: 1-(2-Methylphenyl)methanamine

SMILES and InChI
Cite this record

Names and Identifiers
ChemSpider Searches

Properties

Experimental data | Predicted - ACD/Labs | Predicted - EPI Suite | Predicted - ChemAxon

Data supplied by datasources and users.

Experimental Physico-Chemical Properties

Experimental Boiling Point: 198-200 °C Alfa Aesar, 198-200 °C SynQuest [3730-1-28], 198-200 °C Alfa Aesar L05243, 198-200 °C Matrix Scientific 057130

Experimental Flash Point: 83 °C Alfa Aesar, 83 °C Alfa Aesar L05243

Experimental Gravity: 0.979 g/mL Alfa Aesar L05243, 0.98 g/l Matrix Scientific 057130

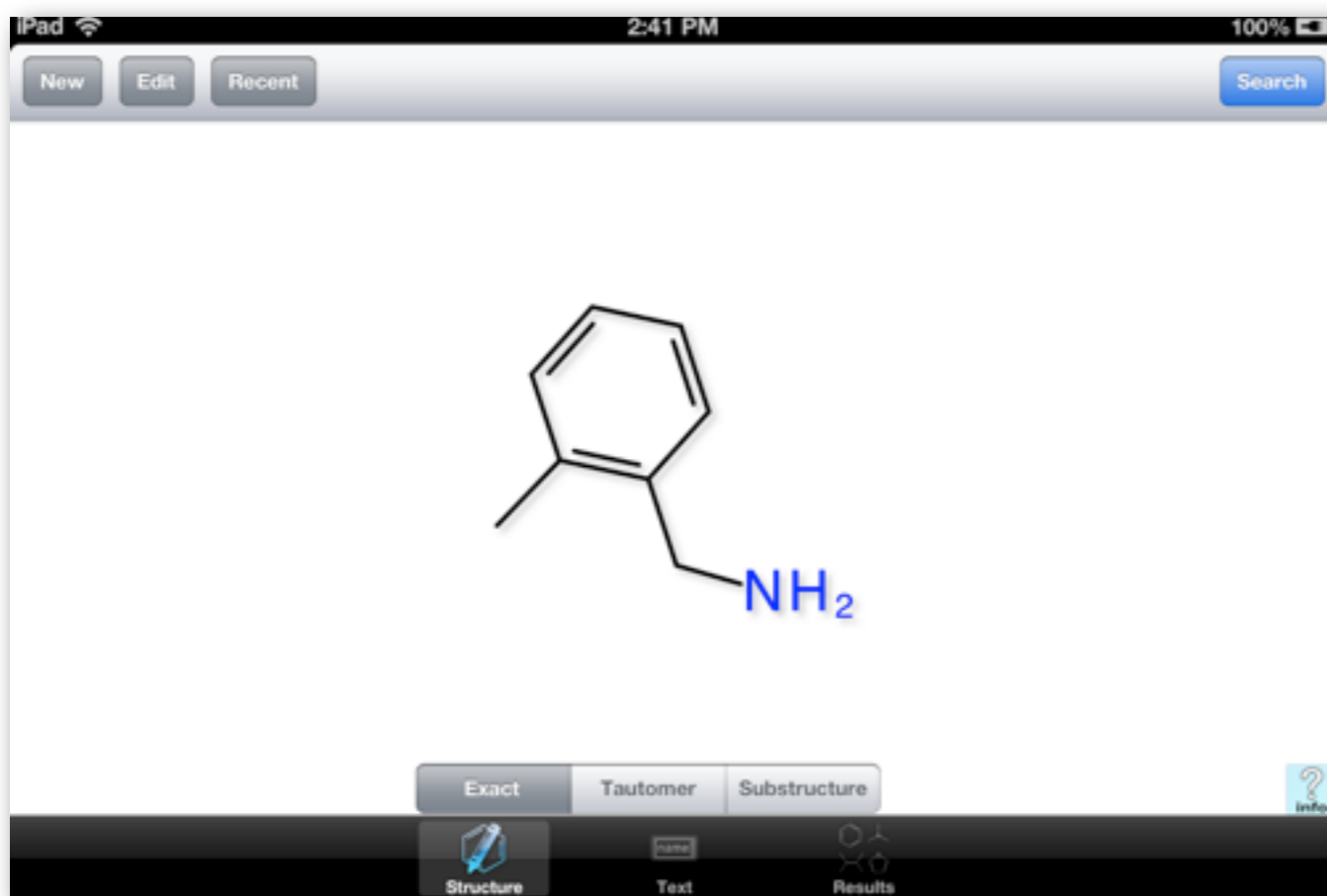
Experimental Refraction Index: 1.544 Alfa Aesar L05243

Miscellaneous

Safety: 20-26-36/37/39-45 Alfa Aesar L05243, 34 Alfa Aesar L05243, CORROSIVE Alfa Aesar L05243, DANGER: CORROSIVE, burns skin and eyes Alfa Aesar L05243, IRRITANT Matrix Scientific 057130

- Paste into **ChemSpider Mobile** app to do structure search
- Open matching **ChemSpider** compound information page

Lookup in ChemSpider



Search term: 6727 (Found by CSID)

o-Xylamine
 ChemSpider ID: 6727
 Molecular Formula: C₉H₁₁N
 Average mass: 121.179604 Da
 Monoisotopic mass: 121.089149 Da

Systematic name: 1-(2-Methylphenyl)methanamine
 SMILES and InChI
 Cite this record

Names and Identifiers
 ChemSpider Searches

Properties

Experimental data | Predicted - ACD/Labs | Predicted - EPI Suite | Predicted - ChemAxon

Data supplied by datasources and users.

- Experimental Physico-Chemical Properties
 - Experimental Boiling Point: 198-200 °C Alfa Aesar, 198-200 °C SynQuest [3730-1-28], 198-200 °C Alfa Aesar L05243, 198-200 °C Matrix Scientific 057130
 - Experimental Flash Point: 83 °C Alfa Aesar, 83 °C Alfa Aesar L05243
 - Experimental Gravity: 0.979 g/mL Alfa Aesar L05243, 0.98 g/l Matrix Scientific 057130
 - Experimental Refraction Index: 1.544 Alfa Aesar L05243
- Miscellaneous
 - Safety: 20-26-36/37/39-45 Alfa Aesar L05243, 34 Alfa Aesar L05243, CORROSIVE Alfa Aesar L05243, DANGER: CORROSIVE, burns skin and eyes Alfa Aesar L05243, IRRITANT Matrix Scientific 057130

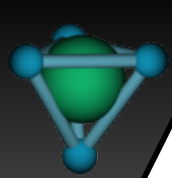
- Paste into **ChemSpider Mobile** app to do structure search
- Open matching **ChemSpider** compound information page

Quantitative Details

The screenshot displays a chemical synthesis software interface with the following data:

Parameter	Reactant 1 (5-bromo-2-hydroxyfuran)	Reactant 2 (N-(2-methylphenyl)ethan-1-amine)	Product (N-(2-bromo-5-(2-methylphenyl)oxy)furan-2-carboxamide)
Chemical Structure	<chem>O=C(O)c1oc(Br)cc1</chem>	<chem>CC1=CC=C(C=C1)CCN</chem>	<chem>CC1=CC=C(C=C1)CCNC(=O)c2oc(Br)cc2</chem>
Molar Mass (g/mol)	190.98	121.18	294.14
Equiv.	1	1	1
Mass (g)	10	6.34516	
Volume (mL)		6.48127	
Moles (mol)	0.0523616	0.0523616	
Density (g/mL)		0.979	
Conc.			
Primary	*		
Progress	61%	39%	0%
Process Mass Intensity	16.345 g / unknown = ?		

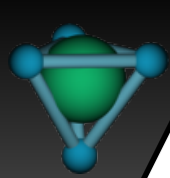
- Enter starting quantity, density: preliminary estimate of amounts used in synthesis



Public Sharing

The screenshot displays the Molsync interface for public sharing. At the top, it says "Upload" and "Prior to sharing on the web, your content needs to be uploaded to <http://molsync.com/MolSync>". Below this is a chemical reaction scheme showing the synthesis of a brominated benzamide derivative from a brominated furan-2-carboxylic acid and a benzylamine derivative. An "Upload" button is positioned below the reaction. The "Link" section provides a public link: <http://molsync.com/share?ds=103>. The interface also shows various input fields for "Moles", "Density", "Conc", "Primary", and "Yield", along with a "Process Mass Intensity" field. A "Close" button is located at the bottom right.

- Upload scheme to *molsync.com*
- Generates publicly sharable link



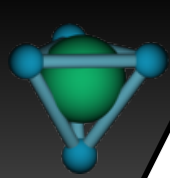
Shared Content

- Stores chemical data in its native form:
 - molecules
 - reaction schemes
 - datasheets
- Can share, tweet, etc.
- Download data in many different forms...

Reaction

Row#1 Download		
		Reactant Equivalents: 1 Primary reactant MF: C ₅ H ₃ BrO ₃ MW: 190.980 Mass: 10 g Moles: 0.0523616 mol
		Reactant Equivalents: 1 MF: C ₉ H ₁₁ N MW: 121.180 Density: 0.979 g/mL Mass: 6.34516 g Volume: 6.48127 mL Moles: 0.0523616 mol
		Product Equivalents: 1 MF: C ₁₃ H ₁₂ BrNO ₂ MW: 294.144 Mass: 15.4019 g Moles: 0.0523616 mol Yield: 100 %

Provided by [Molecular Materials Informatics, Inc](#)
Visit the [iTunes AppStore](#) for the mobile MolSync app



Shared Content

- Stores chemical data in its native form:
 - molecules
 - reaction schemes
 - datasheets
- Can share, tweet, etc.
- Download data in many different forms...

Download DataSheet

Choose Format

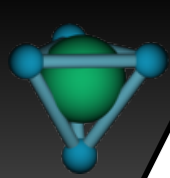
- DataSheet XML
- MDL RDF (reactions + data)
- MDL RXN (single reaction)
- PNG image (raster)
- SVG picture (vector)
- Encapsulated PostScript (vector)
- ZIP (multiple PNG files)
- ZIP (multiple SVG files)
- HTML with embedded SVG
- Microsoft Word
- Microsoft Excel

Graphic Options

Sizing:

Angstroms-to-Points:

Rendering:



Numerous Formats

- Sharing site stores pure data, creates preview graphics dynamically... and interconversions

Molecules

- SketchEI
- MDL MOL
- CML

Reactions

- DataSheet XML
- MDL RXN
- MDL RDF

Collections

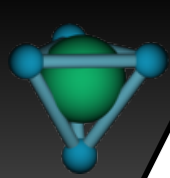
- DataSheet XML
- MDL SDfile

Raster Graphics

- PNG
- PNG ZIP

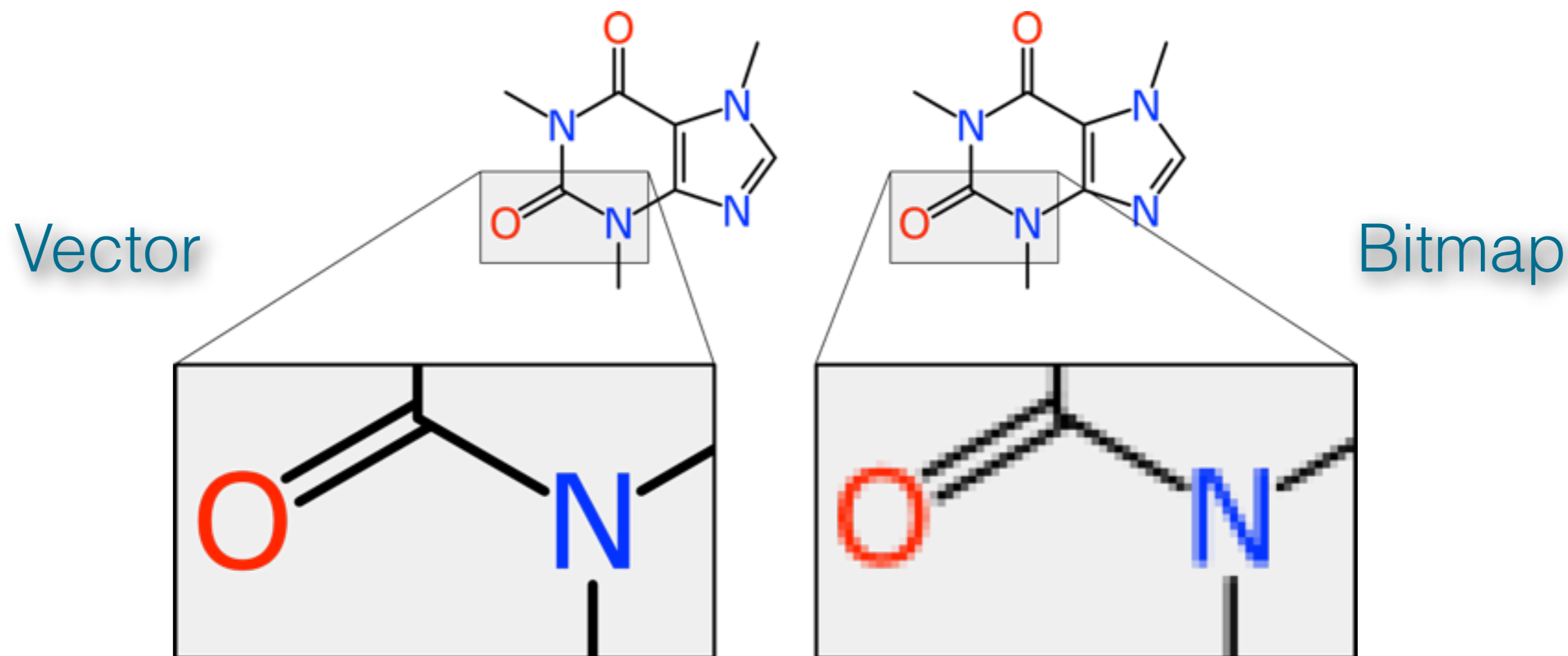
Vector Graphics

- SVG
- SVG ZIP
- HTML SVG
- EPS
- MS Word
- MS Excel



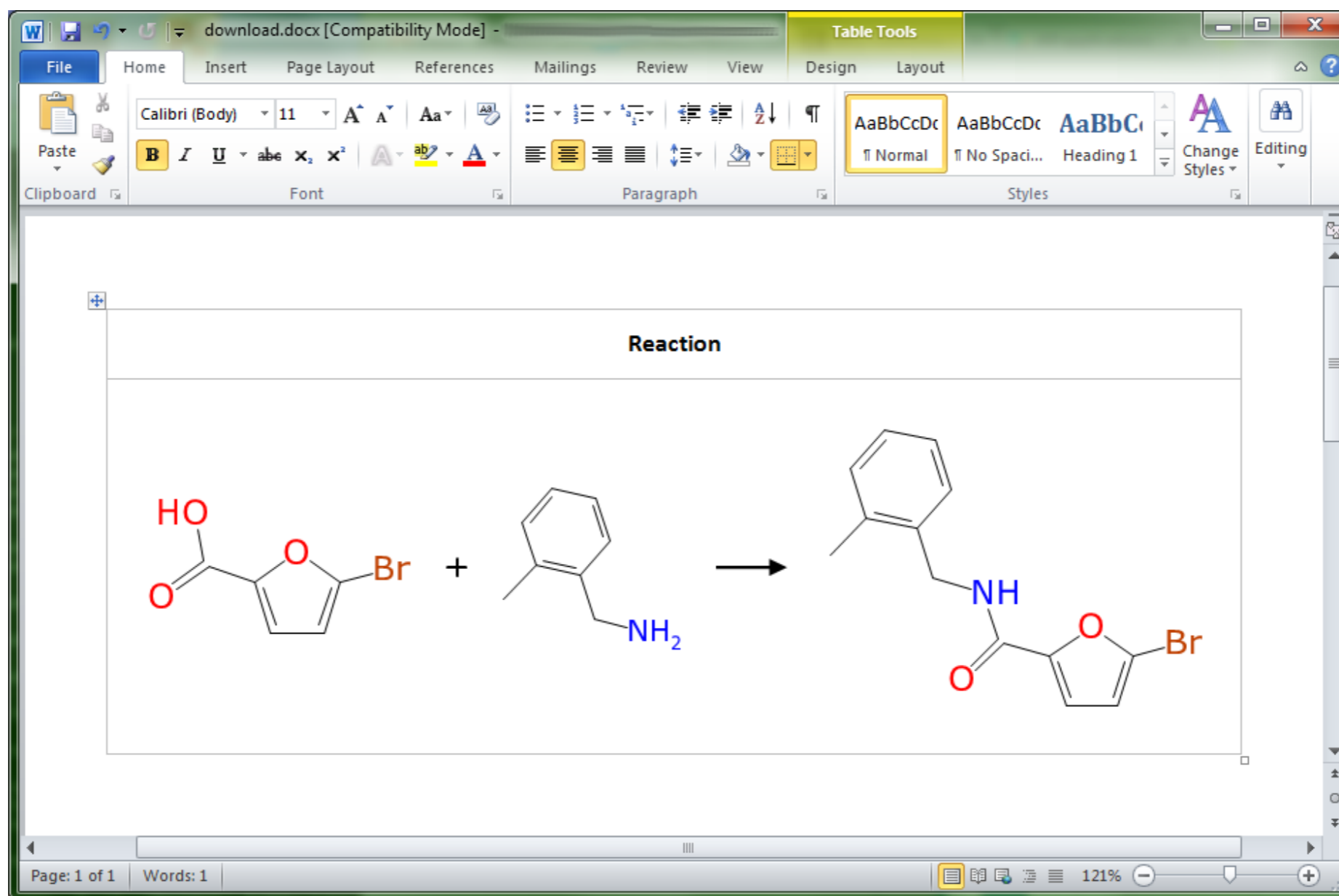
Vector Graphics

- Quality is perfect at all resolutions: screen, projector, web, printer, PDF...



Microsoft Office Documents

- Independent implementation of **OOXML** spec
- Embedded graphics use vector **DrawingML**



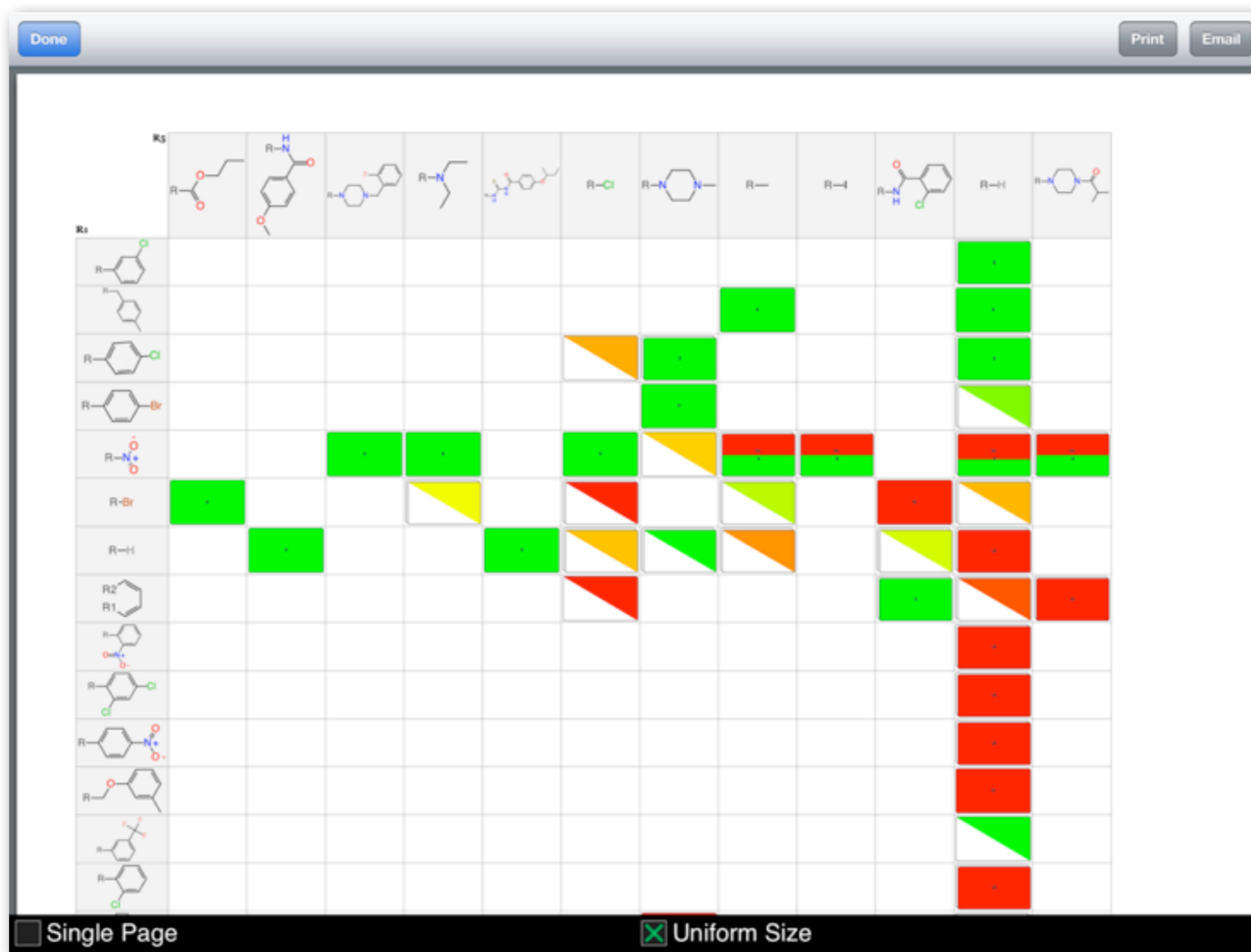
SAR Table Documents

Anti-TB Scaffold: Known

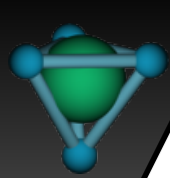
Scaffold	R1	R2	R3	R4	R5	R6	Molecule	Activity
	R-Br	R-H	R-H	R-H		R-H		0
			R-H	R-H		R-H		0
	R-H	R-H	R-H	R-H		R-H		0
	R-H	R-H	R-H	R-H		R-H		0
			R-H	R-H		R-H		0
			R-H	R-Cl		R-H		0
			R-H	R-H		R-H		0

- Scaffold/substituent tables for manuscripts

Matrix PDF



- Multi-page layout: print directly or email PDF



Tweeting

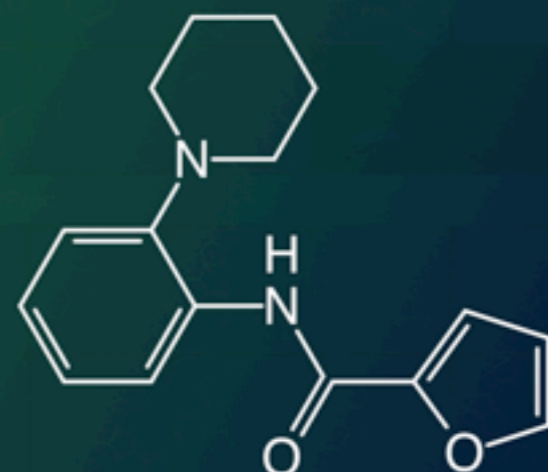
- Can tweet directly from apps like **MMDS**
- First upload data to **molsync.com**
- Then emit tweet with link
- Observe list of predefined hashtags:

#tuberculosis

Twitter Account
Confirm the account from which to tweet:

aclarkxyz
 oddtinfo
 ODDTTest

Content
Prior to tweeting, this data will be uploaded to molsync.com, and made publicly available:



DataSheet with 81 rows and 7 columns

Text

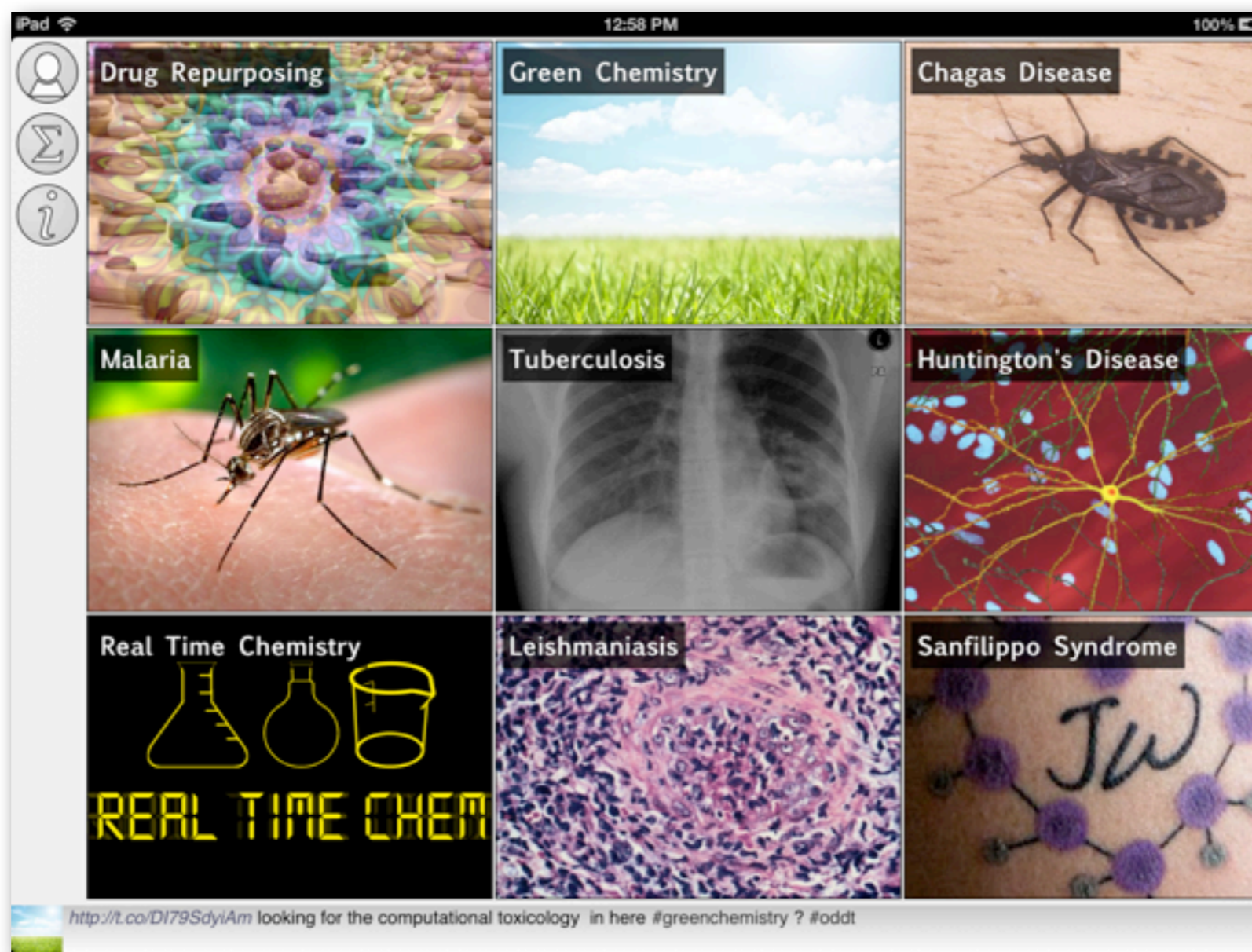
Proposed leads from scaffold analysis

ODDT
Selecting one of the following topics will add hashtags for Open Drug Discovery Teams project:

None
 #realtimechem
 #tuberculosis
 #malaria
 #hivaids
 #huntingtons
 #greenchemistry
 #acsgci

Cancel Tweet

Open Drug Discovery Teams



- Curation of open data, e.g. Twitter & RSS feeds
- Rare & neglected diseases, precompetitive areas

Harvested Tweet

The screenshot shows an iPad interface with a Twitter feed for the topic "Tuberculosis". The feed includes:

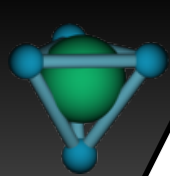
- A news article titled "Tuberculosis fears have health officials fielding calls" with a URL: <http://www.cbc.ca/news/canada/new-brunswick/tuber...>
- A tweet from @aclarkxyz: "(link) #tuberculosis #ODDT Proposed leads from scaffold analysis" with a URL: <http://molsync.com/share/?ds=105>. The tweet includes a table of chemical structures:

Molecule	CASRN	PubChemID	ChEMBL	CSID	Vendors	Activity_predicted

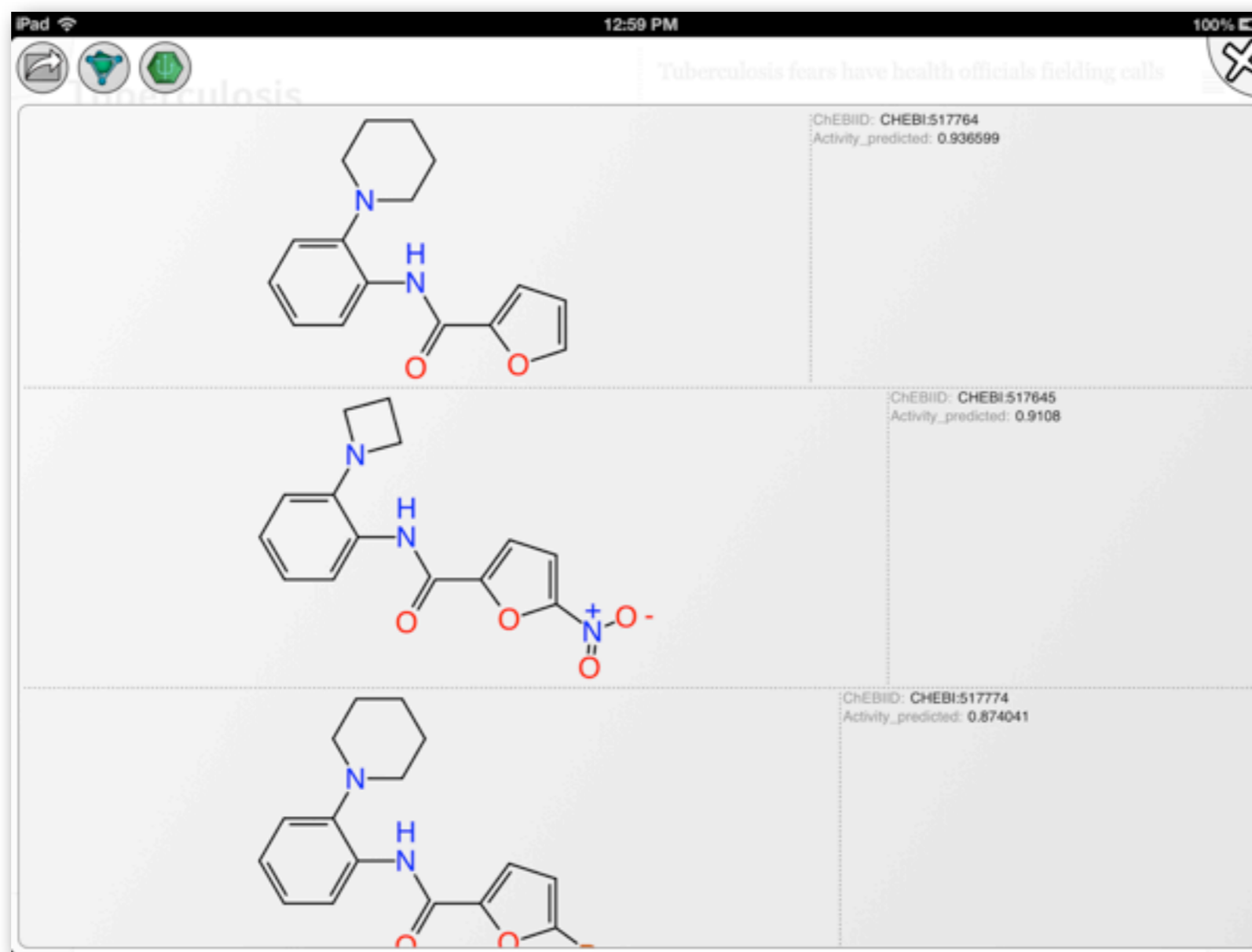
- A tweet from @aerasglobaltb: "The fight against #tuberculosis is not over (link) via @thehill @MarioDB @RepSires" with a URL: <http://thehill.com/blogs/congress-blog/healthcare/324...>
- A tweet titled "How long can mycobacteria tuberculosis culture survive in 4 ..." with a URL: http://www.researchgate.net/post/How_long_can_my...

A green arrow points to the tweet containing the chemical structures.

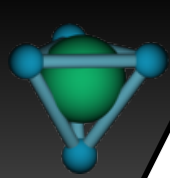
- Tweet got *harvested* into Tuberculosis topic
- Inline preview browsed, with other thumbnails



Chemistry Aware



- The app understands chemical data
- Users can use the data, *open with* in apps...

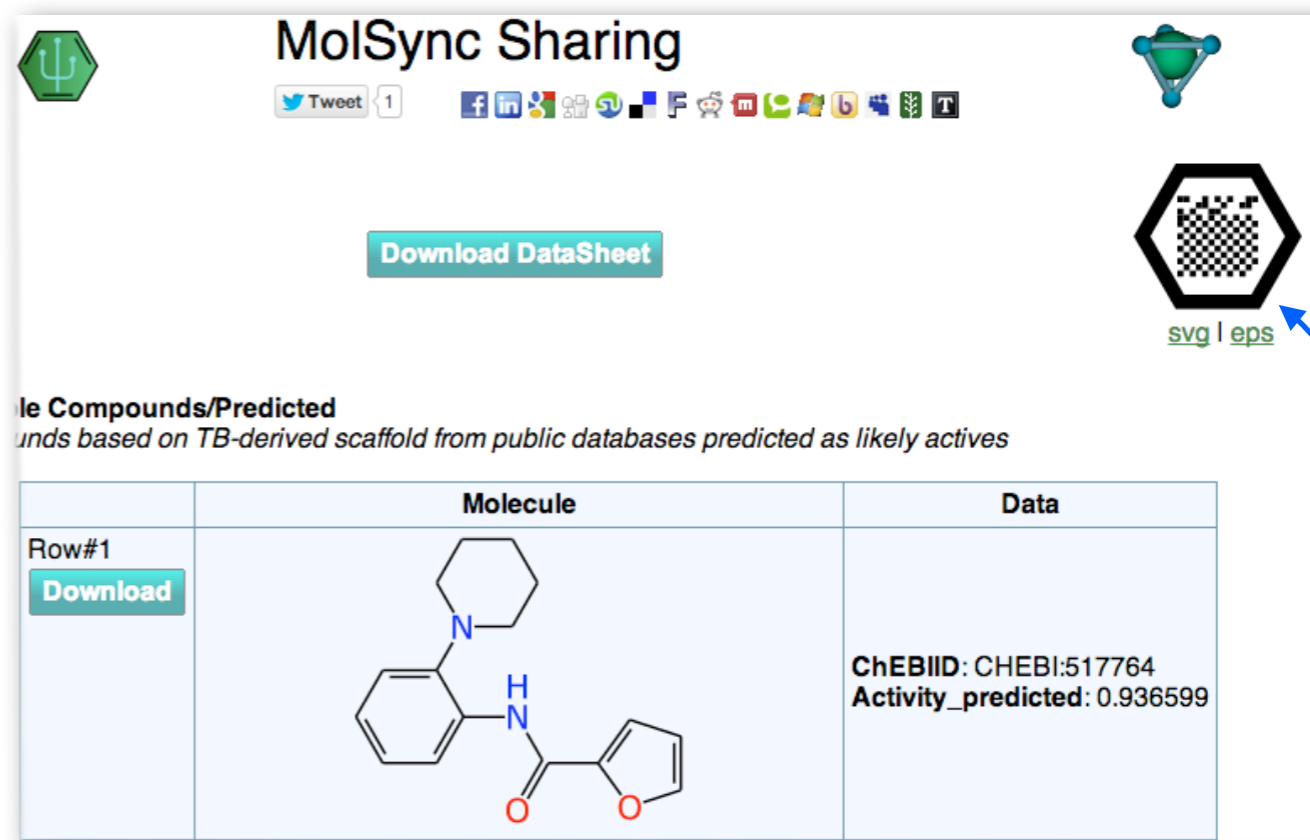


Rinse & Repeat

- Defer to experimentalists: measure activity for selected compounds with predictions
- Return to **SAR Table** app:
 - provide activity results
 - visualise matrix plot with actual data added
 - rebuild model, regenerate predictions
 - iteratively refine hypothesis
- Done using \$50 worth of software and tablet
- Workflow is appropriate for use by experimentalists

Living Molecules

- Going beyond desktop-era functionality
- Devices have:
 - long battery life
 - bluetooth
 - location awareness
 - **camera!**

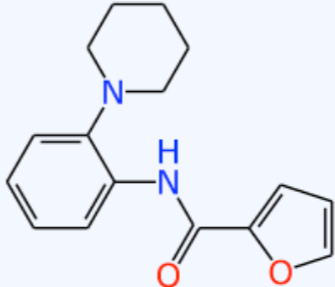


MolSync Sharing

Download DataSheet

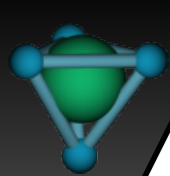
Download DataSheet

Multiple Compounds/Predicted
Compounds based on TB-derived scaffold from public databases predicted as likely actives

	Molecule	Data
Row#1 Download		ChEBIID: CHEBI:517764 Activity_predicted: 0.936599

svg | eps

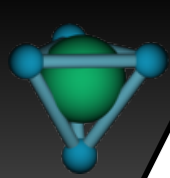
- Observe the strange symbol: **molecular glyph**



Molecular Glyphs



- Recognition sequence: point camera
- View content, import, manage, utilise, export



Conclusion

- Cheminformatics workflows historically the role of specialists: expensive and/or complex
- Mobile apps are much cheaper and much more accessible to experimentalists
- Mobile+cloud can:
 - **replace** simple-to-medium tasks
 - **coexist** with complex tasks run on desktop software
- Other advantages:
 - anywhere/anytime portability
 - excellent collaboration and sharing
 - non-existent installation or maintenance burden

Acknowledgments

<http://molmatinf.com/venice.html>

- Antony Williams
- Sean Ekins
- Joel Freundlich
- NETTAB / Bioinformatics
Italian Society
- InfoChem, Eidogen–Sertanty
- Inquiries to
info@molmatinf.com

**MOLECULAR
MATERIALS
INFORMATICS**

<http://molmatinf.com>

<http://molsync.com>

<http://cheminf20.org>

@aclarkxyz

