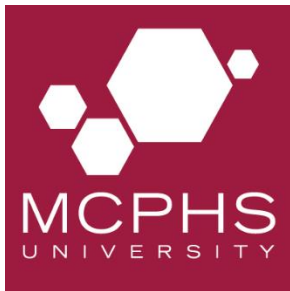


# Development of a Drug Discovery Simulation Laboratory Exercise in the Pharmaceutical Sciences Graduate Program Curriculum

---



*Chase Smith, Ph.D.*

*Assoc. Prof. of Medicinal Chemistry*

*School of Pharmacy (Worcester / Manchester)*



# Outline

- Overview of Curriculum (Worcester/Manchester)
  - Graduate Coursework / Laboratory Sequence
  - Laboratory Exercise Development
- Software & Data Selection
- Laboratory Workflow & Examples
  - Results / Issues
- Future Directions

# PharmD Program Overview

- 300 Students / year
- DE Worcester / Manchester
- PharmD (Accelerated)
- Coursework in MedChem, Pharmacology & Toxicology
- Many w/ a Chemistry or Biochemistry B.S. Degree

➤ **Need for Electives**

➤ **Mobile Technology**



# Graduate Program Overview

- 2-4 Students / year
- Worcester
- Programs in Pharmacology or Pharmaceutics
- M.S. & Ph.D.
  
- Courses DE Boston
- Introduction to Pharmaceutical Sciences



# Material Covered in Lecture

- ***“Introduction to Pharmaceutical Sciences” Lecture Course***
- ***Medicinal Chemistry Focused Topics:***
  - ❑ **Acid / Base Properties / Ionization State**
  - ❑ **Review of Organic Functional Groups**
  - ❑ **Water Solubility / LogP / LogD**
  - ❑ **Polar Surface Area (PSA)**
  - ❑ **Rule of 5**
  - ❑ **SAR (Sterics / Conformation / Electronics / Stereochemistry)**
  - ❑ **QSAR / 3D-QSAR**
  - ❑ **Bioisosteres / Scaffold Hopping**
  - ❑ **Principles of Drug Metabolism**
  - ❑ **ADME**

# Laboratory Exercise Development

- **6 Lab Rotations (5 Weeks Each)**
  - Follows the “Introduction to Pharmaceutical Science” Course
  - Covers basic lab techniques in Pharmacology & Pharmaceutics
  - Also designed to aid the student in choosing a Faculty Mentor
  - 1 Rotation in MedChem
- **Problem: How to provide a meaningful MedChem lab experience during the rotation that would have application in a Pharmacology or Pharmaceutics program**
- **Decision: Introduce a simulated MedChem discovery program**

# Laboratory Exercise Development

## 1. Software Selection

- Optibrium StarDrop™
  - iPad Apps
- 
- iPads not universally owned
  - Apps available still in infancy
  - Multiple Apps & Horsepower
  - Familiarity w/ StarDrop™
  - Glowing Molecule™
  - Possibility of adding modules in the future if successful
- 
- <http://www.optibrium.com/stardrop/>



# Laboratory Exercise Development

## 2. Screening Data Selection

- ChEMBL – NTD Database
- Malaria Screening Data

### ○ Screening Results Available

### ○ Already using literature example in lecture

- *Nature* 465, pp. 305–310 (20 May 2010)

### ○ <https://www.ebi.ac.uk/chemblntd>

The screenshot shows the ChEMBL-NTD website. The header includes the EMBL-EBI logo and navigation links for Services, Research, and Training. The main navigation bar features buttons for ChEMBL, Downloads, UniChem, Malaria Data, and ChEMBL-NTD. The ChEMBL-NTD Home page is displayed, featuring a welcome message, a list of deposited sets (Set 1: GSK TCAMS to Set 13: St. Jude Children's Research Hospital Dataset), and a section for the GSK TCAMS Dataset. The St. Jude Children's Research Hospital logo is visible, along with a detailed description of the NTD dataset and its use in research.



# Lab Exercise Development & Workflow

- **Download Data**

- Scrub Data

- Remove unneeded data columns
- Remove metals, salts...

- Fragment into smaller libraries

- Faster processing of datasets
- An attempt to prevent the same chemical series to be found by multiple students
- Provide students with secondary screening library to search for like scaffolds to develop SAR

- Libraries with approx. 1000 compounds each



# Lab Exercise Development & Workflow

- **Distribute Data**

- Demonstrate StarDrop™ Software
- Explain Objective of Exercise
  - Perform primary screen
  - Identify 2 hit series based on:
    - Physicochemical Properties
    - Activity



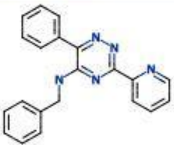

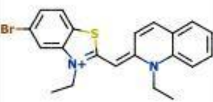
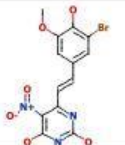
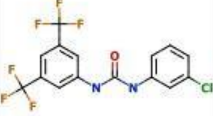
- **1<sup>st</sup> Exercise**

- Students receive fragmented datasets
- Students run calculations and choose 2 hit series

# Initial Student Calculations & Data Triage

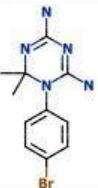
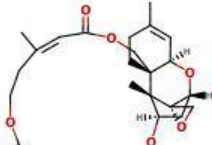
- Students Run Selected Calculations

The screenshot displays the StarDrop software interface. On the left, a sidebar lists 'Available Models' with various calculation options checked, including logS, logS @ pH7.4, logP, logD, 2C9 pKi, hERG pIC50, BBB log([brain]:[blood]), BBB category, HIA category, P-gp category, 2D6 affinity category, PPB90 category, MW, HBD, HBA, TPSA, Flexibility, and Rotatable Bonds. The main window shows a table with five rows of data, each representing a different chemical structure. The columns are: STRUCTURE (with a chemical structure image), GNF-Pf identifier, Pf Proliferation Inhibitor, W2 Pf Proliferation Inhib, and Huh7 Cytotox for Pf Inhil.

	STRUCTURE	GNF-Pf identifier	Pf Proliferation Inhibitor	W2 Pf Proliferation Inhib	Huh7 Cytotox for Pf Inhil
1		GNF-Pf-100	0.1821	0.326	0.63
2		GNF-Pf-1000	0.547	1.086	10
3		GNF-Pf-1001	0.645	1.084	10
4		GNF-Pf-1003	1.25	0.987	9.11
5		GNF-Pf-101	3.22	9.99	19.46

# Initial Screening Hits by Students

- **General Trends in Results from the 1<sup>st</sup> Exercise**
  - Students tended to focus choices on Activity
    - Generally observed Ro5
    - Little concept of scaffold searching and SAR
    - Little use of Visualization vs. Spreadsheet View

	STRUCTURE	GNF-Pf identifier	Pf Proliferation Inhibitor	W2 Pf Proliferation Inhib	Huh7 Cytotox for Pf Inhib	MW	logP	HBA
1		GNF-Pf-1945	0.000592	0.533	10	296.2	0.9297	5
2		GNF-Pf-2151	0.00391	0.00918	0.2064	484.5	1.14	8

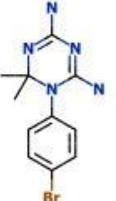
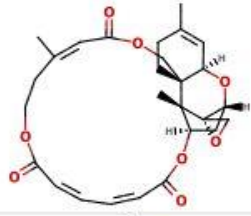
# Follow-on Screening Triage

- **2<sup>nd</sup> Exercise – Utilize Scoring Function**
  - Students are made to consider all physicochemical parameters in addition to activity – *Composite Score*
  - Students Again Asked to Choose Top 2 Hits
    - Based partially on SAR from structure searches

The screenshot shows the StarDrop software interface. The main window displays a table with columns for chemical structures and various properties. A 'StarDrop' dialog box is open, showing the 'Scoring' profile 'Second Search' with the following properties and values:

Property	Desired Value	Importance
HBA	≤ 10	High
HBD	≤ 5	High
MW	≤ 500	High
logP	≤ 5	High
Pf Proliferation Inhi...	-inf -> 0	High

The main table shows the following data:

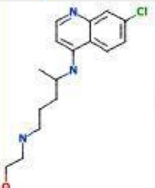
Structure	GNF-Pf identifier	Pf Proliferation Inhibition	W2 Pf Proliferation Inhib	Huh7 Cytotox for Pf Inhi	MW	logP	
	GNF-				10	296.2	0.9297
	GNF-				0.2064	484.5	1.14

# Follow-on Screening Triage

- **General Trends in Results from the 2<sup>nd</sup> Exercise**
  - Overall, the students tended to find different hit compounds as compared to their first search
  - New appreciation of looking at additional information beyond activity

StarDrop

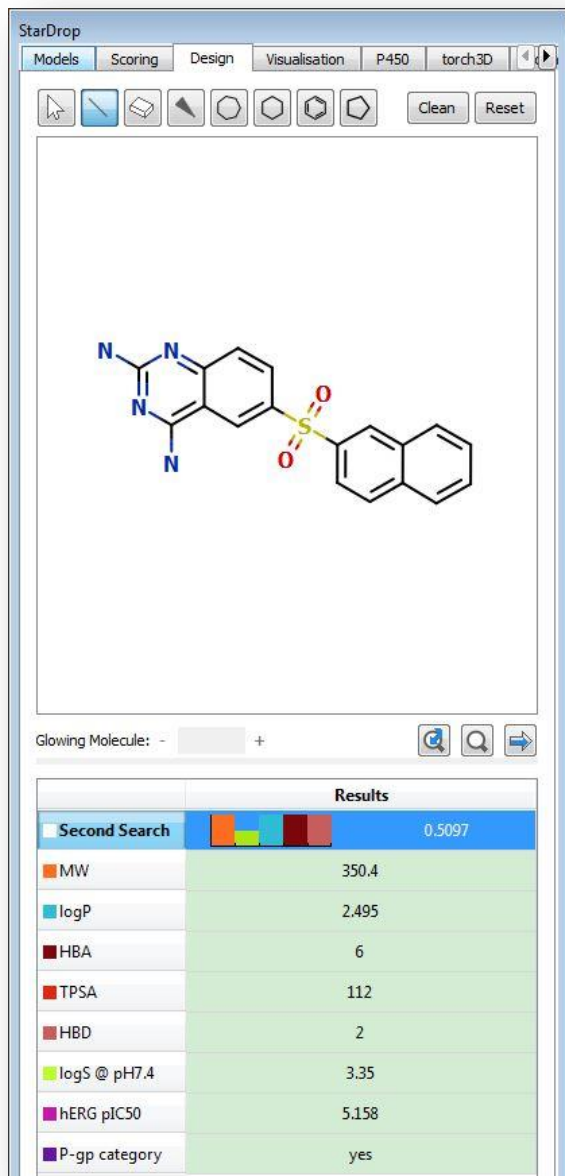
File Edit View Data Set Tools Custom Scripts Help

Second Search	STRUCTURE	GNF-Pf identifier	Pf Proliferation Inhibitor	W2 Pf Proliferation Inhib	Huh7 Cytotox for Pf Inhib	MW	logP
1	 GNF-Pf-1435	0.5484	0.01294	0.01398	0.1733	394.4	2.578
2	 GNF-Pf-1952	0.5009	0.02693	0.736	13.98	307.8	2.382
3	 GNF-Pf-2448	0.4978	0.0973	1.709	49.48	350.4	2.495

# Transition to Compound Optimization

- **Transition to 3<sup>rd</sup> Exercise**
  - Students informed that only enough resources are available to work on a single Series
    - **Student then must begin to suggest modifications to potentially fix most serious compound liabilities**
  - Evaluation of choice involves discussion of all characteristics of the chosen series
    - Activity / Selectivity
    - SAR (usually 3-10 analogs)
    - Calculated physicochemical parameters

# Transition to Compound Optimization



The screenshot shows the StarDrop software interface. At the top, there are tabs for Models, Scoring, Design, and Visualisation. Below the tabs is a toolbar with various icons and buttons for 'Clean' and 'Reset'. The main area displays a chemical structure of a quinazolinone derivative with a naphthalene-1-sulfonyl group. Below the structure, there is a 'Glowing Molecule' section with a search bar and navigation icons. At the bottom, a table titled 'Results' shows various molecular descriptors and their values.

	Results
Second Search	0.5097
MW	350.4
logP	2.495
HBA	6
TPSA	112
HBD	2
logS @ pH7.4	3.35
hERG pIC50	5.158
P-gp category	yes

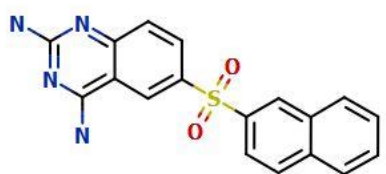
- **Student Selected Single Hit Series**
  - 3<sup>rd</sup> Exercise very open ended
  - Wide latitude given
  - Informed that any changes need to fit within available SAR if possible
  - *Most trying portion of lab exercise*
  - *Weakness in chemistry background difficult to overcome*
  - *Great deal of coaching required*
  - *Glowing Molecule™*



# Transition to Compound Optimization

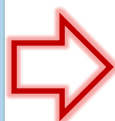
StarDrop

Models Scoring Design Visualisation P450 torch3D



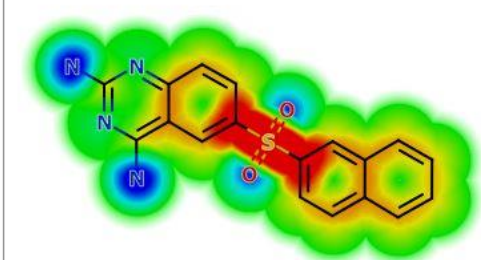
Glowing Molecule: - +

	Results
Second Search	0.5097
MW	350.4
logP	2.495
HBA	6
TPSA	112
HBD	2
logS @ pH7.4	3.35
hERG pIC50	5.158
P-gp category	yes



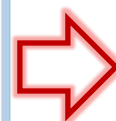
StarDrop

Models Scoring Design Visualisation P450 torch3D



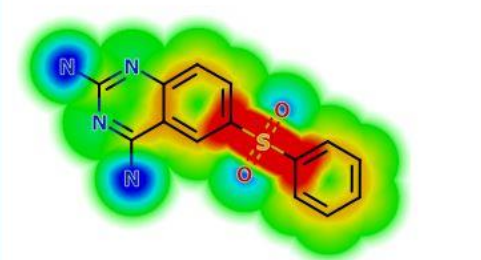
Glowing Molecule: - +

	Results
Second Search	0.5097
MW	350.4
logP	2.495
HBA	6
TPSA	112
HBD	2
logS @ pH7.4	3.35
hERG pIC50	5.158
P-gp category	yes



StarDrop

Models Scoring Design Visualisation P450 torch3D



Glowing Molecule: - +

	Results
Second Search	?
MW	300.3
logP	1.475
HBA	6
TPSA	112
HBD	2
logS @ pH7.4	3.764
hERG pIC50	4.605
P-gp category	no

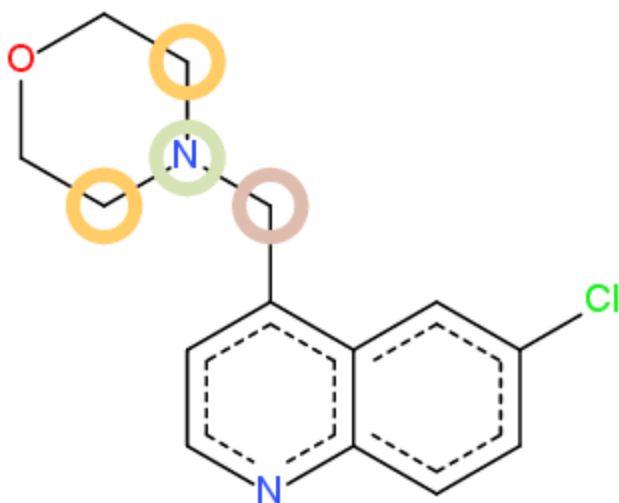
# Evaluation of Proposed Compounds

- **Wrap-Up Exercises**
  - Use of hit series and newly proposed analogs
  - Method to benchmark new analogs and illustrate that they are still very early on the discovery timeline
  - Provides a means to review Metabolism
    - **SMARTCyp Web Service**
    - **SAR Table<sup>©</sup> Model Building**

# Evaluation of Proposed Compounds

- SMARTCyp Web Service

- Molecular series uploaded to server
- Analogs judged in comparison to original compounds
- “**Gamification**” of Results



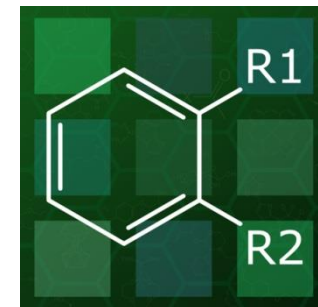
P. Rydberg, D. E. Gloriam, J. Zaretski, C. Breneman and L. Olsen, *ACS Med. Chem. Lett.*, **2010**, 1, 96-100; P. Rydberg, D. E. Gloriam and L. Olsen, *Bioinformatics*, **2010**, 26, 2988-2989; P. Rydberg and L. Olsen, *ACS Med. Chem. Lett.*, **2012**, 3, 69-73; P. Rydberg and L. Olsen, *ChemMedChem*, **2012**, 7, 1202-1209; P. Rydberg et al., *Angew. Chem, Int. Ed.*, **2013**, 52, 993-997; and P. Rydberg et al., *Mol. Pharmaceutics*, **2013**, 10, 1216-1223.

	Standard	CYP2C	CYP2D6			
<b>1: null</b>						
	Rank	Atom	Score	Energy	Accessibility	2DSASA
1	1	C.12	33.55	41.1	0.78	33.17
2	2	C.10	35.87	41.1	0.56	19.57
3	3	N.11	37.12	42.6	0.67	3.75
4	4	C.13	53.59	62.2	0.89	37.53
5	5	N.6	67.67	75.6	0.89	20.4
6	6	C.4	68.99	77.2	0.89	27.38
7	7	C.3	71.67	80.8	1	28.22
8	8	C.18	73.71	80.8	0.78	21.6
9	9	C.8	79.91	86.3	0.67	26.32
10	10	C.7	84.45	92.0	0.78	33.21
11	11	C.2	991.68	999	0.89	5.28
12	12	C.5	992.49	999	0.78	7.23
13	13	C.17	993.47	999	0.67	4.99
14	14	C.9	994.46	999	0.56	2.37

# Evaluation of Proposed Compounds

- SAR Table

- iPad Application / Cheminformatics Tools
- Ability to generate Schemes & Models



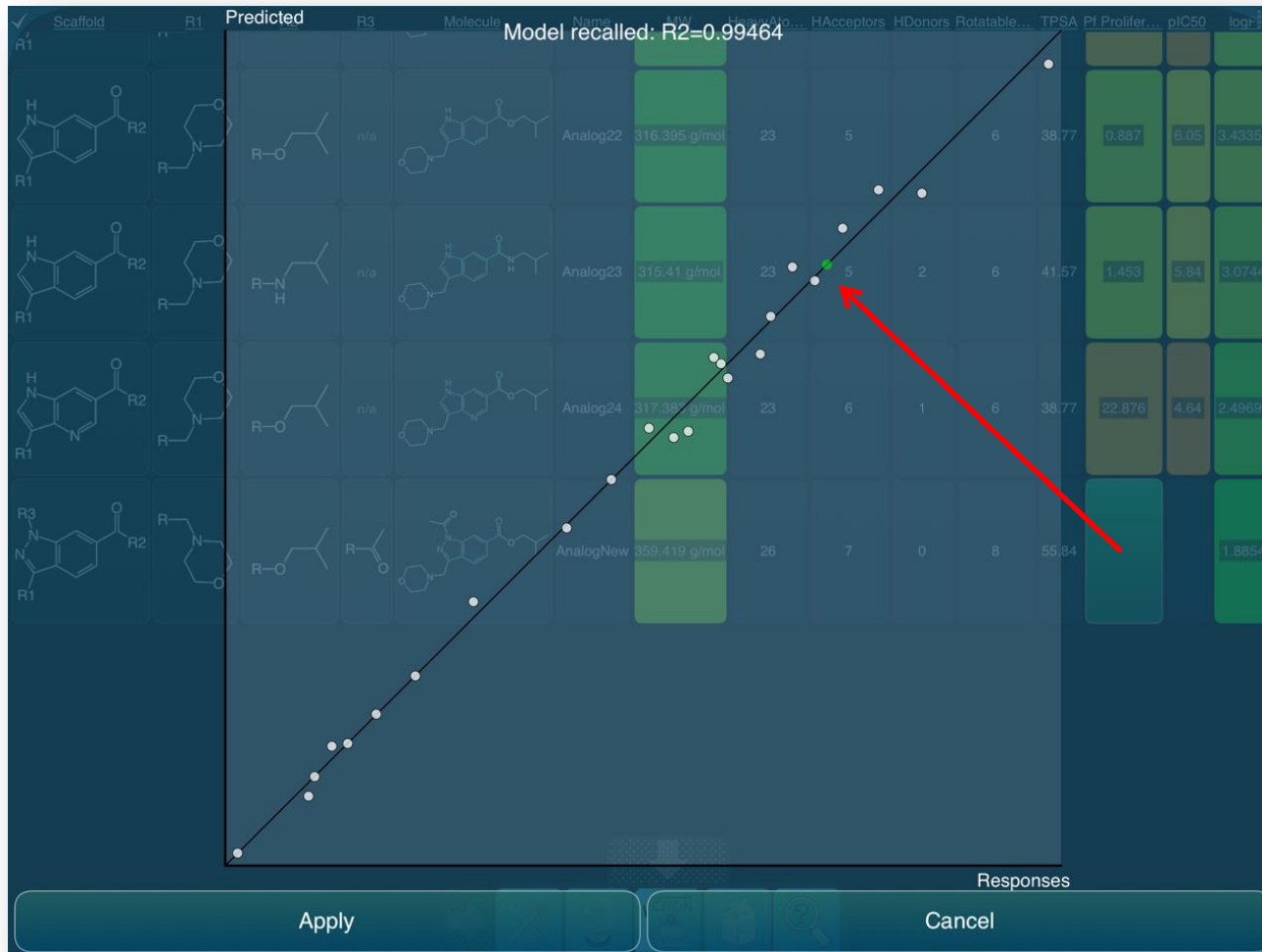
Scaffold	R1	R2	R3	Molecule	Name	MW	HeavyAto...	HAacceptors	HDonors	Rotatable...	TPSA	Pf Prolifer...	pIC50	logP
			R-H		Analog1	304.424 g/mol	22	3	1	6	46.53	15.921	4.8	6.68814
					Analog2	346.461 g/mol	25	4	0	8	52.6	11.248	4.95	6.51667
			R-H		Analog3	306.397 g/mol	22	4	1	6	55.76	3.657	5.44	4.61409
					Analog4	348.433 g/mol	25	5	0	8	61.83	2.745	5.56	4.44262
			R-H		Analog5	307.385 g/mol	22	5	1	6	59	1.633	5.79	2.77273
					Analog6	349.421 g/mol	25	6	0	8	65.07	0.65	6.19	2.67401

Used to generate predicted biological activity of new analogs.

© 2011-2015 Molecular Materials Informatics  
<http://molmatinf.com/>

# Evaluation of Proposed Compounds

- SAR Table



1. Build a scheme for any of the properties calculated or experimentally derived.
2. Application builds model and makes prediction.

# Evaluation of Proposed Compounds

- SAR Table

Scaffold	R1	R2	R3	Molecule	Name	MW	HeavyAto...	HAcceptors	HDonors	Rotatable...	TPSA	Pf Prolifer...	pIC50	logP
			n/a		Analog22	316.395 g/mol	23	5	0	6	38.77	0.887	6.05	3.43352
			n/a		Analog23	315.41 g/mol	23	5	2	6	41.57	1.453	5.84	3.0744
			n/a		Analog24	317.383 g/mol	23	6	1	6	38.77	22.876	4.64	2.49699
					AnalogNew	359.419 g/mol	26	7	0	8	55.84	(0.774581 μM)	(5.846 96)	1.8854

# Current Status of Lab Exercise

- **Lessons Learned**

- Positive response from students
- Fills a gap in graduate curriculum
- Effective method to review and reinforce abstract physicochemical properties covered in lecture
- Identity of the Screening Target not important
- Need to allow an iterative approach to any products (calculation runs, target selections, etc.) turned in by the students
- ***Baseline chemistry knowledge weak and will not be correctable in the timeframes allotted***

# Future Plans

- **Scale and Expand Scope of Laboratory Exercise**
  - Recycle current concept into Elective Course format
    - Pharm.D. Students
    - Scale to ~10 – 25 Students (*Need more data sets*)
    - Seed Screening Databases w/ Active Compounds (?)
  - Require additional simulated drug discovery tasks
    - Increase drug optimization cycles and requirements
    - hERG Pharmacophore Matching
    - Expand Metabolism Profiling / CYP Profile
    - Include Molecular Modeling / Docking
- **Incorporate mobile software & platforms**



# Next Steps

