



## Reaction-based enumeration: Lessons learned in designing a workflow that chemists want to use

27<sup>th</sup> March 2023 – ACS National Meeting, Indianapolis – CINF

Tamsin Mansley – [tamsin@optibrium.com](mailto:tamsin@optibrium.com)

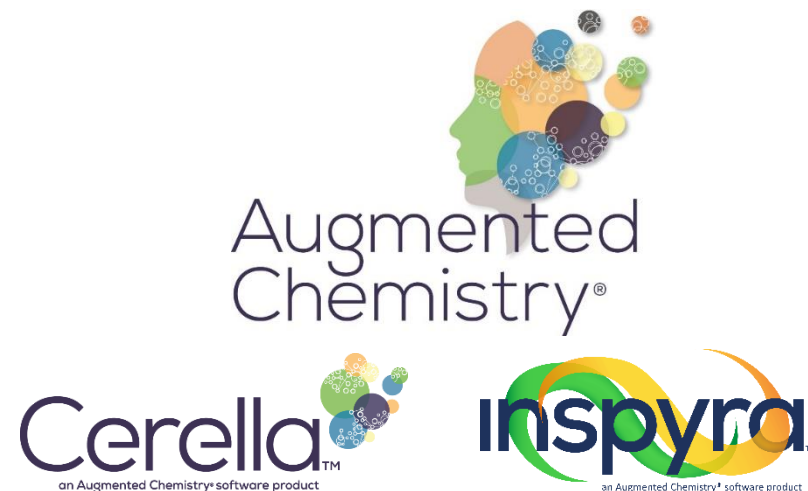
# Agenda

---

- Introduction to Optibrium
- Scaffold and Reaction-based Library Enumeration
- Reaction-based Library Enumeration
  - Reactions
  - Reagents
  - Products
  - Beyond Chemistry
- Conclusions

# Introduction to Optibrium

[www.optibrium.com](http://www.optibrium.com)



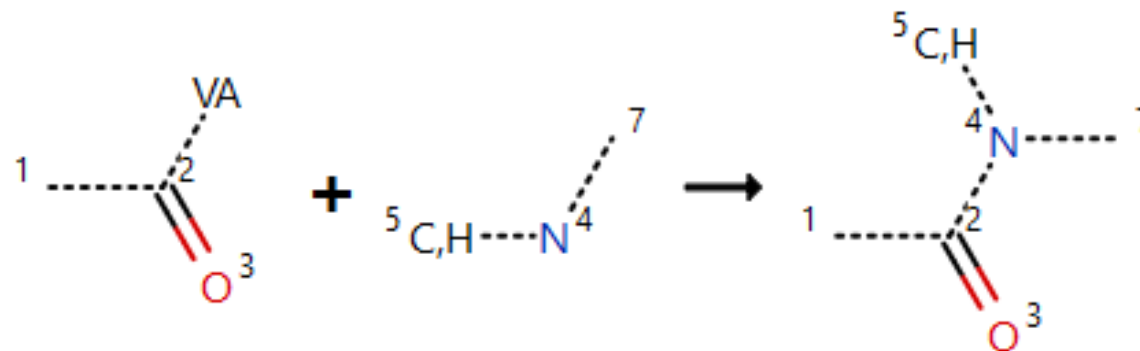
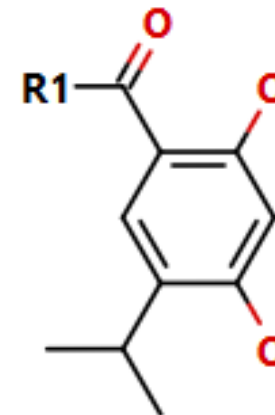
- Optibrium creates elegant software solutions for small molecule design, optimisation and data analysis
  - R&D of novel technologies to guide decisions and improve efficiency in drug discovery
- Two platforms forming an integrated ecosystem
  - StarDrop™
  - Augmented Chemistry®
- Global customer base from top-ten pharma to small biotech and academia
  - >180 customers worldwide
  - Adoption in other chemistry fields, e.g. animal health, agrochemicals, etc.



# Types of Library Enumeration

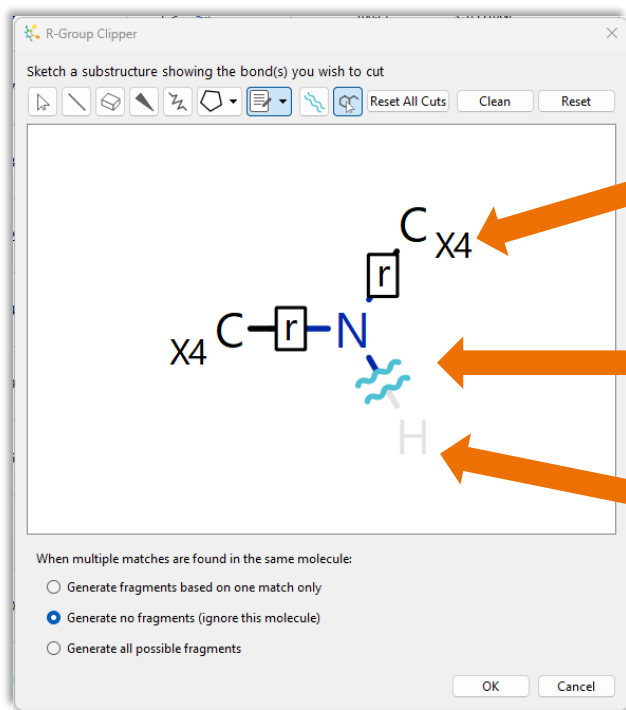
---

- Library enumeration involves joining fragments
  - R-groups, linkers and scaffolds
- Scaffold-based enumeration
  - Defined from the product
  - Uses pre-prepared lists of fragments
- Reaction-based enumeration
  - Defined from a reaction
  - Uses lists of reagents
  - Creates fragments on the fly



# Reagent clipping

- Flexible fragment definition
  - From lists of reagents or building blocks
    - o Metadata retained with fragment
  - Save and reuse fragment libraries



Set atom and bond constraints

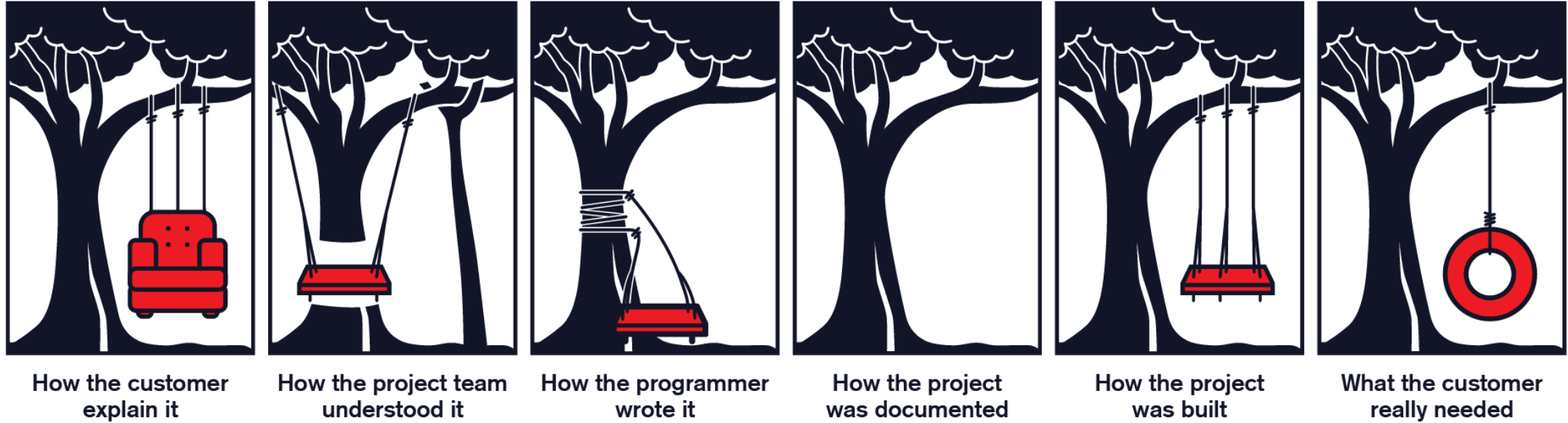
Define the bond to cut

Define the fragment to exclude

New column of fragments with metadata

VID	SMILES	Fragment1_0	eMolecules URL	Quantity	MWT	Units
1 96276481			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	1	201.27	g
2 48572294			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	100	188.27	mg
3 44469267			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	1	202.29	g
4 36251563			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	0.25	202.29	g
5 26985702			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	1	188.27	g
6 17413314			<a href="https://www.emolecules.com/">https://www.emolecules.com/</a>	5	190.29	g

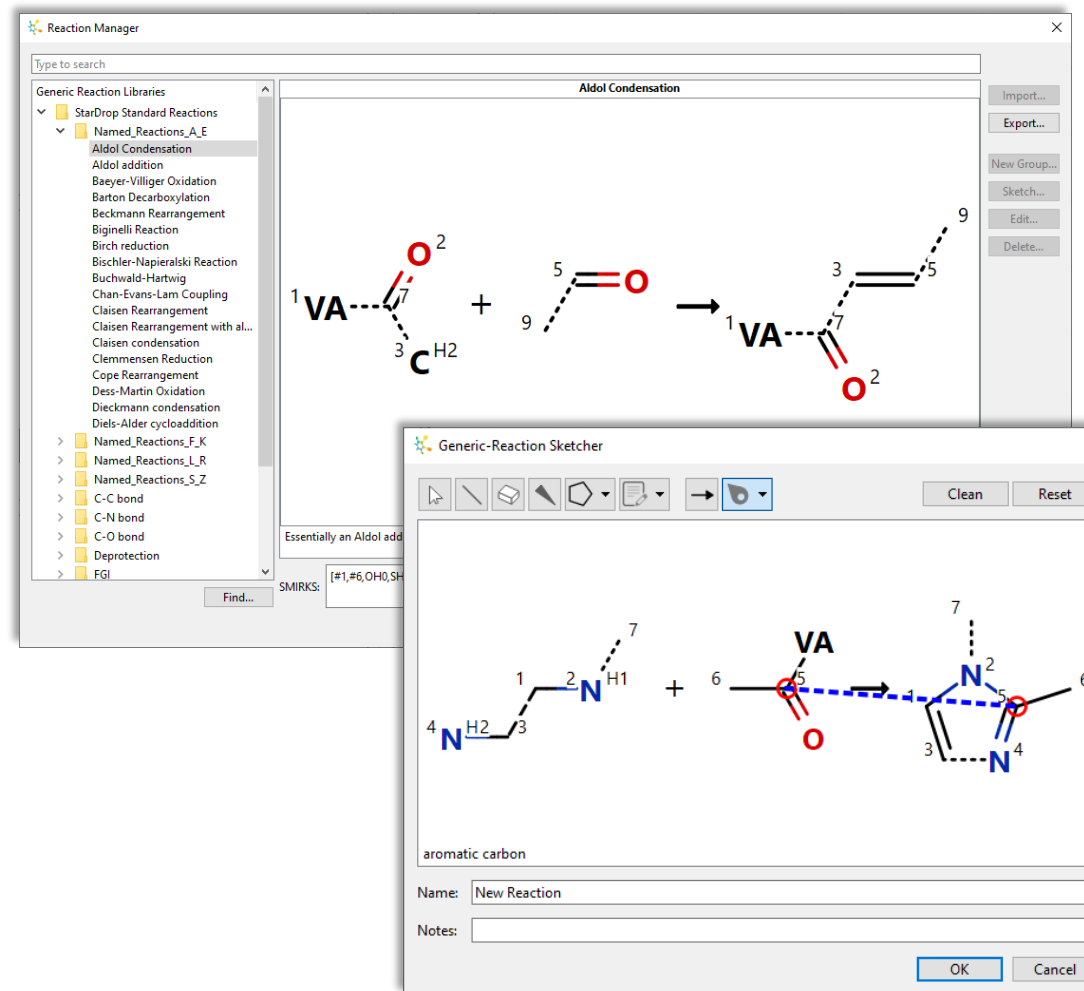
# What a Scientist Really Wants



- Importance of working with users at every step
  - Scope, test and deliver
- User-friendly, flexible workflow that meets user requirements

# Reactions

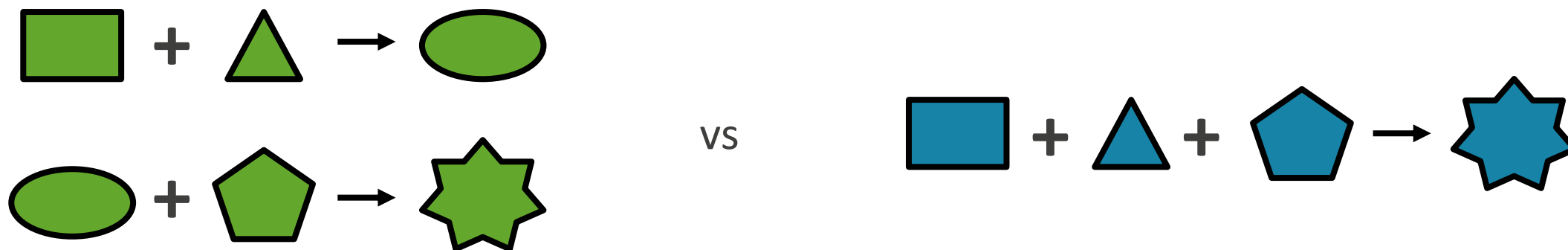
- Approximately 100 pre-defined common reactions
  - Create centrally-managed and individual reaction libraries
- Import RXN files
- Edit existing reactions or define your own
  - On-the-fly
  - Save and share



## Flexible Reaction Definition

---

- No requirement to mimic the real life lab process
- Combine multiple steps in a single enumeration

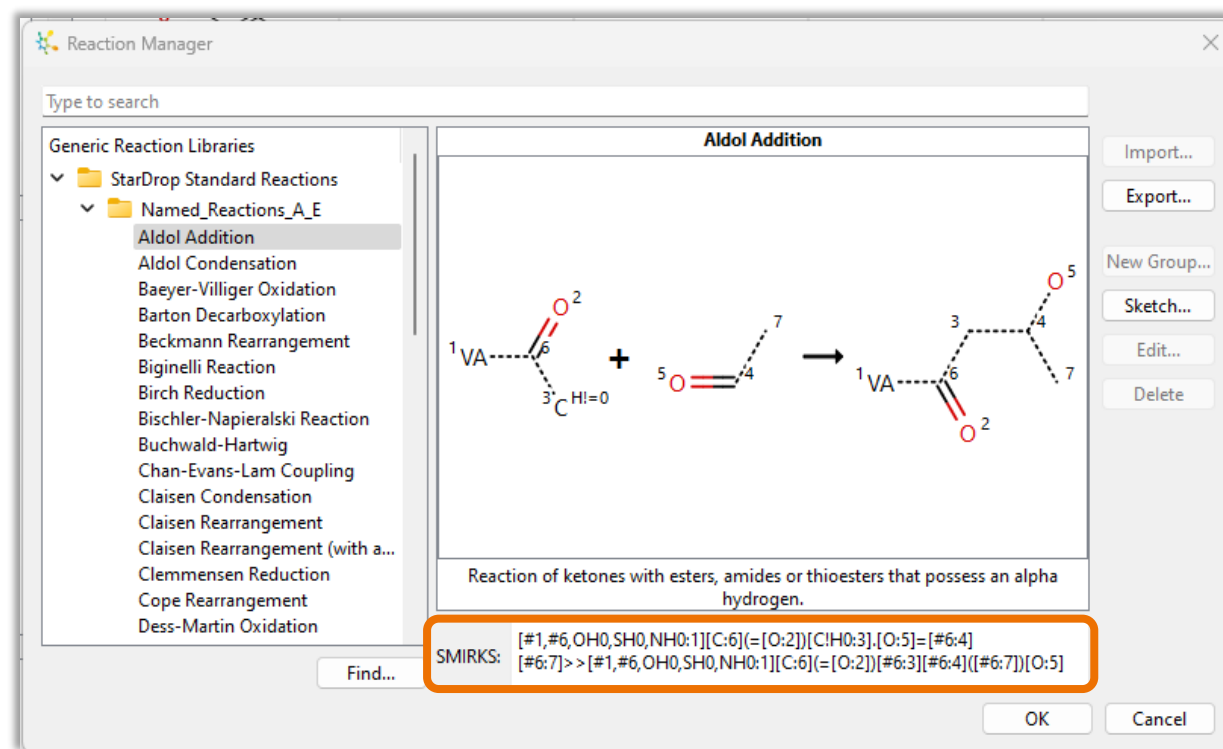


- For example, BOC-deprotection followed by amide formation



# SMIRKS

- SMIRKS is the language the computer uses to define a reaction
  - [#1,#6,OH0,SH0,NH0:1][C:6](=[O:2])[C!H0:3].[O:5]=[#6:4][#6:7]>>[#1,#6,OH0,SH0,NH0:1][C:6](=[O:2])[#6:3][#6:4]([#6:7])[O:5]
- Encodes reaction, variable atoms and atom/bond constraints
- Also represent SMIRKS graphically
  - Enables graphical editing
  - Support multiple levels of understanding
- Export SMIRKS to other platforms



# Reagents

- Reagent lists as StarDrop data sets
  - Minimal preparation
  - Can be reused flexibly
- Configure and filter on the fly
  - Detection of unsuitable reagents
  - Selection criteria (filters)
  - Define regioselectivity

Reaction Based Enumeration

### Select Reagents

General Amide Formation

Reagent Group 1   Reagent Group 2

Choose data set: Carboxylic acids

	SMILES	Canonical SMILES	MolPort ID	ID	Unverified Amount	Verified
1			MolPort-016-937-956	<a href="#">16937956</a>	15	
2			MolPort-016-939-029	<a href="#">16939029</a>	400	
3			MolPort-016-965-516	<a href="#">16965516</a>	2.6e+03	2
4			MolPort-017-007-387	<a href="#">17007387</a>	1e+04	1

Displaying 83 of 83 rows

Display suitable reagents only

Selection Criteria: Set... Clear

Suitable reagents: 83

Reagents matching multiple sites: 12

Use one site  
 Use all sites  
 Use none  
 Choose sites: Select...

< Back   Next >   Cancel

# Reagents

- Reagent lists as StarDrop data sets
  - Minimal preparation
  - Can be reused flexibly
- Configure and filter on the fly
  - Detection of unsuitable reagents
  - Selection criteria (filters)
  - Define regioselectivity

The screenshot displays the 'Reaction Based Enumeration' software interface. At the top, the title bar reads 'Reaction Based Enumeration'. Below it, the 'Select Reagents' window is active, showing a chemical reaction scheme for 'General Amide Formation'. The reaction involves a carboxylic acid derivative (with atoms labeled 1, 2, 3, and VA) reacting with an amine (with atoms labeled 5, C, H, N, 4, 7) to form an amide product (with atoms labeled 1, 2, 3, 4, N, 5, C, H, 7). Below the reaction scheme, there are tabs for 'Reagent Group 1' and 'Reagent Group 2', and a dropdown menu for 'Choose data set:' set to 'Carboxylic acids'. A toolbar contains icons for 'SMILES', 'Canonical SMILES', 'MolPort ID', 'ID', 'Unverified Amount', and 'Verified Amount', along with a 'Selection Criteria' button. The 'Selection Criteria' dialog box is open, showing options to 'Keep reagents where:' with radio buttons for 'All criteria are met' (selected), 'At least one criterion is met', and 'No criteria are met'. A list of chemical classes is shown, including Ester, Carb\_x\_ates, Aldehyde, Acyl halide, Anhydride, Aniline primary, Aniline secondary, Aniline tertiary, Imide, Isocyanate\_isothiocyanate, and Nitro. A 'Smiles' dropdown is set to 'contains none of'. A 'Select...' button is visible next to the list. At the bottom of the dialog are 'Add' and 'Clear' buttons, and an 'OK' button at the bottom right of the main window.

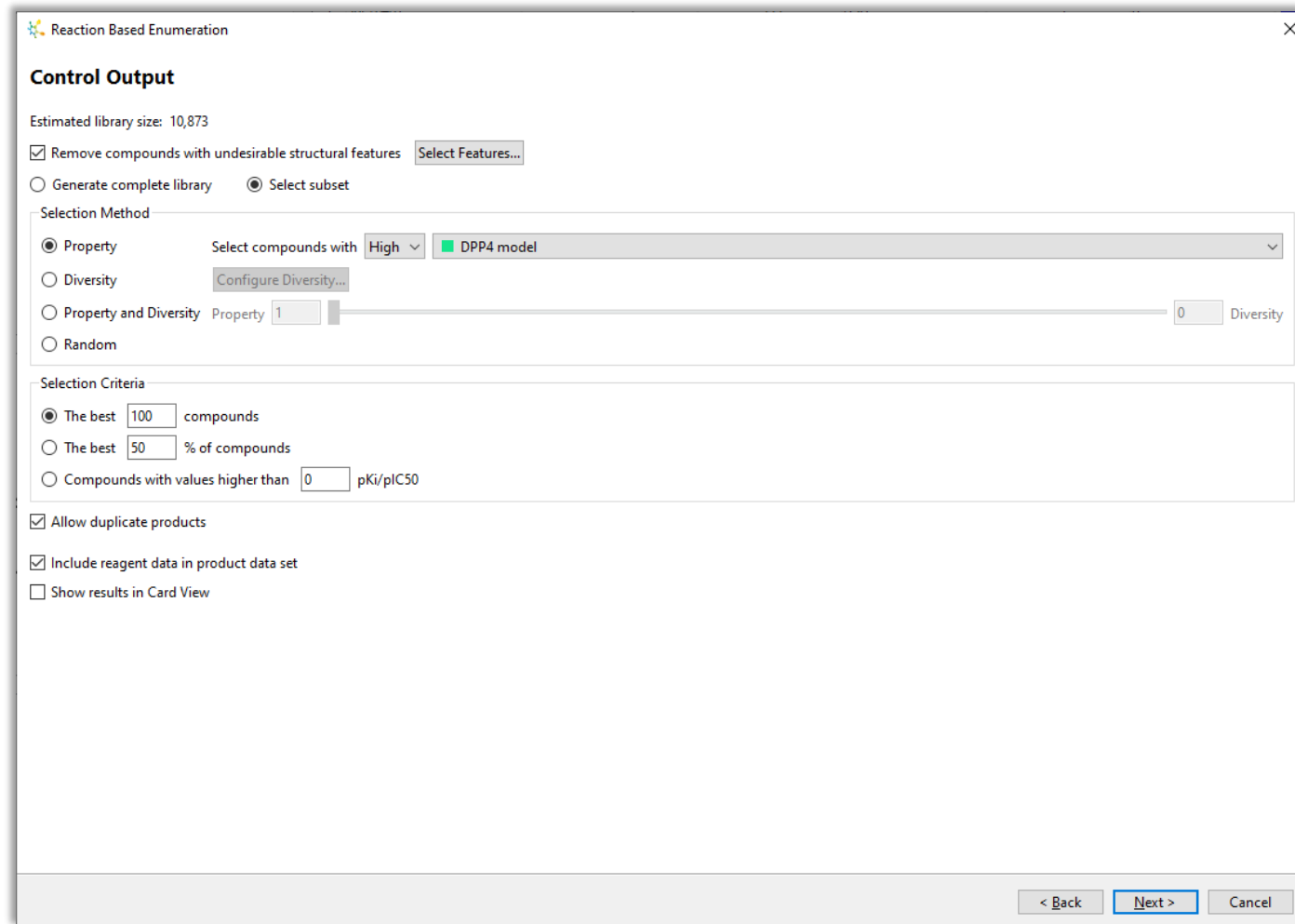
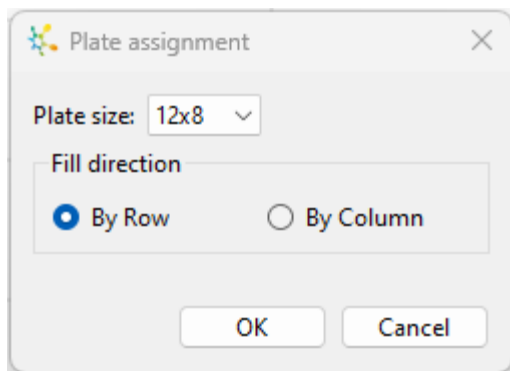
# Reagents

- Reagent lists as StarDrop data sets
  - Minimal preparation
  - Can be reused flexibly
- Configure and filter on the fly
  - Detection of unsuitable reagents
  - Selection criteria (filters)
  - Define regioselectivity

The screenshot displays the 'Reaction Based Enumeration' software interface. The main window is titled 'Select Reagents' and shows a reaction scheme for 'General Amide Formation' with a starting material labeled 'VA' and a carbonyl group labeled  $^5\text{C,H}$ . A 'Reaction Site Selection' dialog box is overlaid on the main window, showing a list of reagents matching multiple sites. The reagents are listed with their chemical structures and IDs: 107563022, 262417225 (highlighted), 38031866, and 50806033. The 'Select preferred site(s)' panel shows the chemical structure of the selected reagent, 262417225, which is a brominated and chlorinated benzene ring attached to a sulfonamide group. The 'Keep reagents where:' section has the radio button for 'All criteria are met' selected. The 'Smiles' checkbox is also checked. The 'Add' and 'Clear' buttons are visible at the bottom of the dialog box.

# Products

- Define selection criteria for products using a property, MPO score or diversity
- Include reagent metadata with products for easy ordering
- Apply to plate well locations



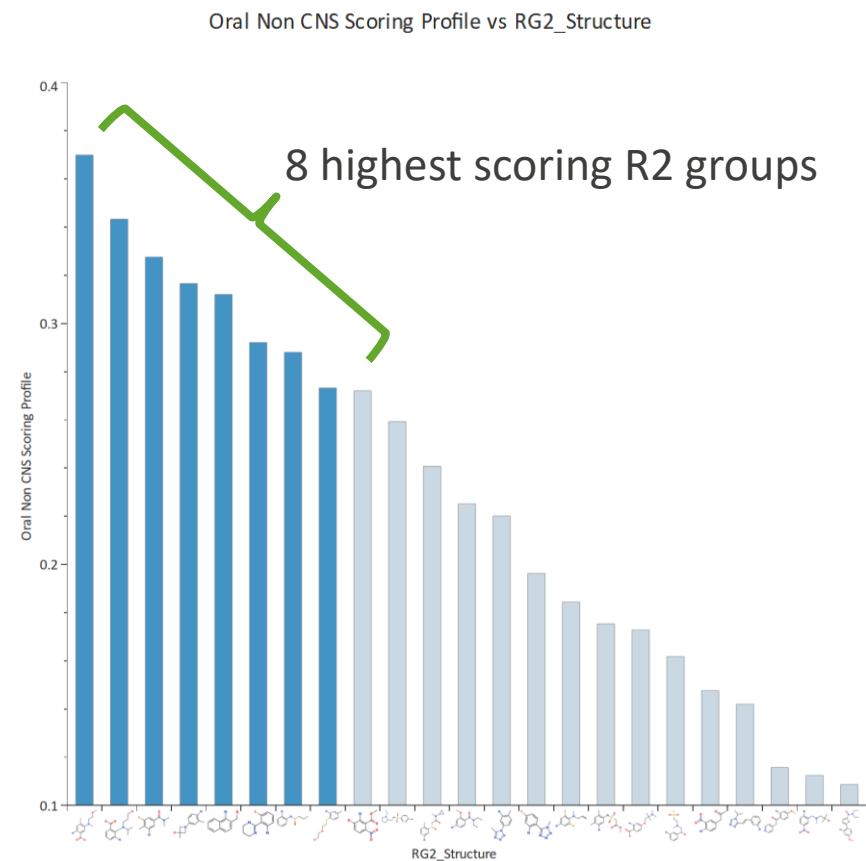
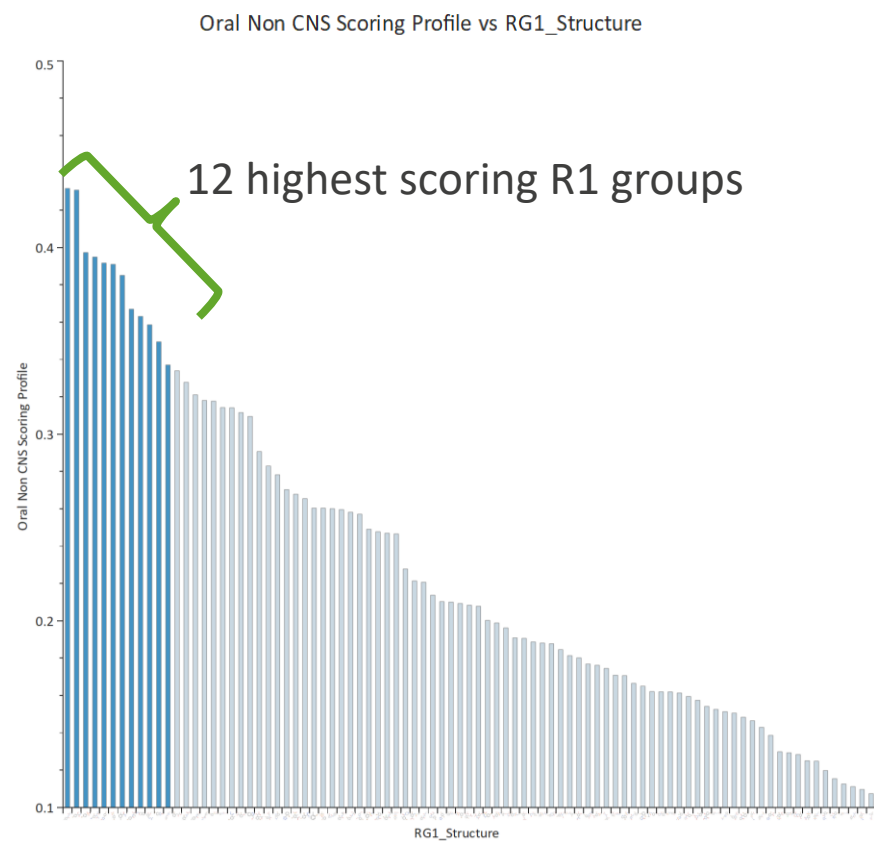
# Above and Beyond Chemistry

---

- Speed of enumeration
- Considering the process a chemist goes through
- Metadata and reagent IDs carry through
- Methods for reagent selection to fit plate dimensions or optimize a property
  - Ability to sample library using fewest reagents.

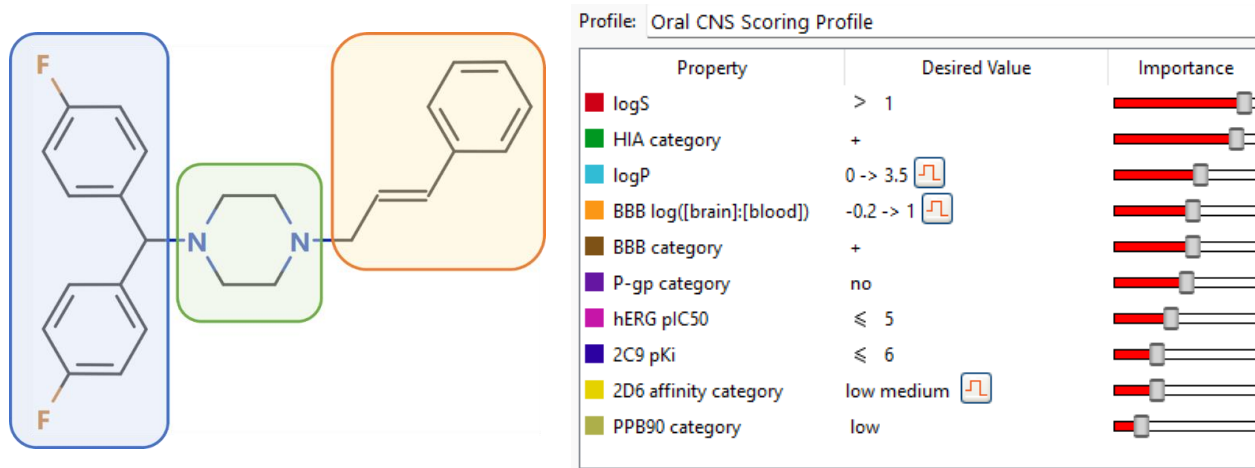
# Prioritising Reagents Using Product Properties

- Select the R-groups generating the highest scores
  - e.g 12 x R1 and 8 x R2 for a 96-well plate

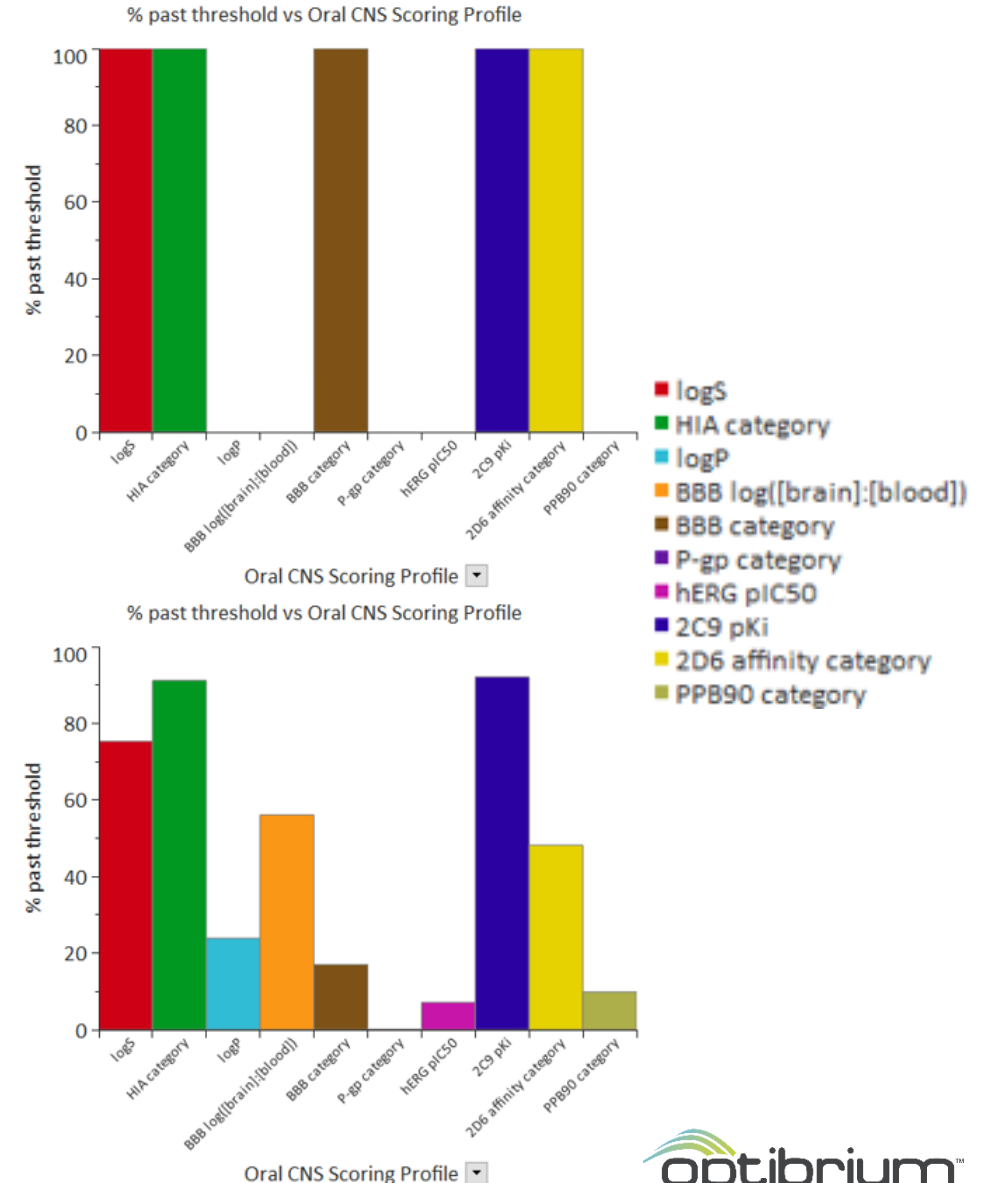


# Flunarizine Library

- 10,530 product library based on flunarizine



- Score for oral bioavailability and CNS penetration
- Some library compounds meet threshold for logP, BBB log([brain]:[blood]), hERG pIC50 or PPB90

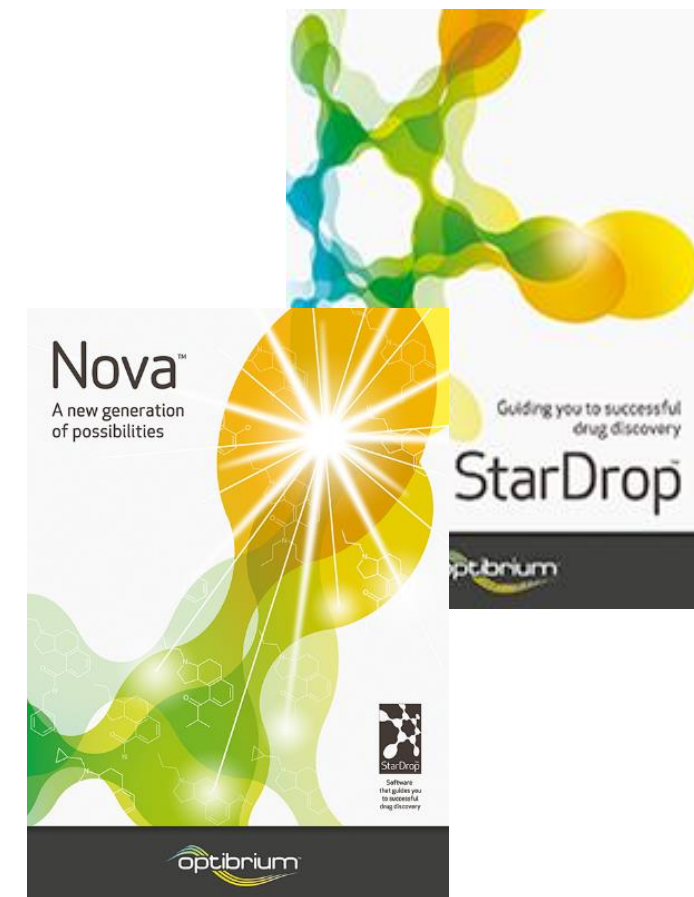




# Conclusions



- Reaction-based enumeration in StarDrop's Nova™ module is:
  - User friendly
    - o Easy to interpret
    - o Supports different levels of experience
  - Flexible
    - o Prepared reaction libraries and filters support major use cases
    - o Ability to add and customise for in-house chemistry and expertise
  - Fast
    - o Combine reactions in a single enumeration
    - o Metadata for easy inventory and building block requests
    - o Select an optimal subset of compounds for synthesis
- User feedback at every step of development is critical



# Acknowledgements

---

## Optibrium

- Fayzan Ahmed
- Matt Segall
- Ed Champness
- Peter Hunt
- Aishling Cooke

## Celgene

- Joe McDonald
- Andrew Burritt

For more information and a demo visit [optibrium.com/project/nova/](https://optibrium.com/project/nova/)

