The Efficient Enumeration of Reaction Networks.

MODGRAPH

Reaction Based Enumeration.

- Developed jointly with Evotec OAI.
- 'Reaction Based' - reactions are predefined.
- Choice of Reactant Classes, aids specificity.
- Flexible reagent selection.
- 'Reaction Network' GUI developed through discussions with Harvard Med. School.
- Integrated with Modgraph Registry and CSC (Chemical Supplier Catalogue).

Library Builder

- Exploits systematic nature of Daylight SMIRKS and SMARTS.
- SMIRKS generated on the fly.
- Many thousands of different SMIRKS. (Only one bug found).
- Selection from Inventory and selection of reaction centre are the same ‘thing’.
- Design of SMARTS is a compromise between trying to select only hits which are ‘useful’, (minimising list size) and excluding structures which are inappropriate.

MODGRAPH

Enumeration Example

2 Step Reaction

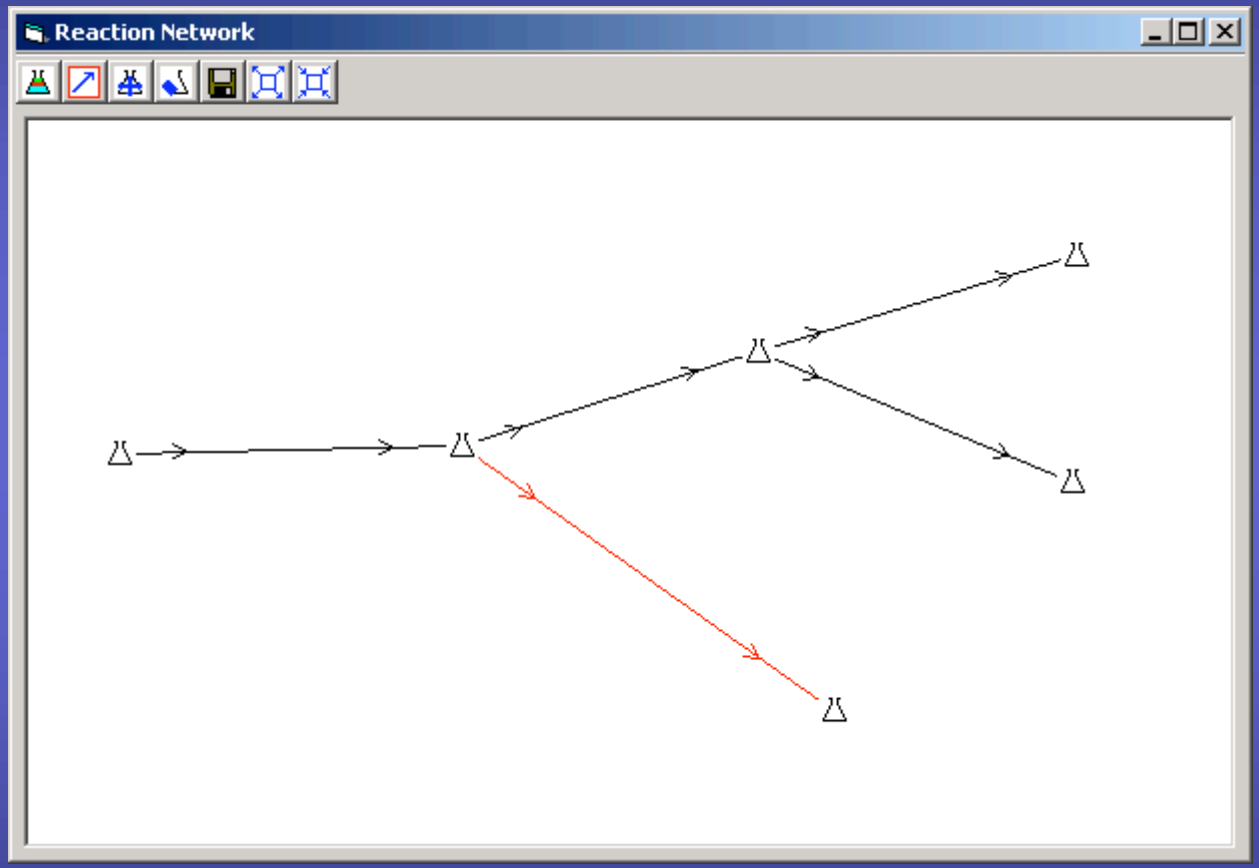
$$A + B >> C$$
$$C + D >> E$$

2 A x 3 B = 6 C x 6 D = 36 E

A1 A2	+	B1 B2 B3	>	C1 (A1B1) C2 (A1B2) C3 (A1B3) C4 (A2B1) C5 (A2B2) C6 (A2B3)	+	D1 D2 D3 D4 D5 D6	>	E1 (A1 B1 D1) E2 (A2 B1 D1) E3 (A1 B2 D1) E4 (A2 B2 D1) E5 (A1 B3 D1) ...

MODGRAPH

Network Design

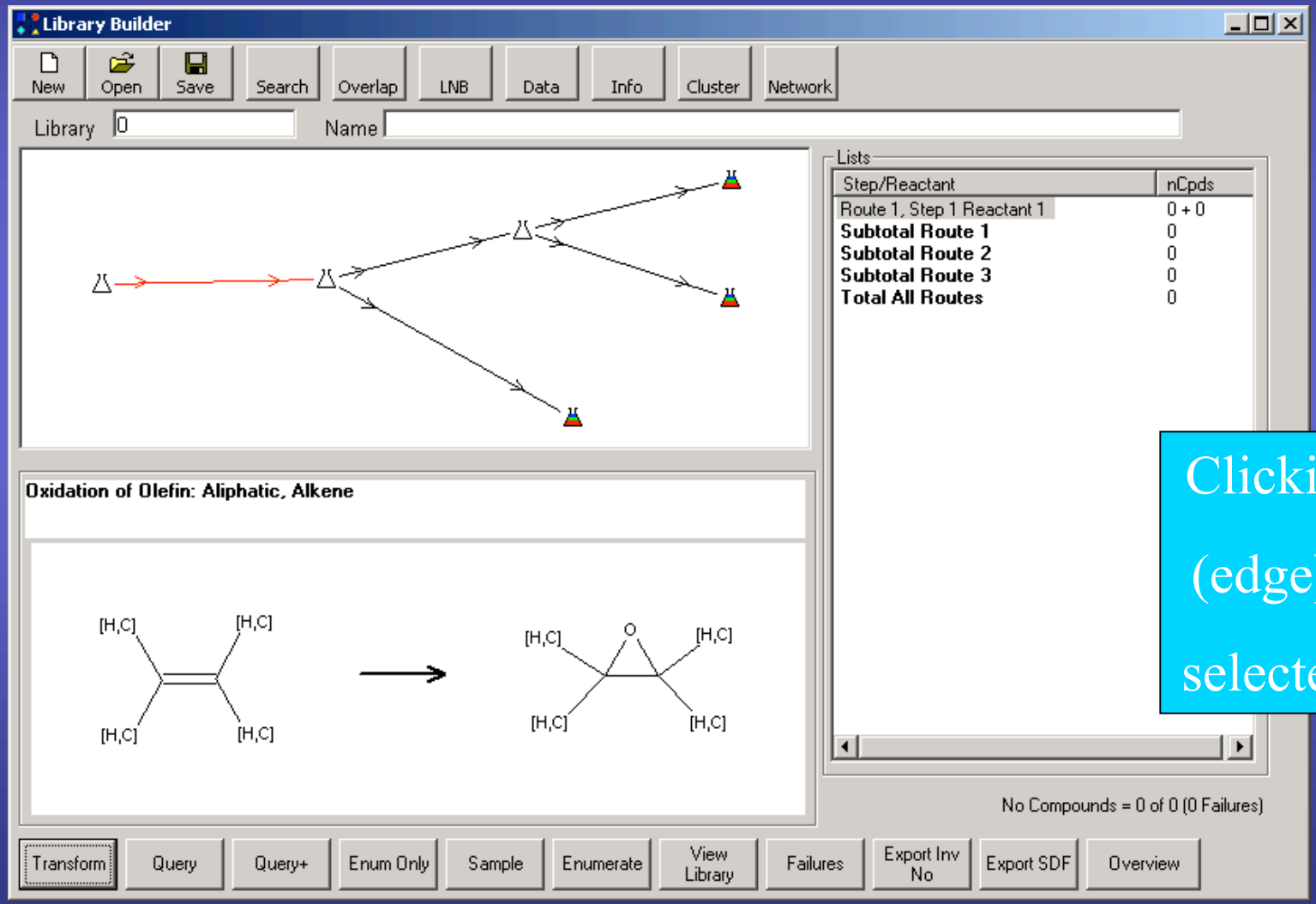


First design the reaction network.

Nodes can be added, deleted and moved, zoom-in, zoom-out. Rules.

MODGRAPH

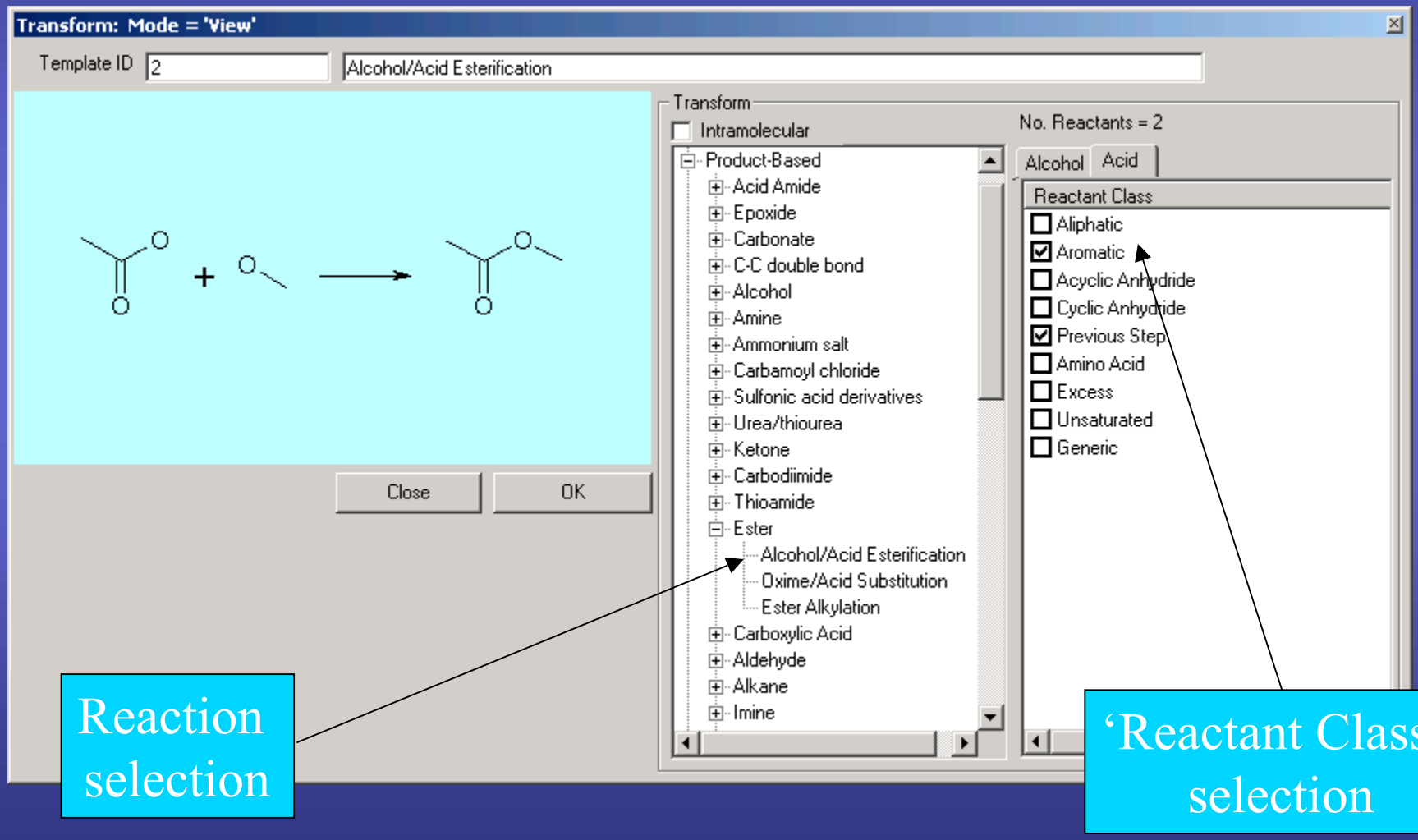
Template Approval



Clicking on step (edge) displays selected reaction

MODGRAPH

Reaction Selection

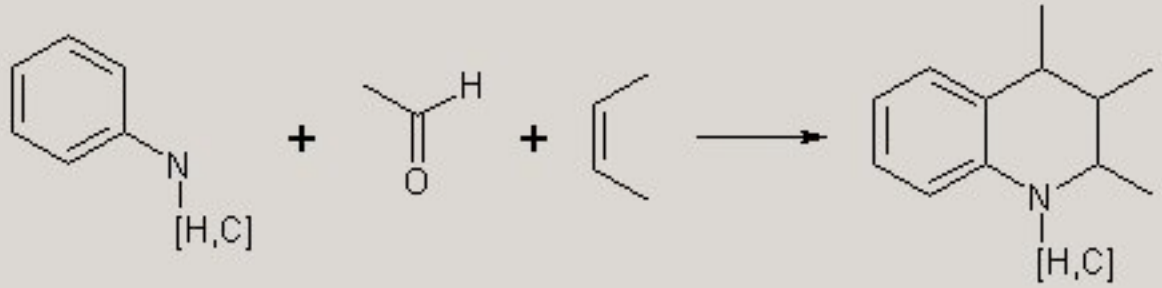


MODGRAPH

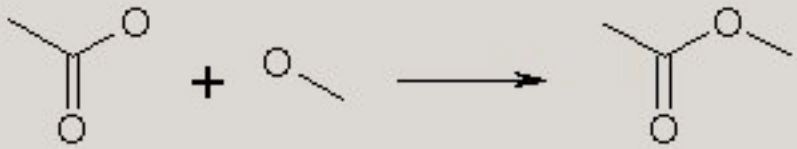
Reaction Overview

Overview: template = WX24V

Quinoline: Aliphatic, Aldehyde + Primary, Aromatic Amine + Aliphatic, Alkene



Alcohol/Acid Esterification: Primary, Alcohol + Aromatic, Previous Step, Acid



Overview can be printed

MODGRAPH

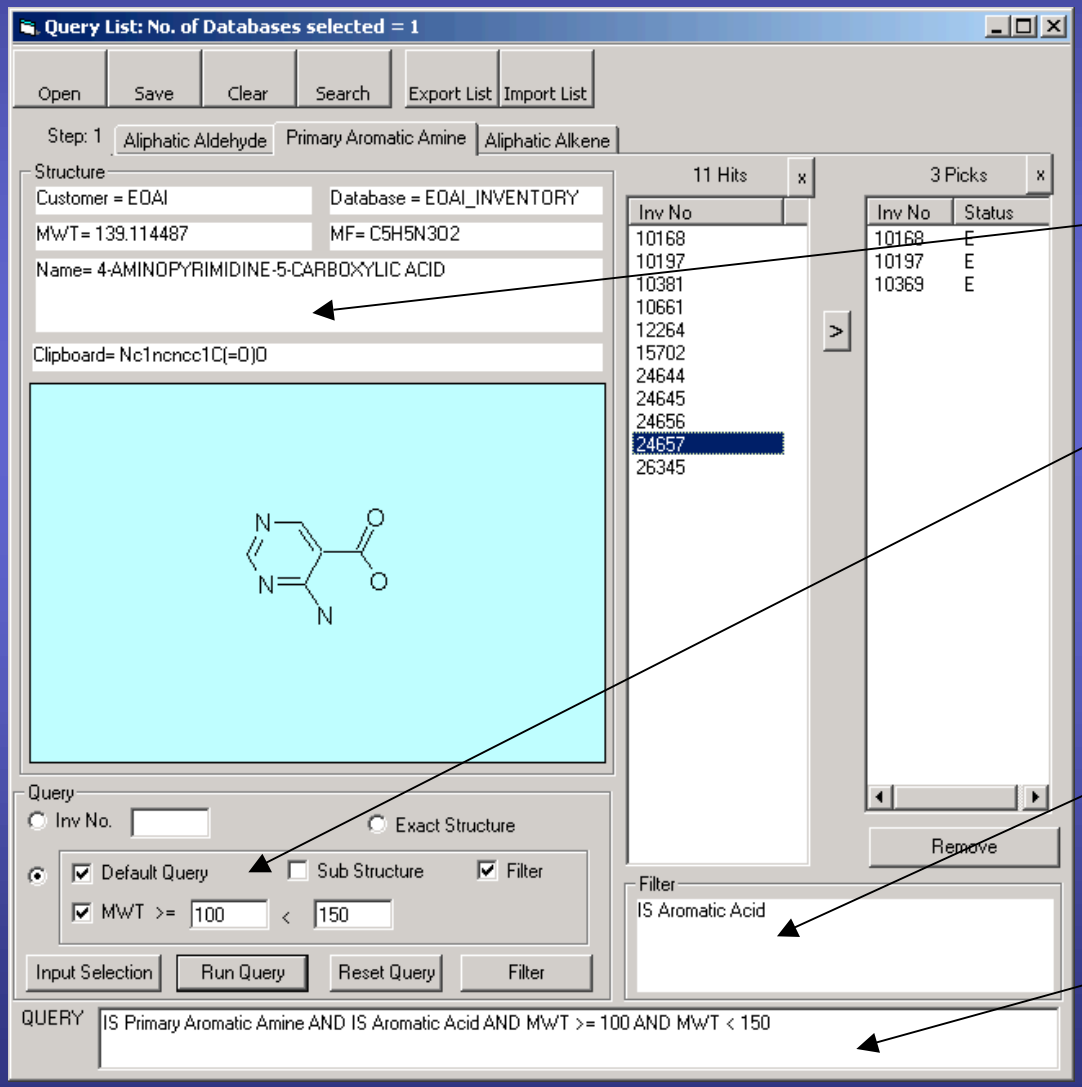
Reactant Selection

Search for reactants is based on either:

- Inventory number
- Exact structure
- Any Combination of
    1. Default SMARTS - the SMARTS based on reactant class selection.
    2. Sub-structure
    3. Filter - applies additional SMARTS based on reactant classes
    4. Molecular weight range.
- Two Stage Selection. Select list and select from list.

MODGRAPH

Reagent Selection - Stage 1.



Multiple, customer specific databases.

Reactant Structure and details.

Default Query

+ Filter

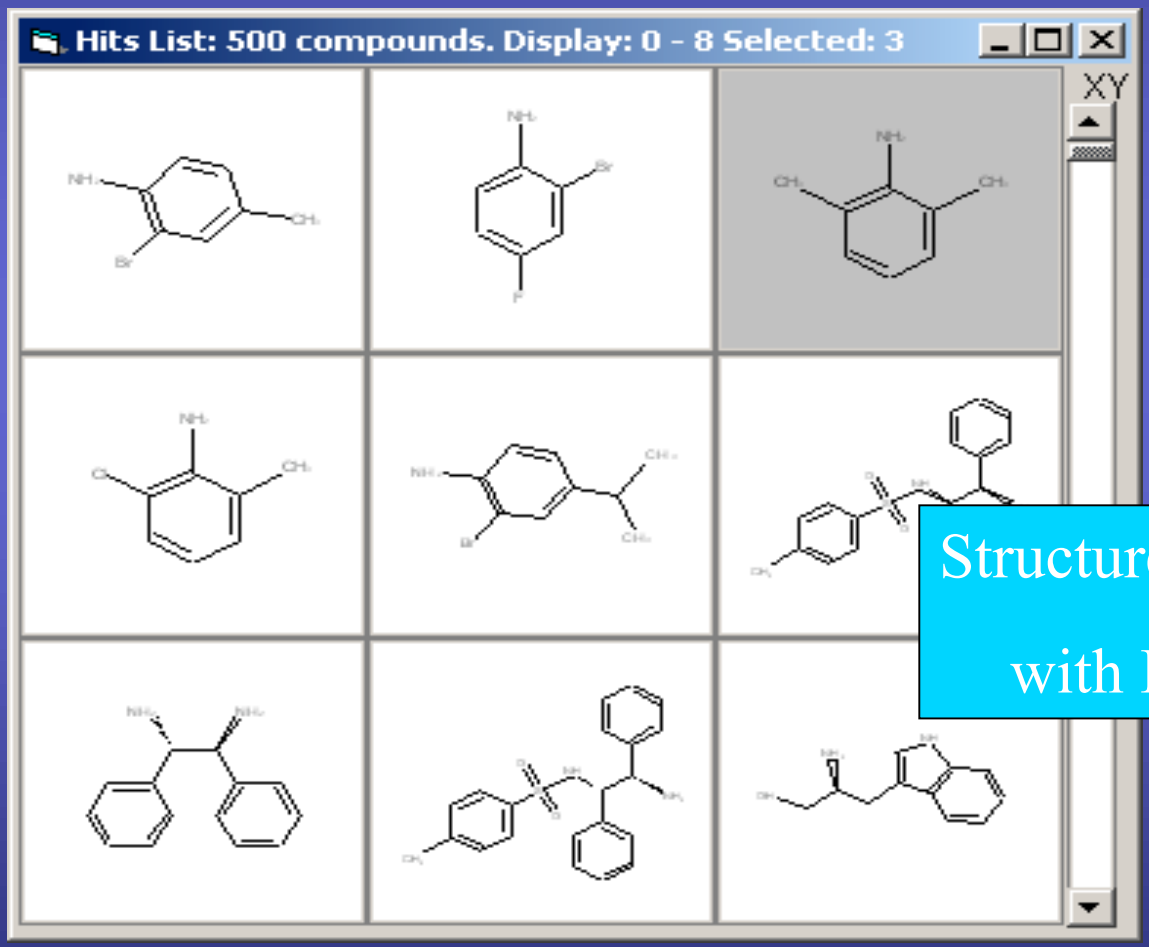
+ MW range

Filter selected by name/class.

English description of Query.

MODGRAPH

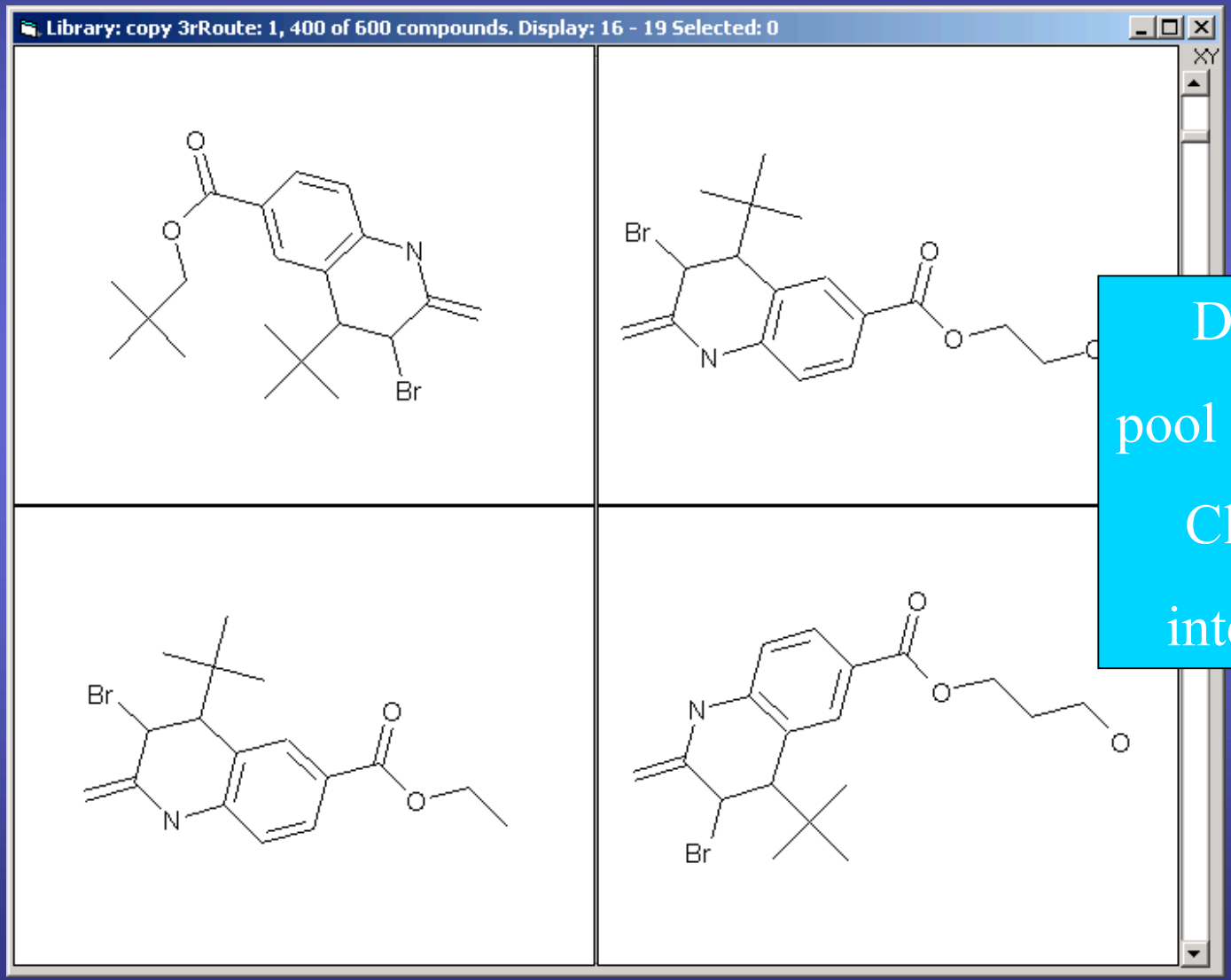
Reagent Selection - Stage 2



Structures are displayed with Inventory No.

MODGRAPH

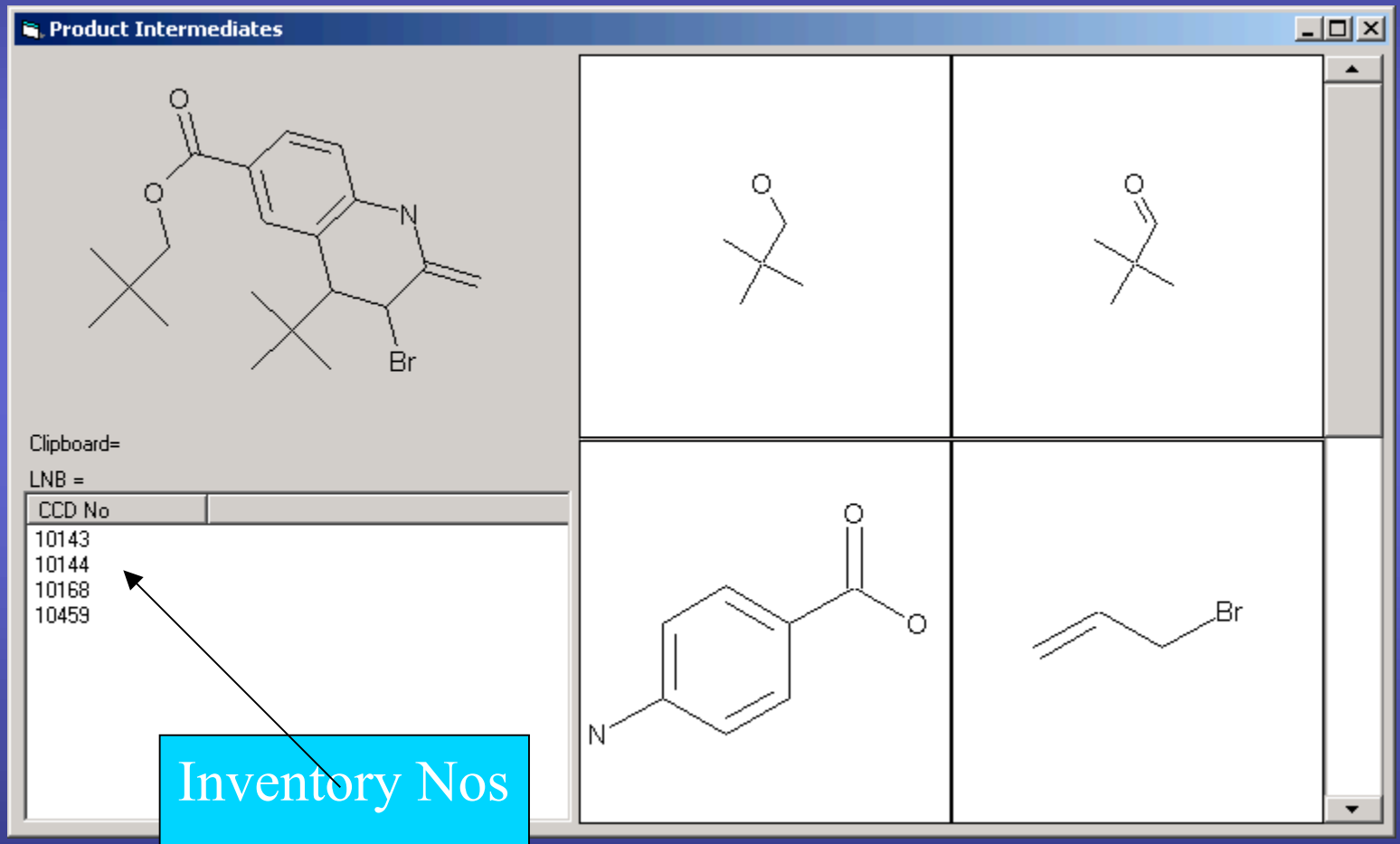
Library View



Display of pool compounds. Click to see intermediates

MODGRAPH

Product Intermediates



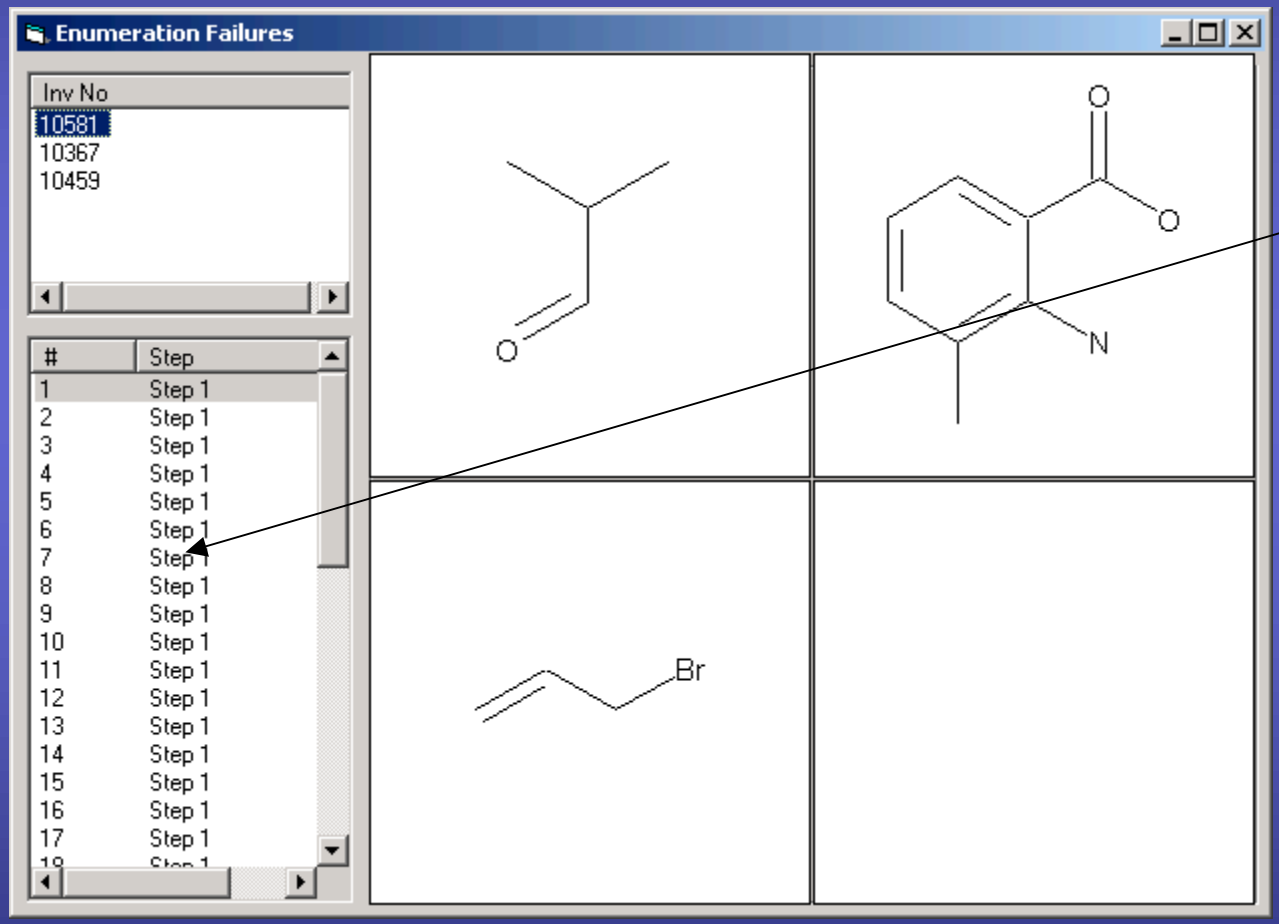
MODGRAPH

Enumeration Algorithm

- Fast and efficient enumeration of very large libraries.
- The library is split into a number of ‘experiments’, such that the number of compounds in one experiment is minimized, subject to no intermediate being synthesized more than once. **Minimum memory usage; maximum speed**.
- Minimising memory usage is more important, since writing to a db is likely to be RDS.
- Library Sample Algorithm - minimum library size consistent with every reactant being used at least once.

MODGRAPH

Failures List

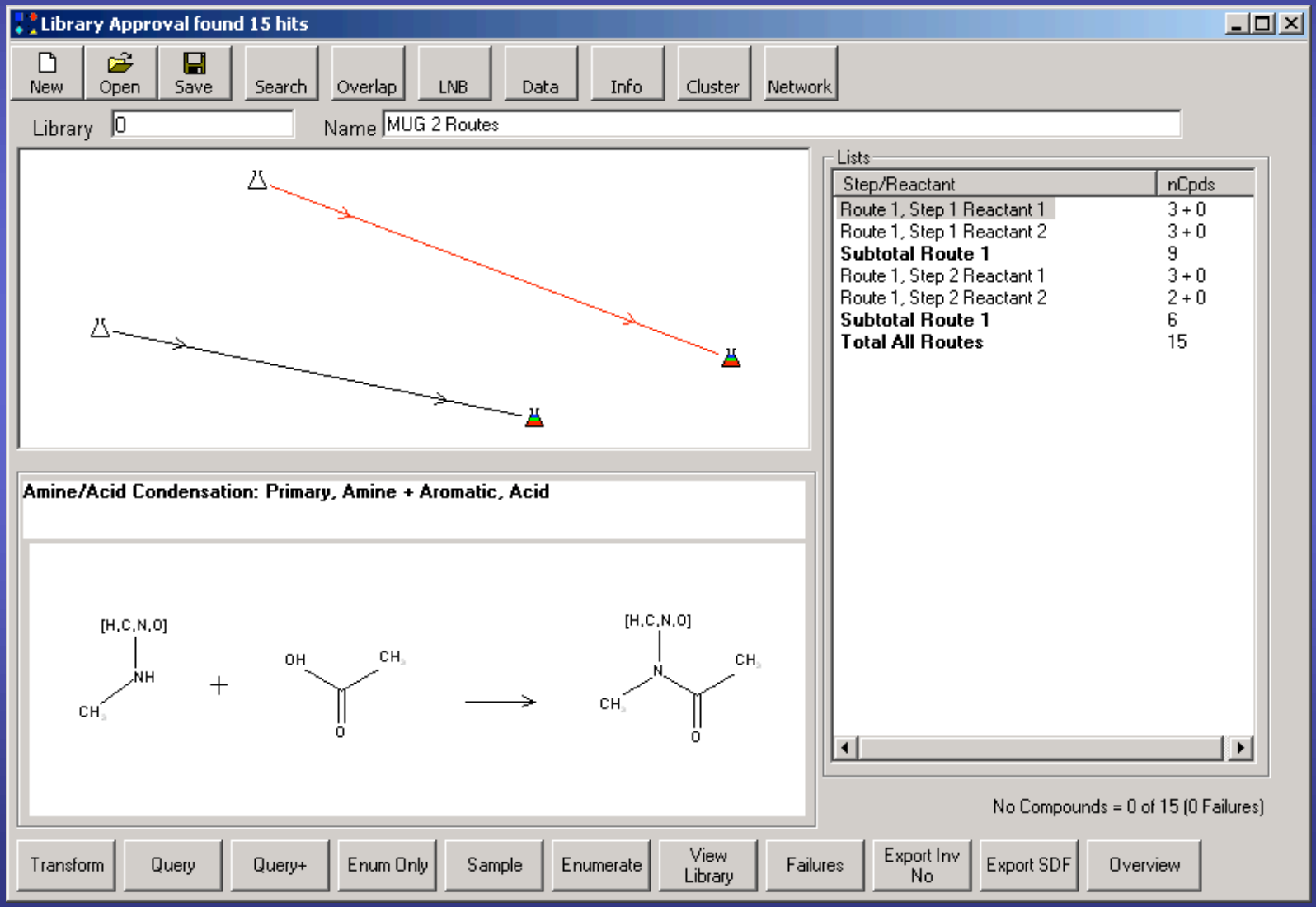


Shows the step no. which failed and reagents up to that step.

Here the failure is due to absence of aromatic H ortho to NH2

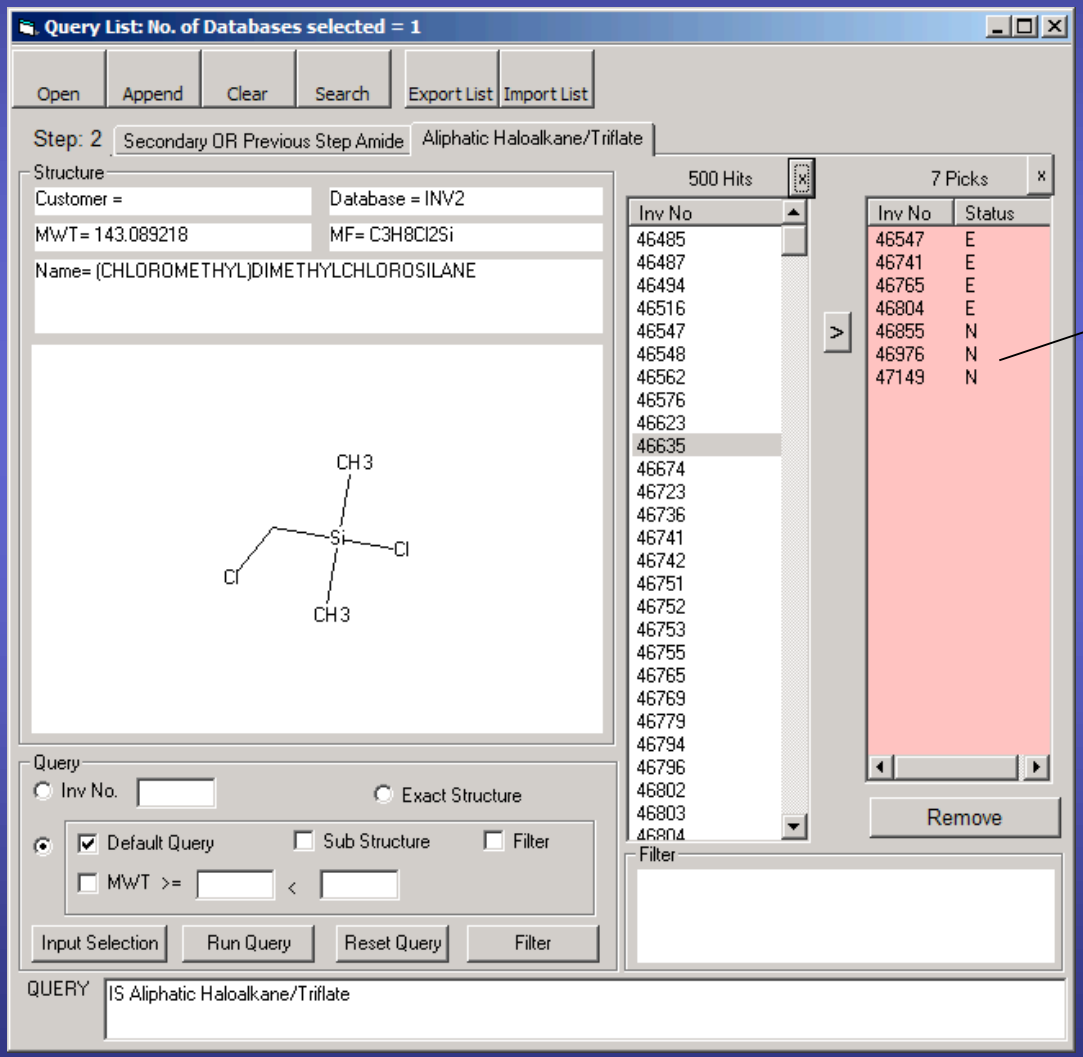
MODGRAPH

Routes can be Disjoint.



MODGRAPH

Query +

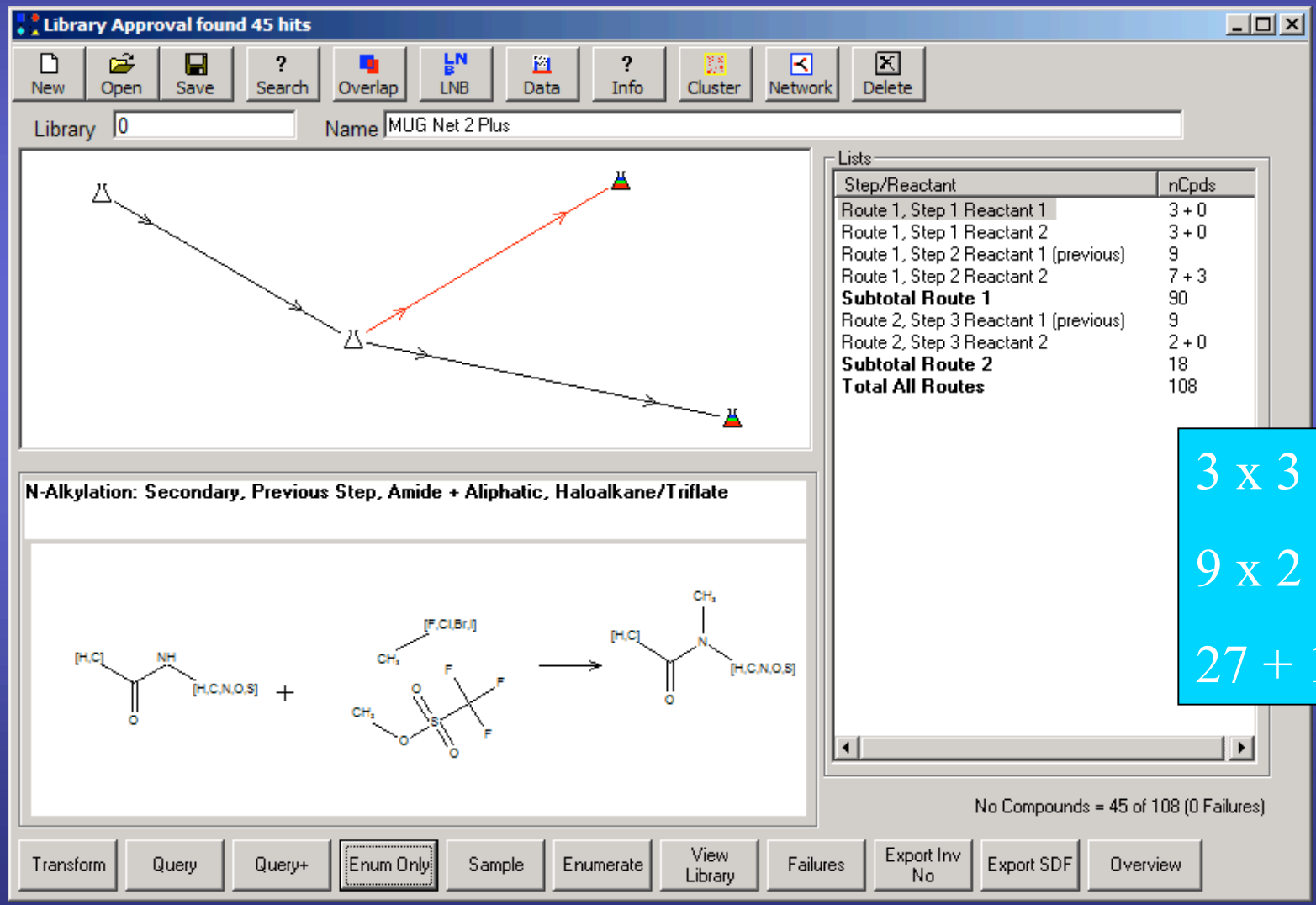


Three extra compounds added.

Avoid re-enumerating whole library. (Time, Data).

MODGRAPH

Query +



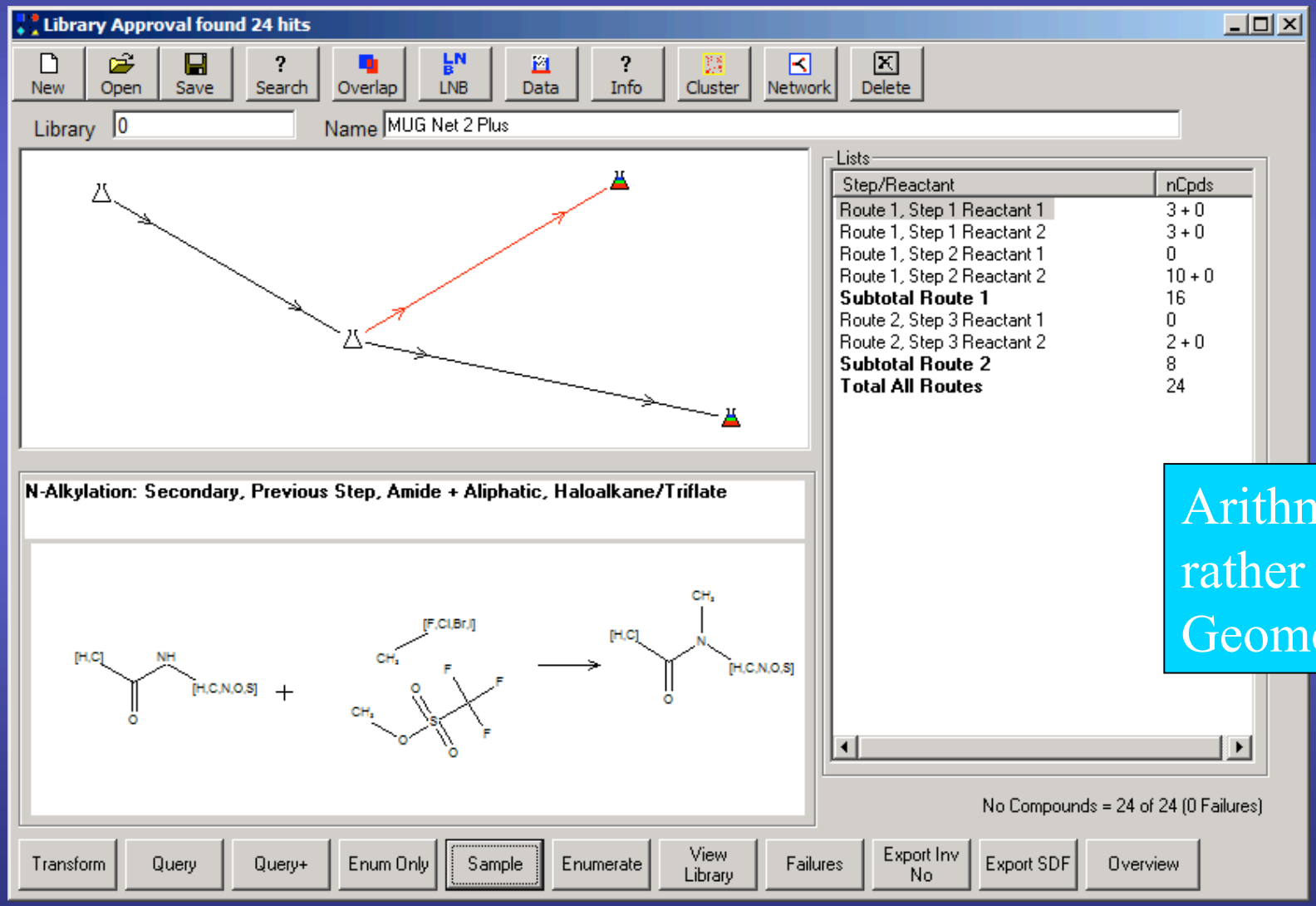
3 x 3 x 3 = 27

9 x 2 = 18

27 + 18 =45

MODGRAPH

Sampling



Arithmetic rather than Geometric.

MODGRAPH

Futures and Options

- Customisation. Integration with other applications/business process. Plate layout, library filtering, LNB assignment etc.
- Variable Sample size.
- Improvements to Reagent selection.
- Specification of Reaction Centre per reactant.

Acknowledgements

- John Bradshaw
- Jack Delaney
- Caroline McGee - formerly Evotec OAI