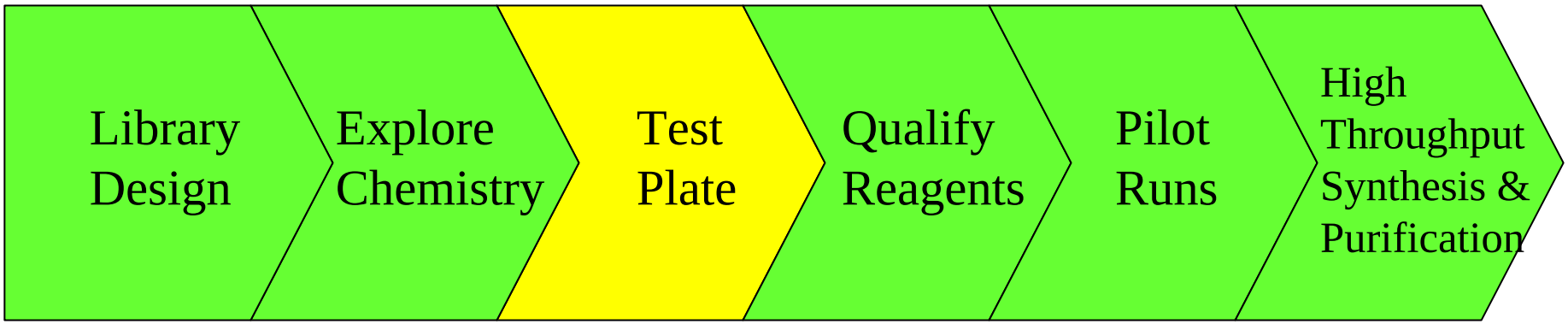

ArQule®

# Designing test plates with maximal information content and diversity for the development of library protocols

Jean E. Patterson, Ying Zhang, Andrew Smellie, Daming Li, David S. Hartsough, Libing Yu, and Carmen M. Baldino.

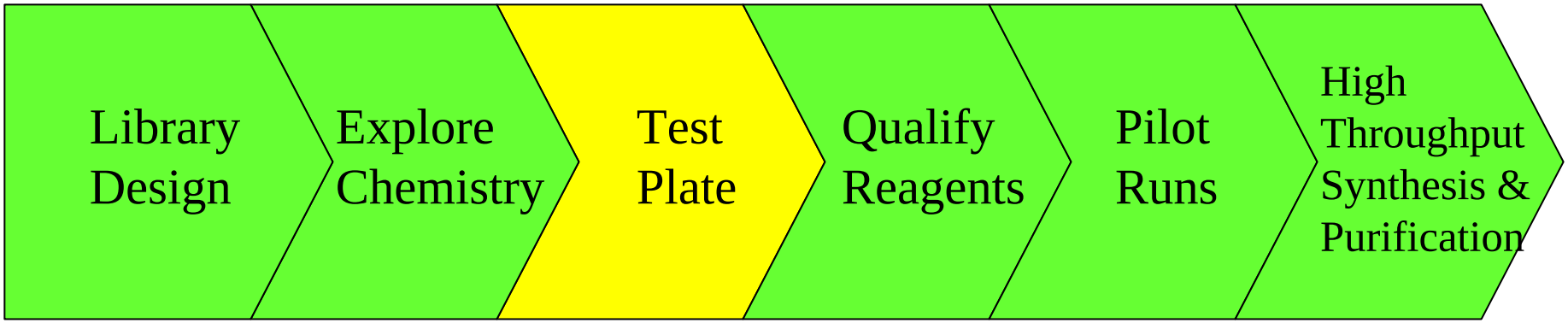
March 15, 2005 ACS Meeting

# Development Workflow for High-Throughput Chemistry Libraries



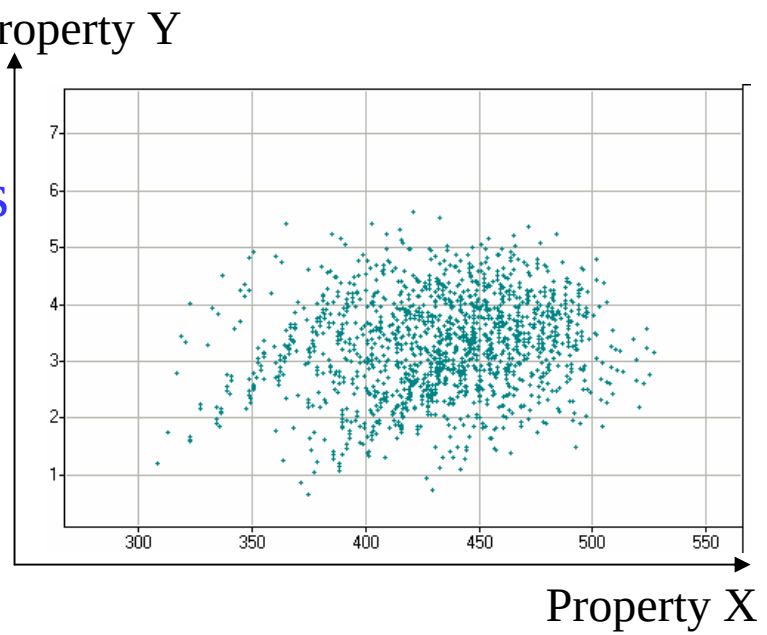

ArQule®

# The Challenge





Which test compounds predict?




*Initial Virtual Library*

*Production Library*

ArQule®

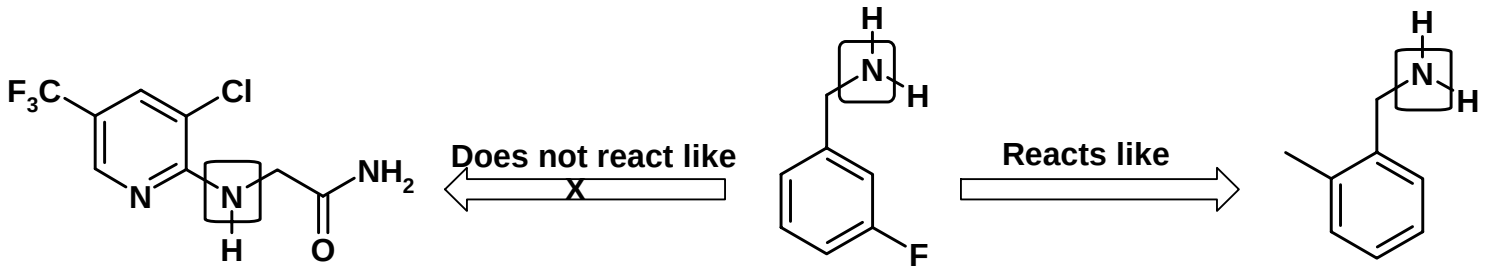
# How do I design a test plate?

- Need to predict how final library will behave
  - Production library overlaps with virtual library, but exact compound set is not yet known
  - Best estimate is a small, diverse set that represents large virtual library
  - Represents behavior of multiple, unrelated properties
    - Product LogD or LogP
    - Product Molecular Weight
    - Reagent Reactivity
    - Reagent and Product Solubility
    - Reagent Availability
  - Prior to complete procedure optimization

ArQule®

# Reagent reactivity can be described by “binning”

- Assumption:  Reagents with similar structural features will have similar reactivity




- Many approaches
  - Med chemists’ “eye”
  - Tanimoto coefficients
  - Daylight fingerprints
  - Many others…

ArQule®

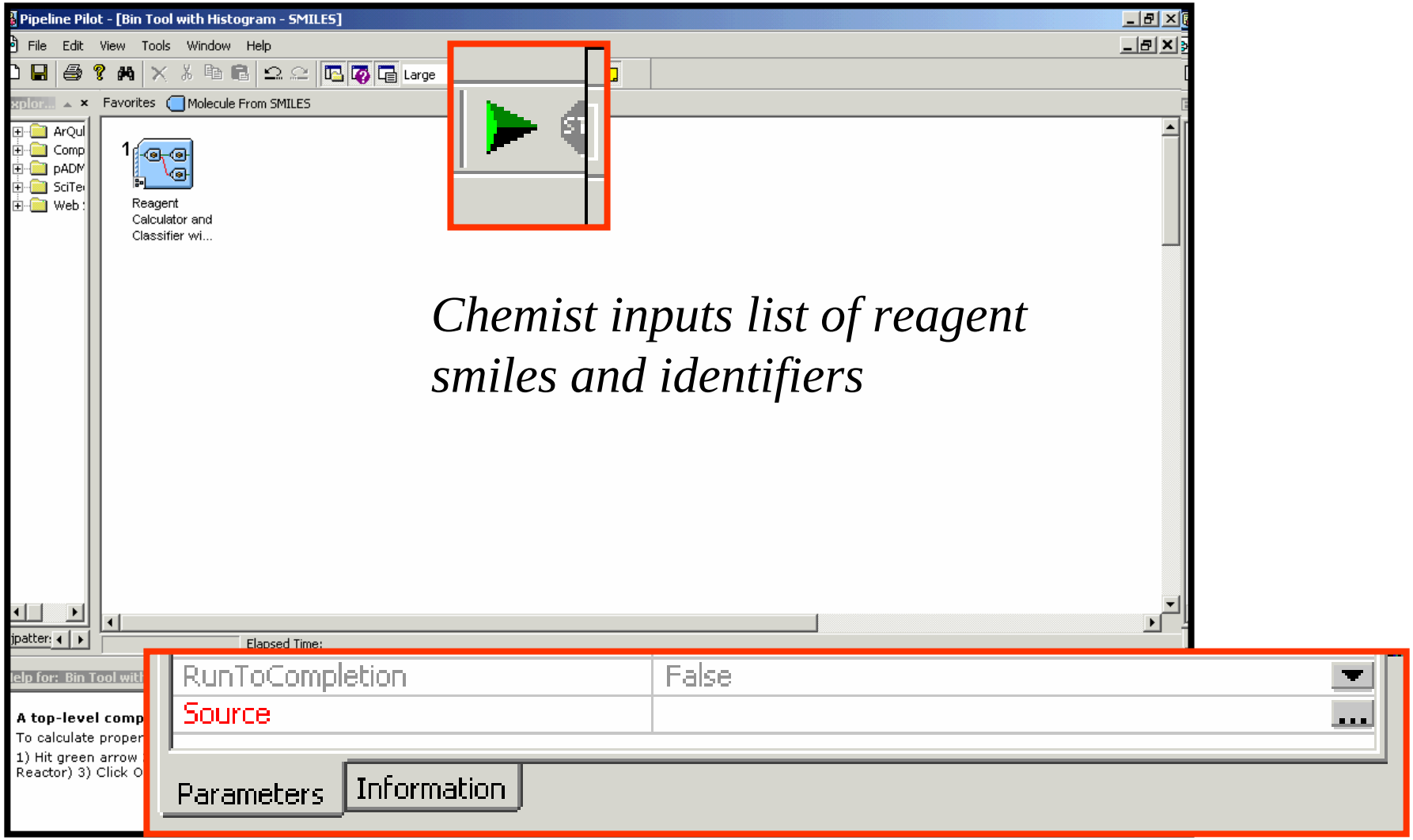
# Automated Binning Tool

- Pipeline Pilot Interface
- Common reactivity bins defined by chemists (SMARTS pattern matching)

  *Primary, secondary, aromatic, heteroaromatic, aliphatic, benzylic, cyclic, EDG, EWG, hindered, tertiary N, aminoalcohols, aminoacids, salts, racemic*
- Sorts reagents in structural order (minimum spanning tree algorithm)
- Excel output of structures, bins, and summary histogram

ArQule®

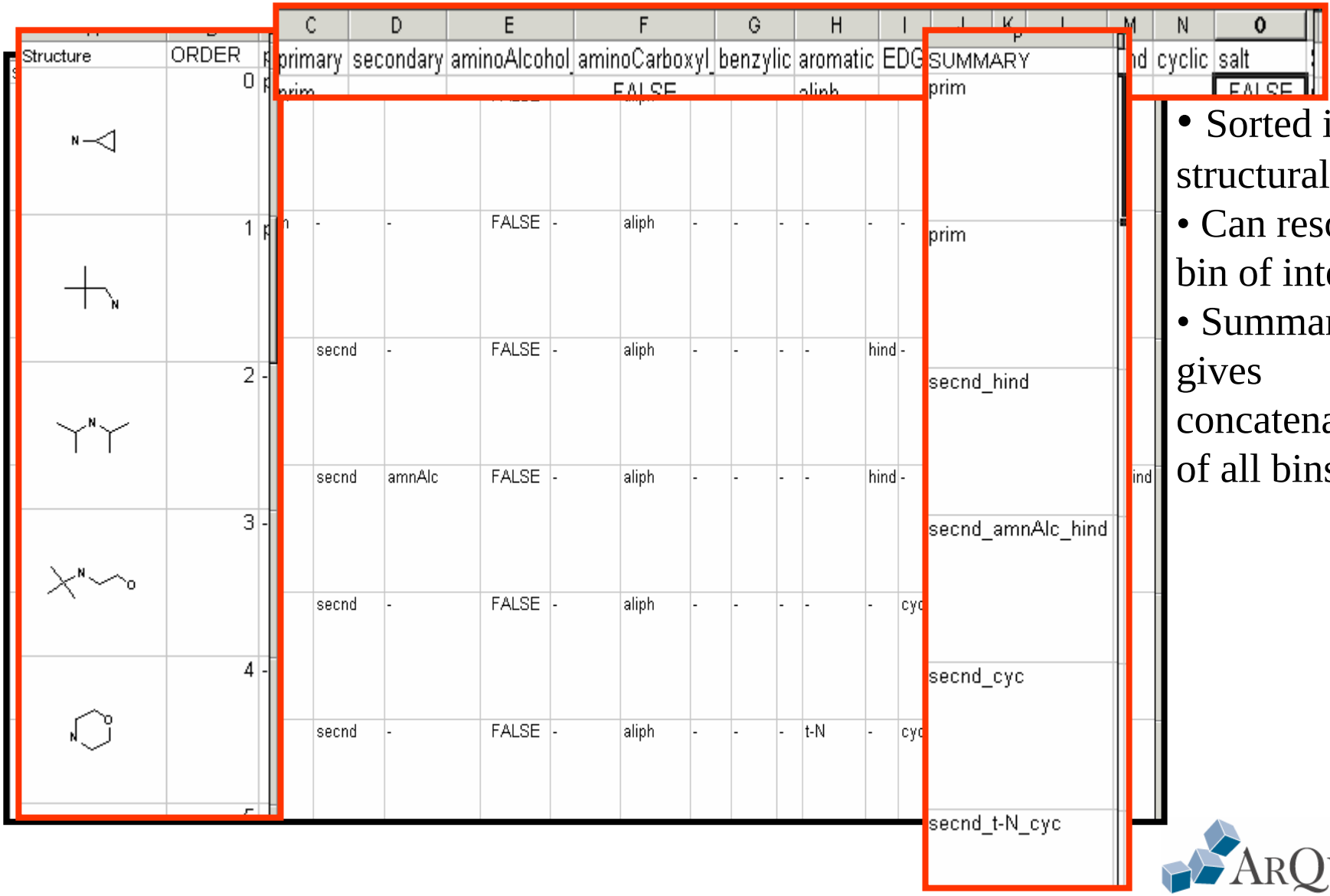
# User-Friendly Interface in Pipeline Pilot




*Chemist inputs list of reagent smiles and identifiers*

ArQule®

# Flexible Excel output



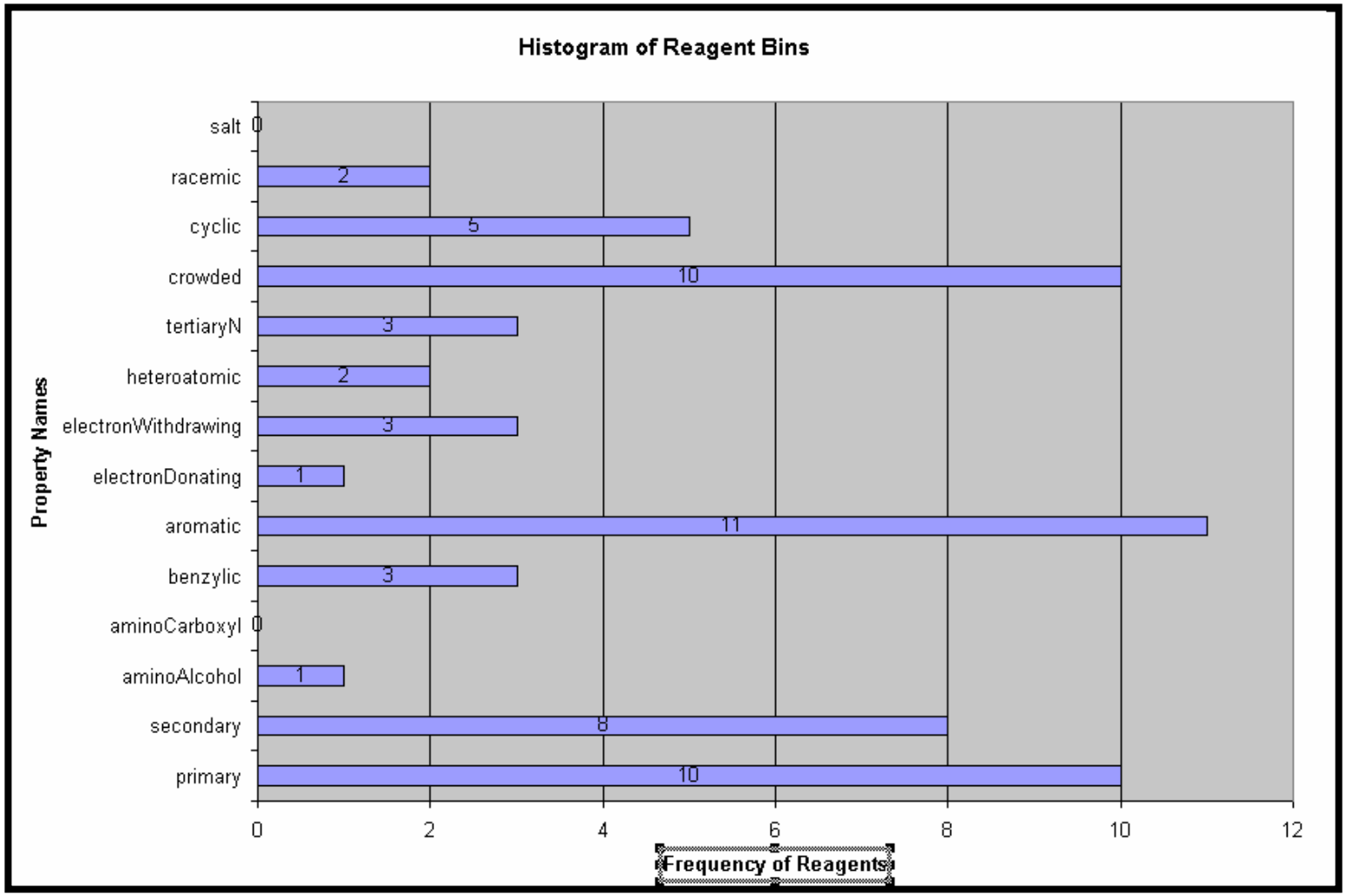
| Structure | ORDER | C primary | D secondary | E aminoAlcohol | F aminoCarboxyl | G benzylic | H aromatic | I EDG | SUMMARY | N cyclic | O salt |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | prim | | | FALSE | | aliph | | prim | | FALSE |
| | 1 | - | - | FALSE | - | aliph | - | - | - | - | - |
| | 2 | secnd | - | FALSE | - | aliph | - | - | hind | - | - |
| | 3 | secnd | amnAlc | FALSE | - | aliph | - | - | hind | - | - |
| | 4 | secnd | - | FALSE | - | aliph | - | - | - | cyc | - |
| | | secnd | - | FALSE | - | aliph | - | - | t-N | cyc | - |

| SUMMARY |
|---|
| prim |
| prim |
| secnd_hind |
| secnd_amnAlc_hind |
| secnd_cyc |
| secnd_t-N_cyc |

- Sorted in structural order
- Can resort by bin of interest
- Summary gives concatenation of all bins

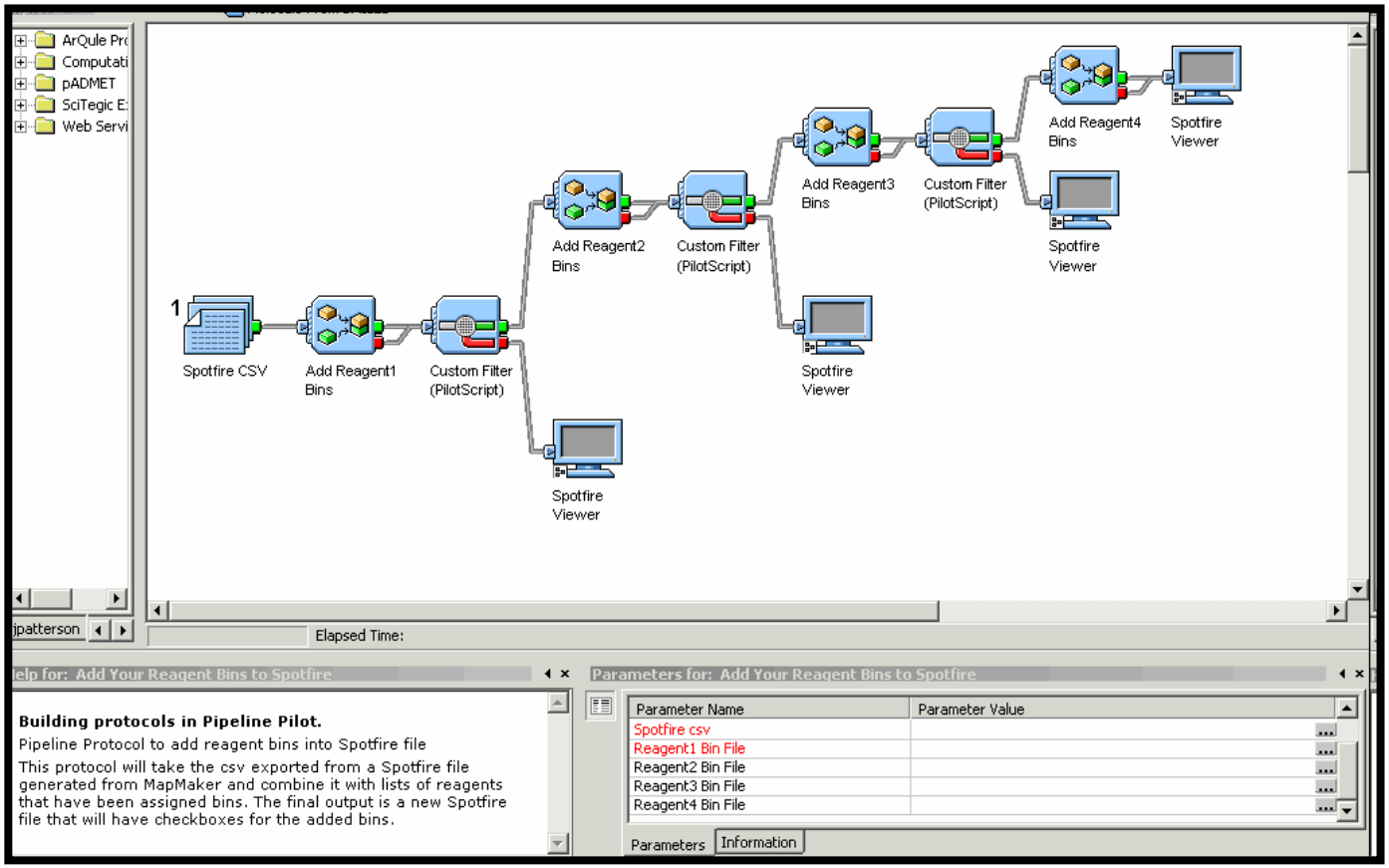
ArQule®

# Histogram output focuses development efforts



# Visualization of Bins

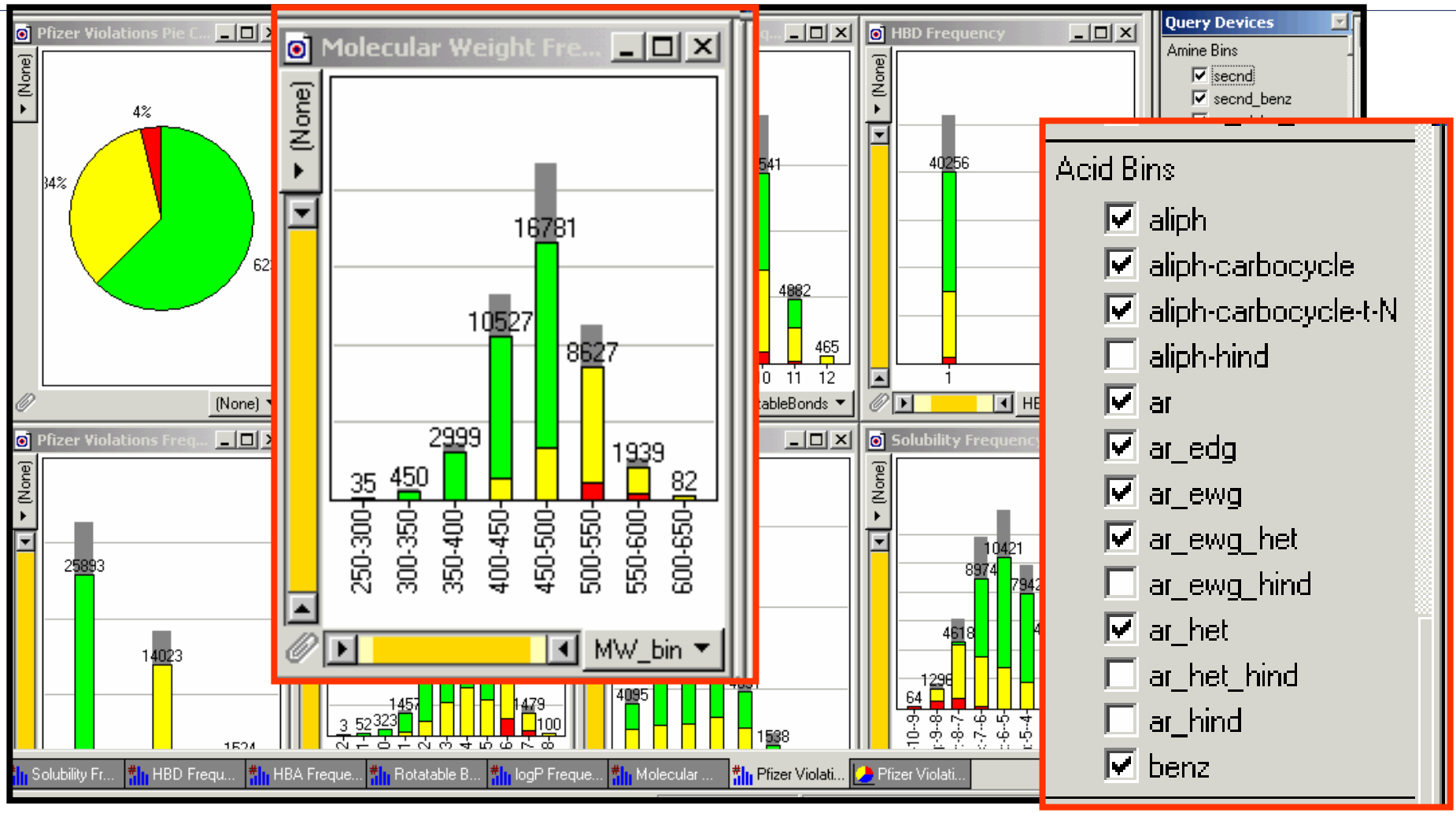
- Bins can be merged with other library design data for powerful, interactive visualizations




- Simple Pipeline Pilot protocol
- Merge library design data (automatically generated from MapMaker™) and data from automated bin tool

ArQule®

# Spotfire Bin Visualization with Library Design Data




- *Ability to interactively deselect reagent bins facilitates development*
- *Reagent reactivity, logP, and MW data – how to get others?*

ArQule®

# Many properties considered in design criteria

- Product LogP – library design (MapMaker™)
- Product Molecular Weight – library design (MapMaker™)
- Reagent Reactivity – bin (automated PP tool)
- Product LogD
  - Related to logP, but varies with pH
  - Correlates to retention time in high-throughput purification
  - Calculated property
- Reagent and Product Solubility - experimental
- Reagent Availability – what's on the chemists' bench (Reactor™)

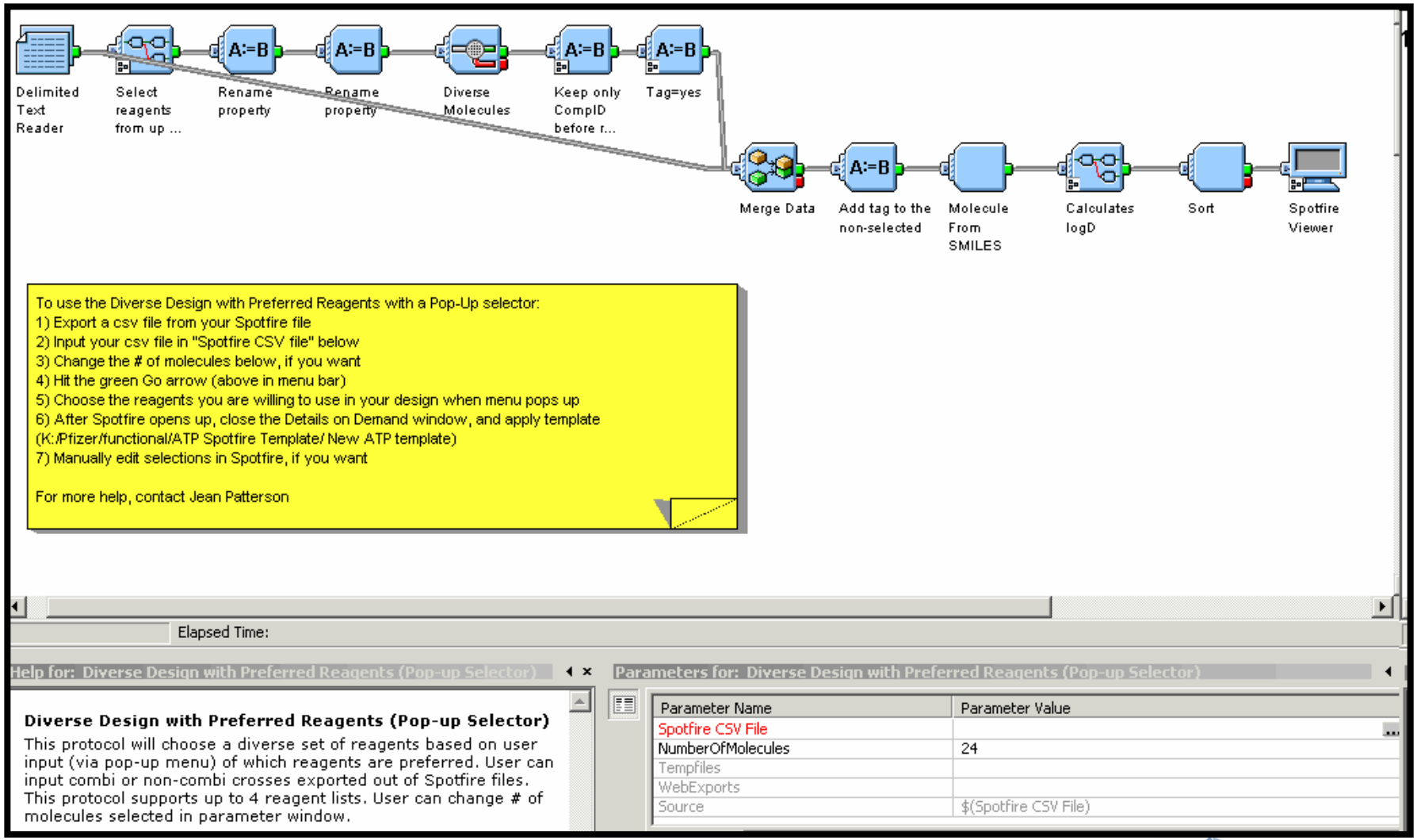
ArQule®

# Design Strategy

- Automate diversity with respect to calculated properties
  - LogP and Molecular Weight
  - LogD – should be diverse if logP is diverse
- Let chemists influence design with respect to non-calculated properties
  - Reagent Reactivity (guided by bins)
  - Reagent and Product Solubility
  - Reagent Availability
- Display design diversity and allow for easy modifications by users

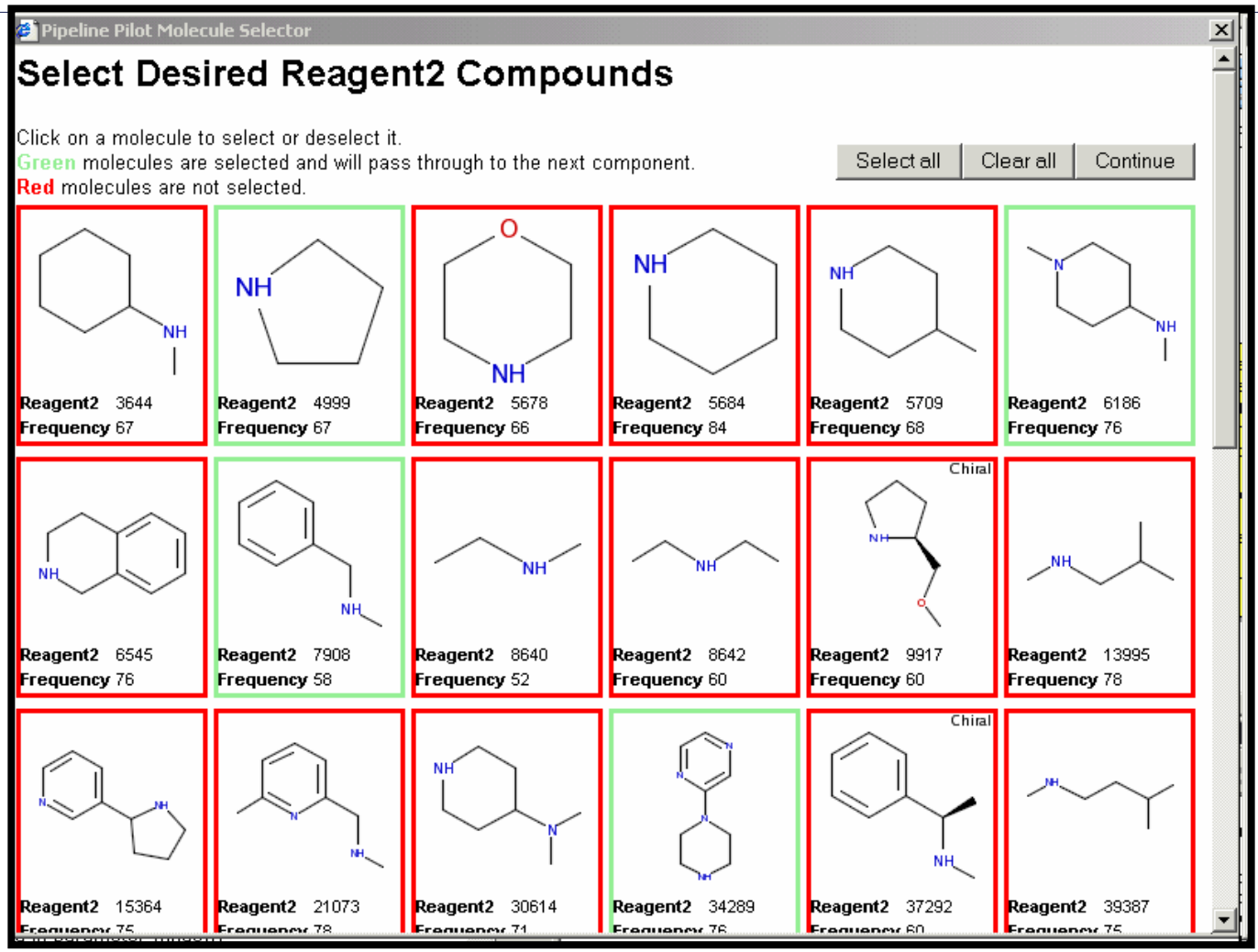
ArQule®

# Test Plate Design Tool




ArQule®

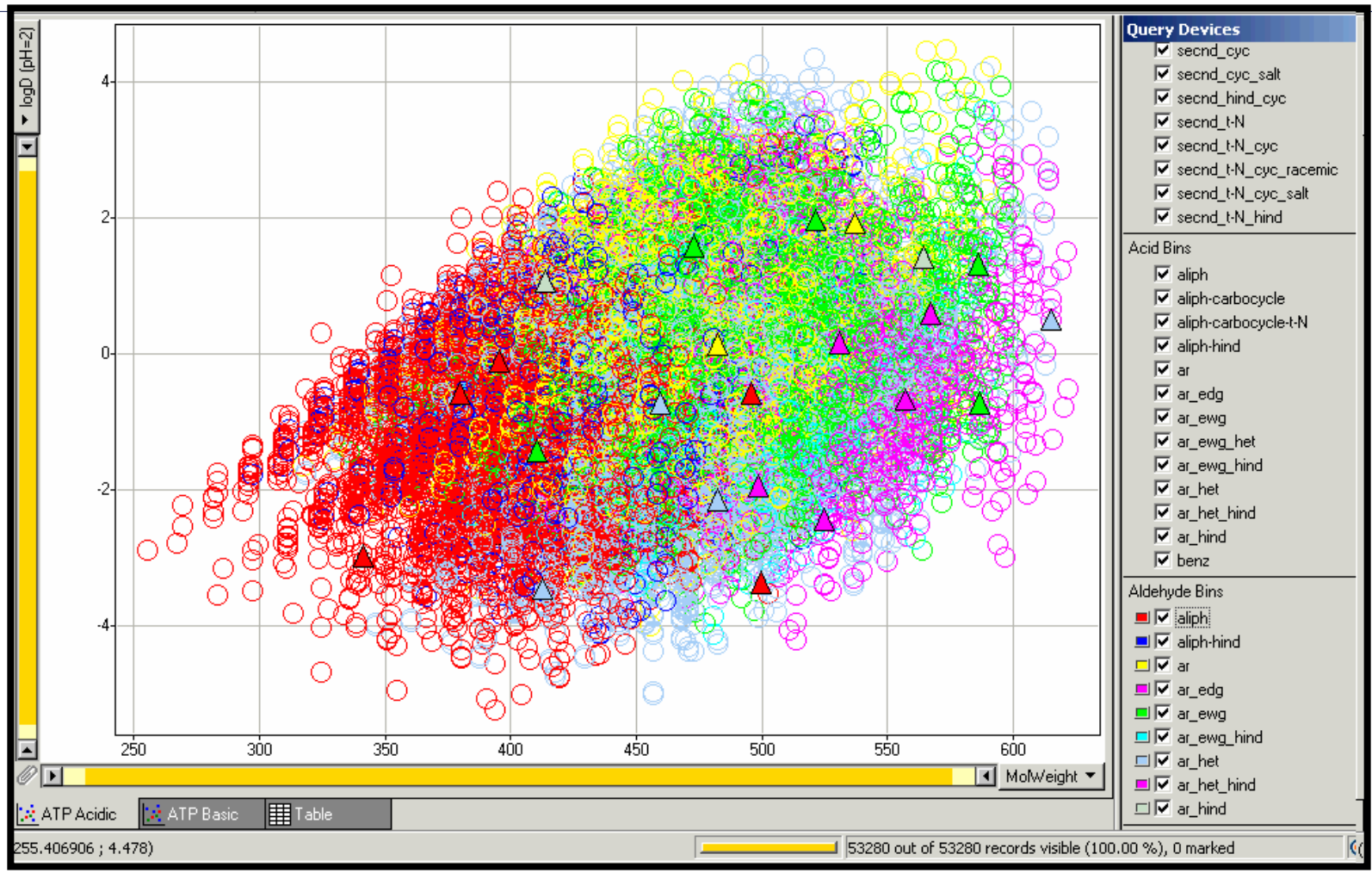
# Selection Window allows users to influence design




- *Users eliminate reagents unwilling to use in design*
- *Next, diverse set chosen wrt logP and MW*

ArQule®

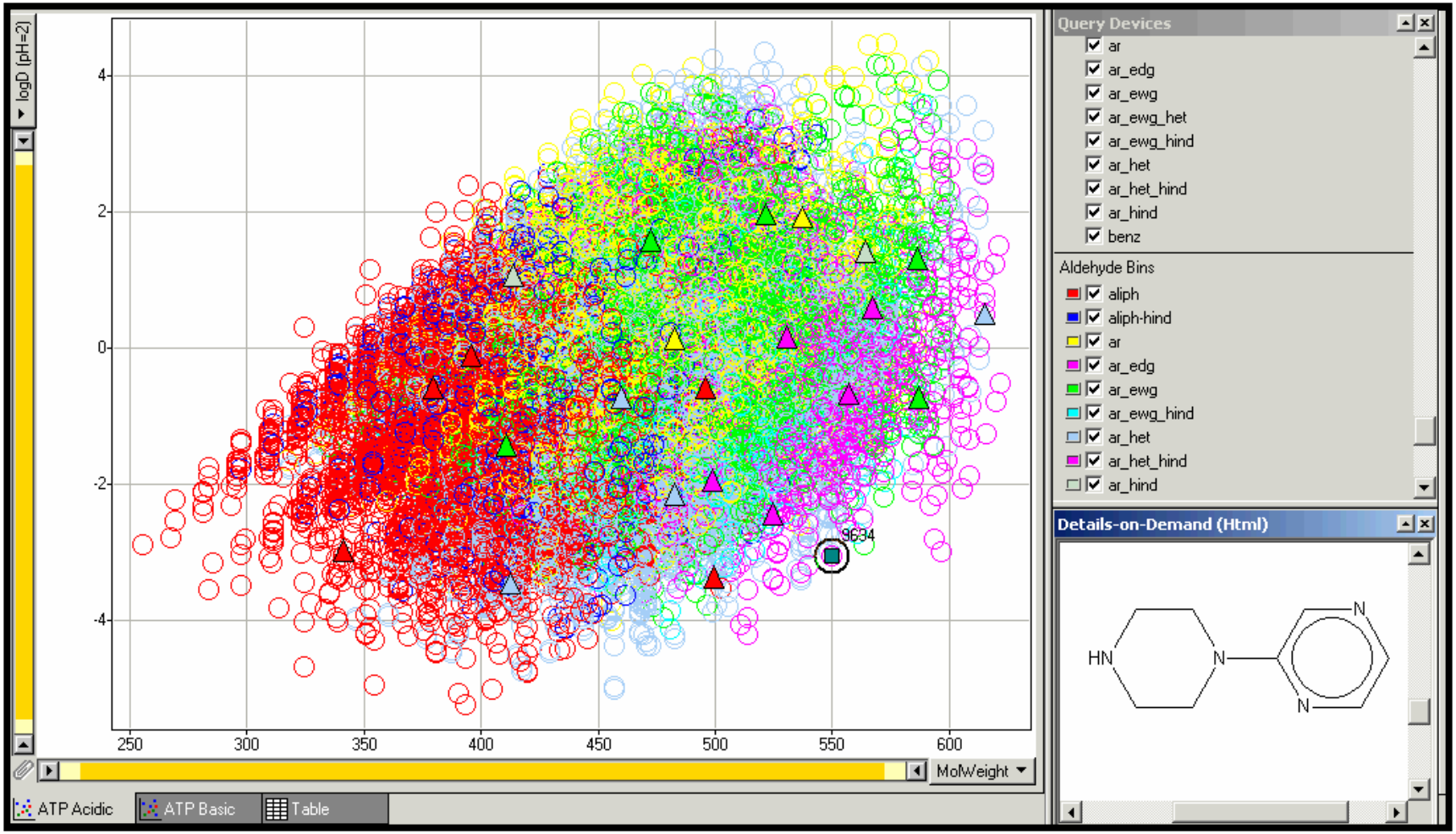
# Diversity can be viewed in Spotfire




- *Selected test compounds are triangles against all virtual library*
- *Scatter plot toggles between properties – logD, logP, MW*
- *Color by reagent bins*

ArQule®

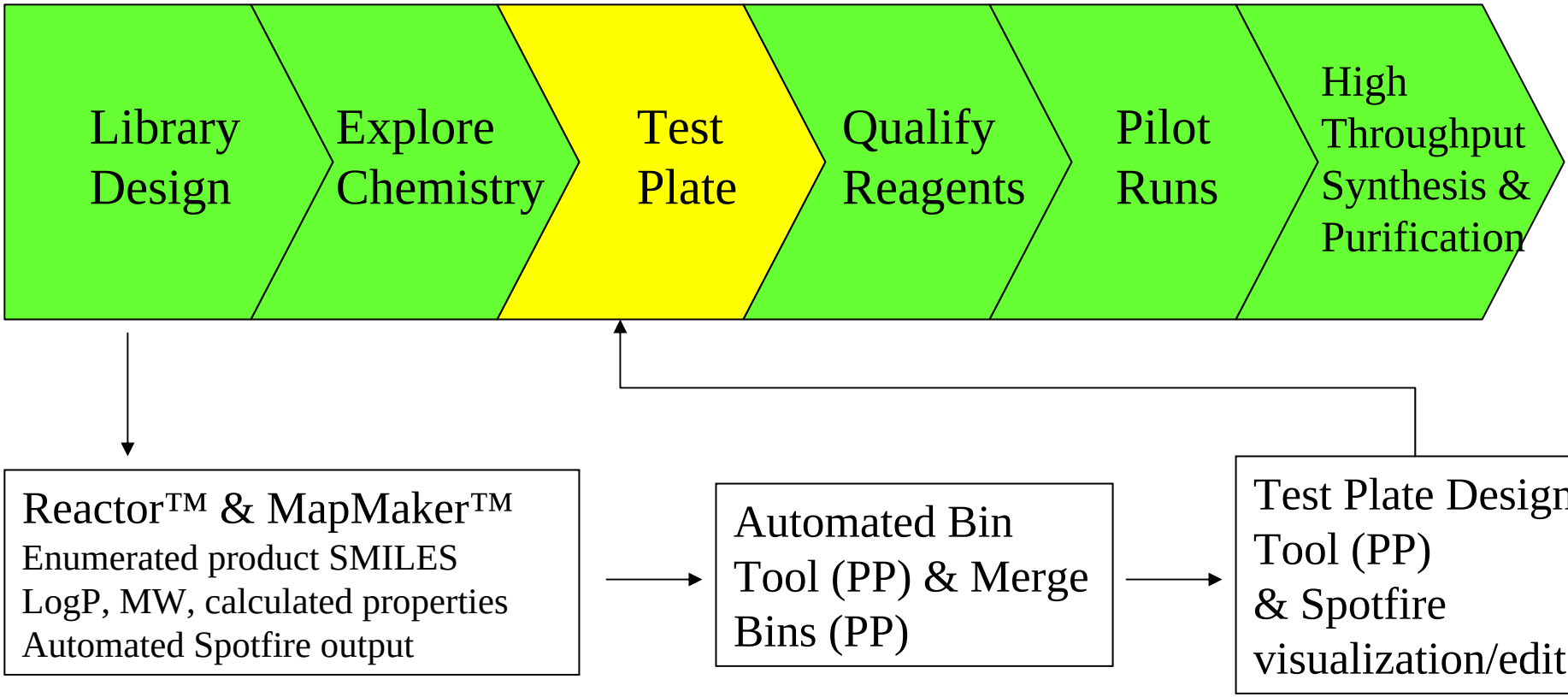
# Design can be modified easily in Spotfire




- *User can mark new records in the appropriate space*
- *Reagent/product structures viewable in Details-on-Demand window*

ArQule®

# Test Plate Design Workflow




ArQule®

# Conclusion

- Test plate design is crucial to the successful development of production-friendly high-throughput chemistry
- Test plate design is a challenging balance of many, non-orthogonal design constraints
- A flexible design process has been developed that involves library design data, reagent binning, Pipeline Pilot protocols, and Spotfire visualizations

ArQule®

# Acknowledgements


- Rob Onorato
- Sergio Rotstein
- ArQule chemistry and analytical departments

# Abstract


- Designing test plates with maximal information content and diversity for the development of library protocols Jean E. Patterson1, Ying Zhang1, Andrew Smellie2, Daming Li, David S. Hartsough3, Libing Yu2, and Carmen M. Baldino4. (1) Department of Chemistry, ArQule, Inc, 19 Presidential Way, Woburn, MA 01801, (2) ArQule Inc, 19 Presidential Way, Woburn, MA 01801, (3) Informatics and Modeling, ArQule, Inc, 19 Presidential Way, Woburn, MA 08101, (4) Chemistry Department, Arqule Inc, 19 presidential way, woburn, MA 01801
- The design of test plates used to develop a parallel synthesis library often has multiple, competing constraints such as a wide range of physico-chemical properties (molecular weights, logP, logD, solubility, HPLC retention times, etc.), structural diversity, and practical limitations such as reagent availability or reactivity that need to be taken into account. At ArQule, we have developed a test plate design process that involves reagent binning, automated Pipeline Pilot protocols that allow chemists maximum flexibility in choosing preferred reagents, and standardized Spotfire visualization tools that enable chemists to interactively view the design diversity and other information with respect to the virtual library.



- Informatics and High Throughput Experimentation
- 11:20 AM-11:45 AM, Tuesday, March 15, 2005 Unknown Location -- Unknown Room, Oral
- Division of Chemical Information

- The 229th ACS National Meeting, in San Diego, CA, March 13-17, 2005


ArQule®