

# Application integration: Providing coherent drug discovery solutions

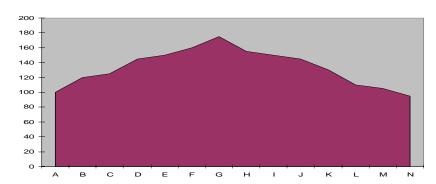
Mitch Miller, Manish Sud, LION bioscience, American Chemical Society 22 August 2002



- Introduction: exploring 'application integration'
- Our integration platform
- Compound prioritization one example
- Information traversal another example

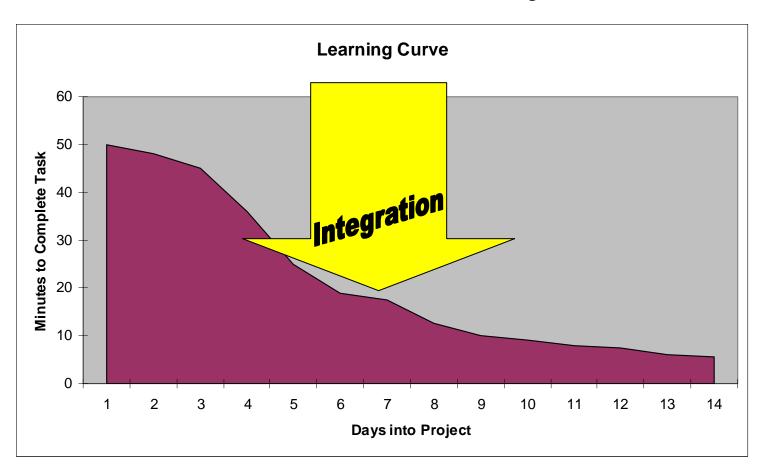


- Application Integration
  - Great idea!
  - What does it mean?
- 'The act or process of integrating'
  - 'To make into a whole by bringing all parts together; unify'
- Mathematical
  - 'A number computed by a limiting process in which the domain of a function, often an interval or planar region, is divided into arbitrarily small units, the value of the function at a point in each unit is multiplied by the linear or areal measurement of that unit, and all such products are summed. '
  - The area under a curve!





- The area under the information-access learning curve!
  - Learning curves go down, according to Gary Anthes in ComputerWorld, July 2, 2001
  - We want to reduce the area under the learning curve





### Application integration:

A situation in which different applications share data and user interface elements, allowing the user to move quickly and efficiently between tasks and *work more productively*.



### What are we integrating?

- Vertical applications:
  - Chemical database searching
  - Compound selection
  - Compound analysis
- Horizontal integration:
  - Assay data browsing and searching
  - Relationship browsing



### **Introducing LION DiscoveryCenter**

- Integration framework for compound discovery
  - Supports all the applications listed
- Guiding principal: 'Show me everything we know about X'
  - uniting information about drug-discovery topics
  - variety of sources public and private
  - provide access to specialist tools
  - capture knowledge of individual scientists
  - jump to related Xs
- Bioinformatics products released last year
- Chemistry: from vision in April to functioning prototype
  - release scheduled for early 2003



- Summary sheet key to 'show me everything'
  - variety of information on a single topic in a single screen
  - described by XML -- easy to extend and customize
  - addressed by URL -- can be transmitted
- Session paradigm:
  - search -> browse -> drill down to summary sheets -> traverse
- 'Favorites' provide pointers to individual summary sheets
  - find out what's happening to the entities you're interested in
- Links between different types of data
  - build on natural relationships between items



- Annotations capture user knowledge
  - pertain to specific sections or to entire entities
  - shared or private
- Follow-up searching integrated into content display
- Fully-functional Java client
  - includes IceStorm browser to render Web content
  - domain-specific viewers for scientific content
- Component-based easy to extend

10



### **Change in Session Paradigm for Chemistry**

- Search -> browse -> analyze -> drill down to summary sheets -> traverse
  - Use computational services to enhance data from databases
  - Possibly load set of compounds from disk file to initiate analysis
- LDC has compute services API to facilitate hooking in computational tools to better understand our compounds
- Examine one case in point:
  - Compound Prioritization

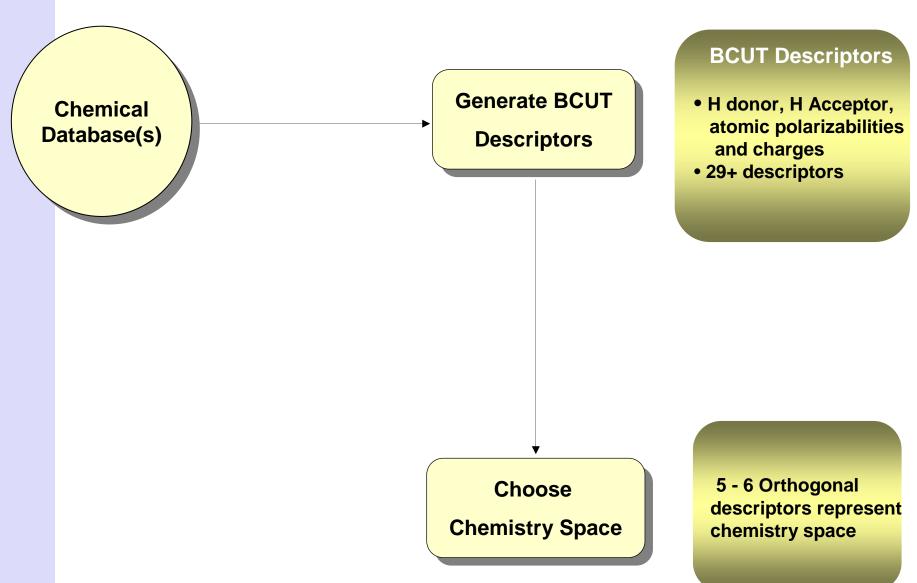


### **Compound Prioritization: An Overview**

- An application currently in use at LION, predating LDC
- A means of assigning a score to a compound
  - how useful will this compound be for us?
- Based on existing models for biological activity and ADME
  - Biological activity models use Receptor Relevant ChemSpaces, generated using DiverseSolutions from R. S. Pearlman and K. M. Smith, University of Texas, distributed by Tripos
  - Predictive ADME using in-house models

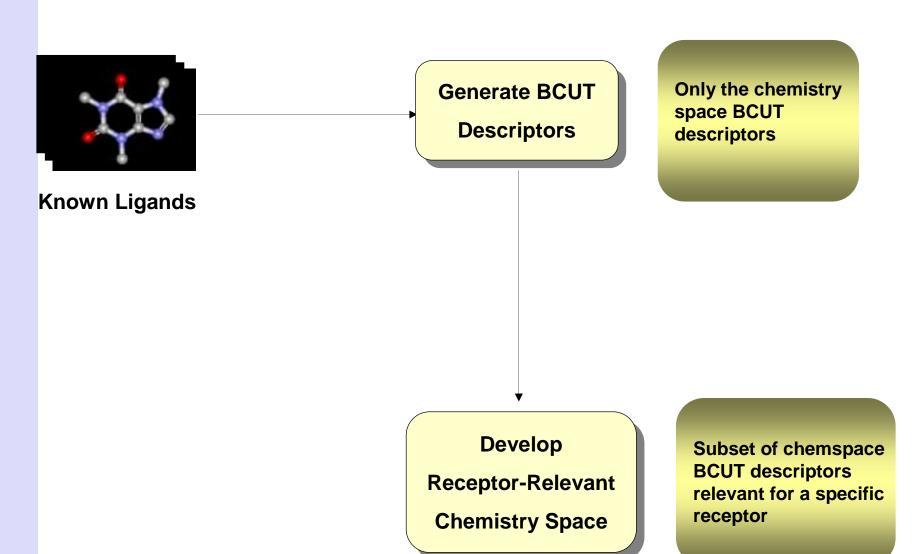


# Receptor-Relevant Chemspaces Part I: Database Chemspace





# Receptor-Relevant Chemspaces Part II: Receptor-specific refinement





- We have generated Receptor-Relevant Chemspaces for about 18 different receptors
- Used to estimate receptor specificity
  - suggest when a compound is similar to known ligands for our receptors
  - Is a proposed compound for one receptor likely to interact with others?
- Compound Prioritization user can select:
  - whether compounds should be inside or outside each of the various RRCs
  - cutoff for RRC
  - weighting

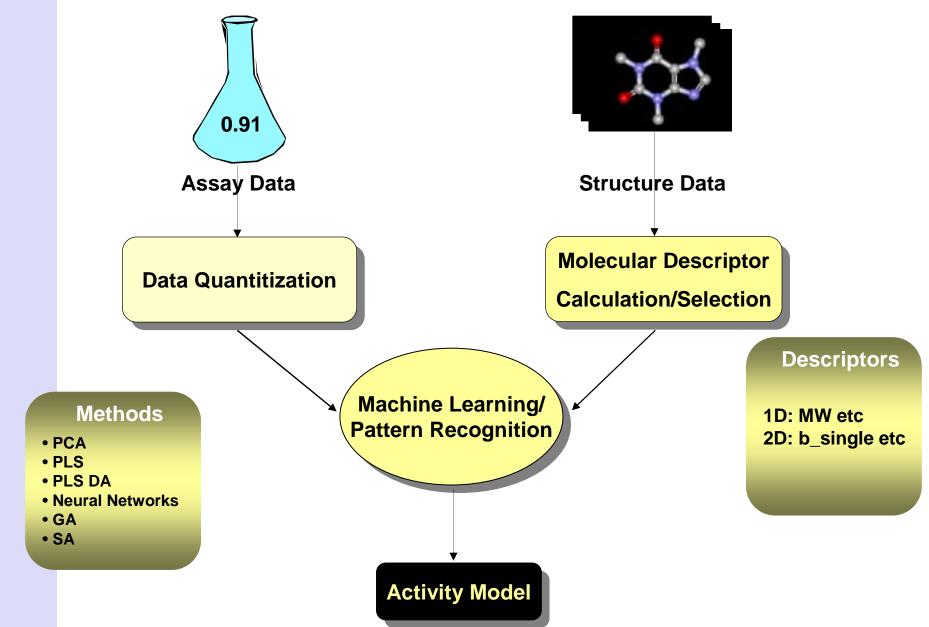


### **LION Predictive ADME Models**

- We predict a variety of ADME parameters:
  - Caco-2 cell permeability (P<sub>eff</sub>)
  - %Fractional dose in portal vein (FDP)
  - Blood-brain barrier penetration (BBB)
  - CYP2D6
  - CYP3A4
- Similar to iDEA product

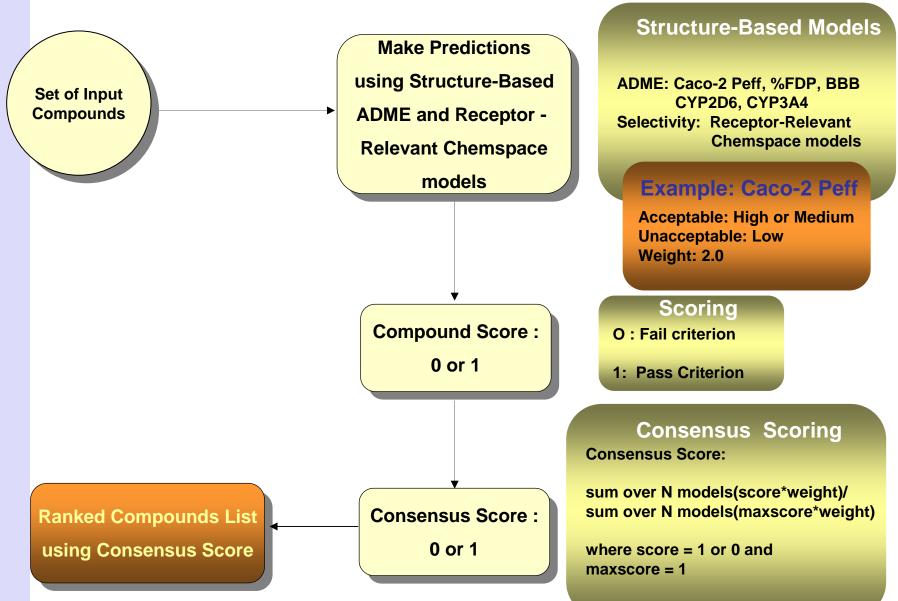


# LION Structure-Based ADMET Predictions: Methodology





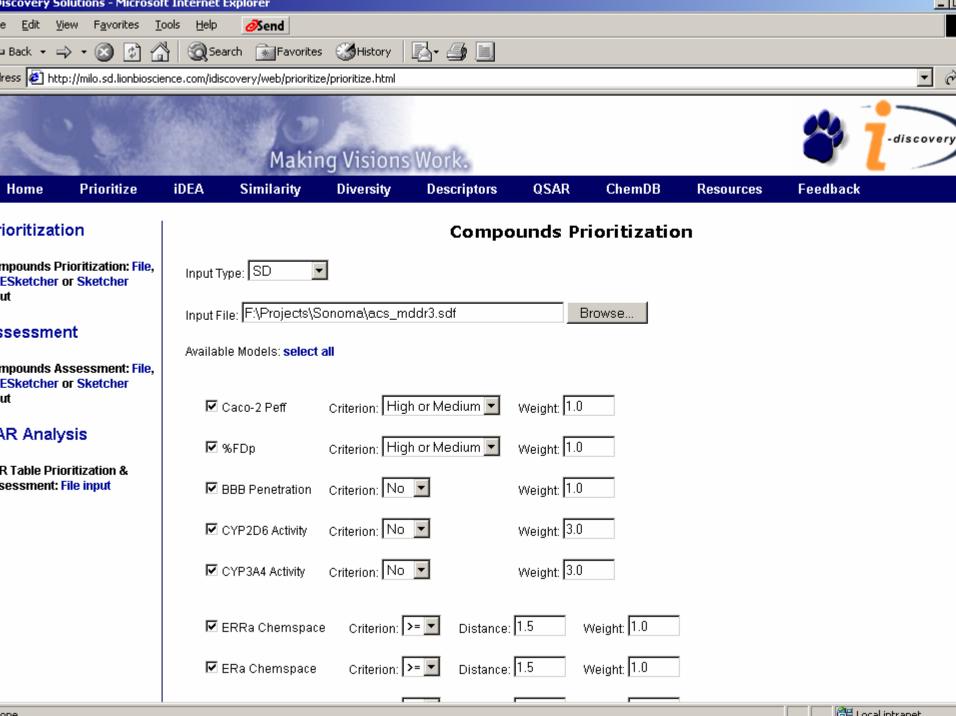
### **Compound Prioritization Schematic**

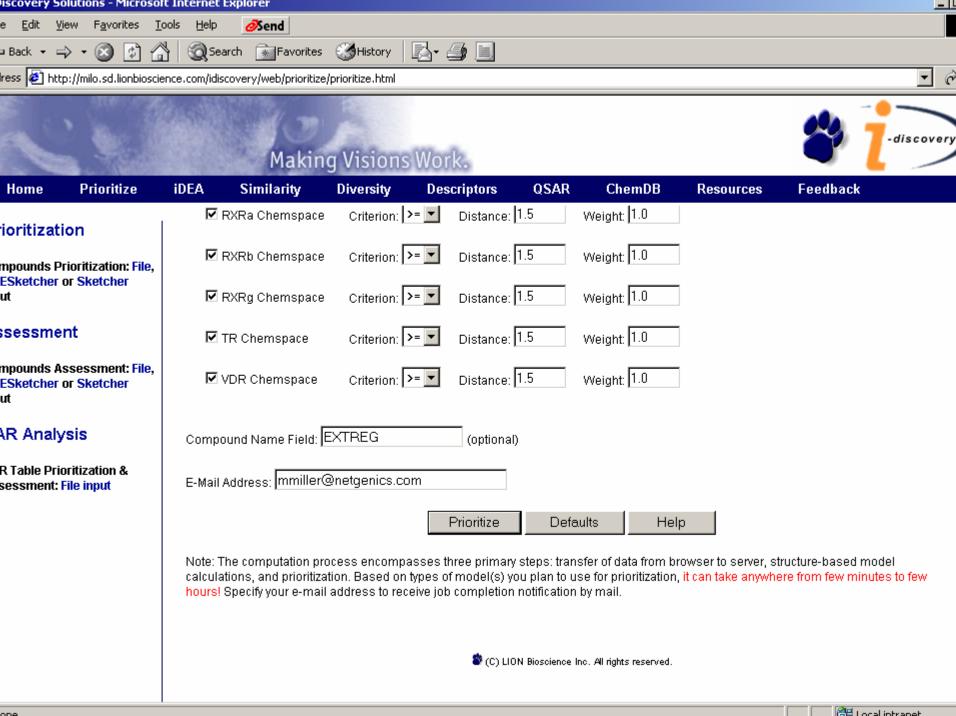


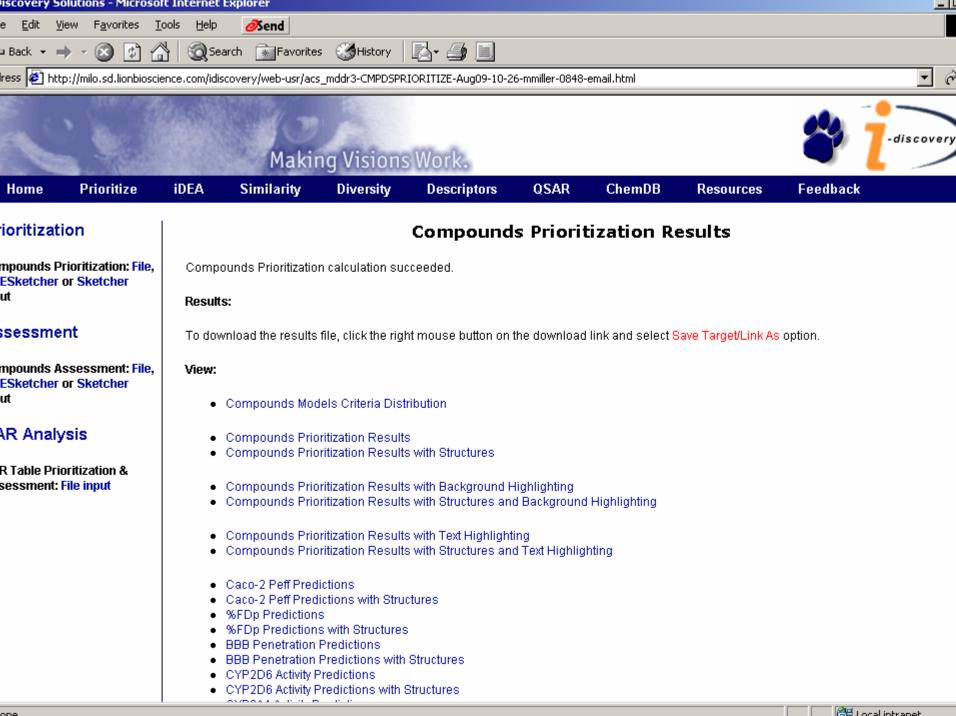


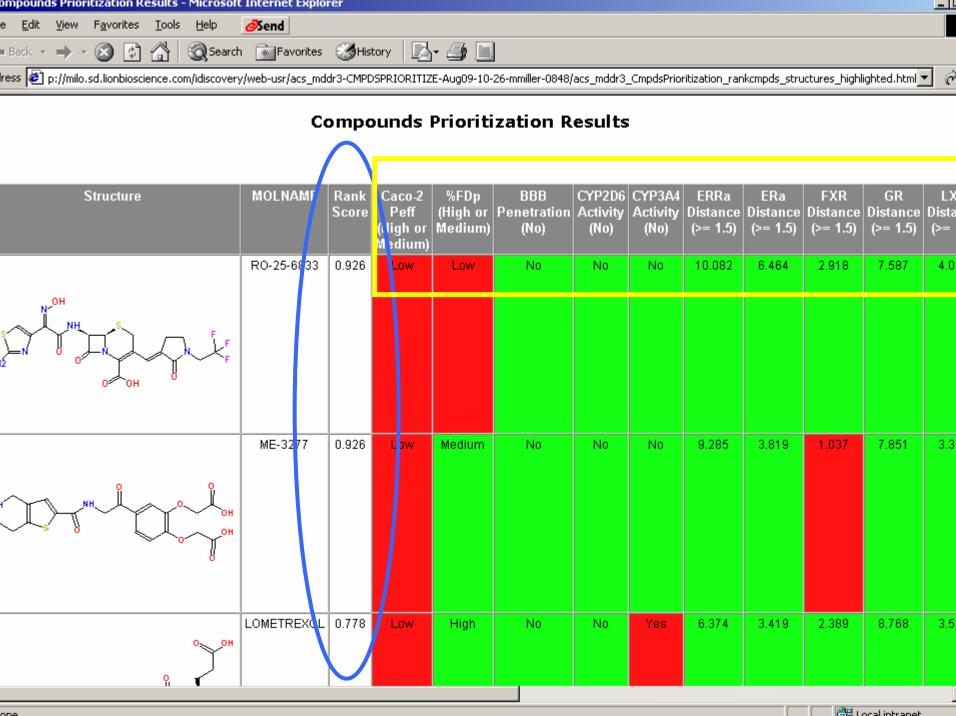
### **Previous Compound Prioritization Workflow**

- Perform search in ISIS/Base
- Export SD file
- Start web application
- Load SD file
- Specify parameters
- View results







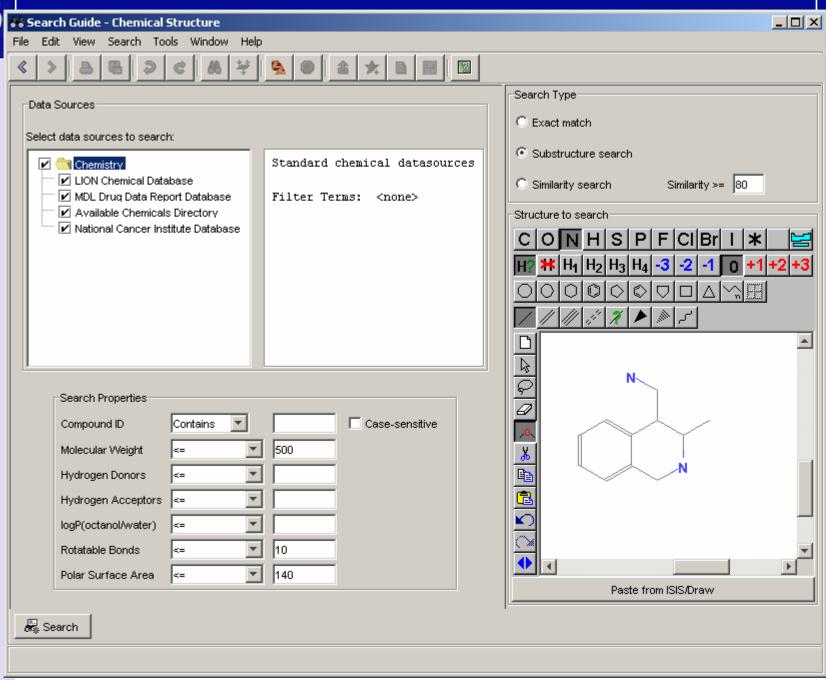


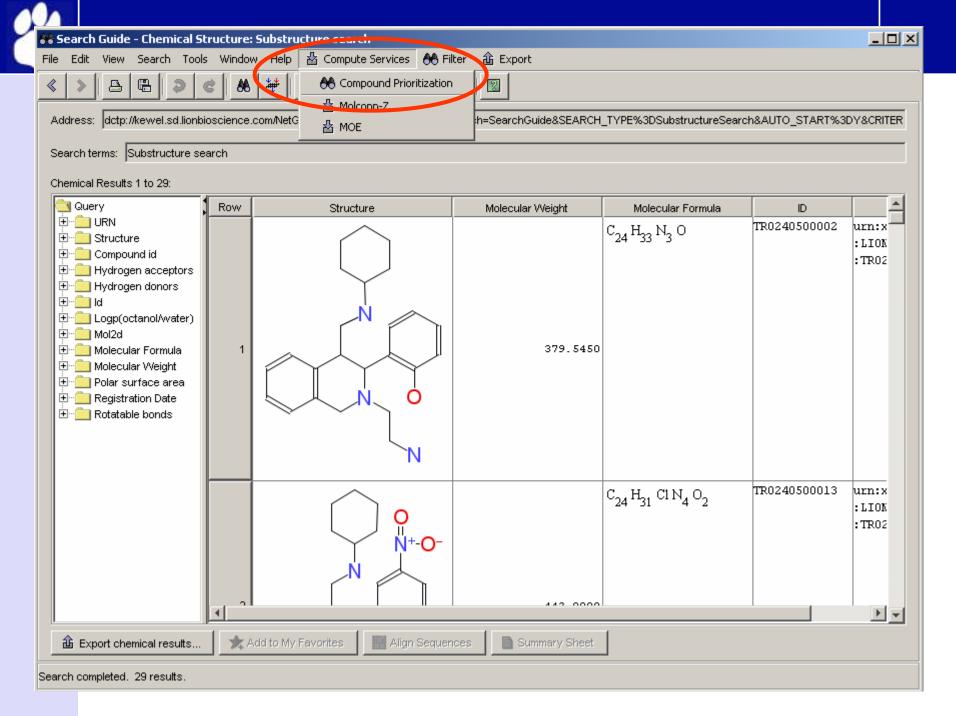


### **Compound Prioritization in LDC**

- Perform search in LDC Search Guide
- Browse results
- Select 'Compound Prioritization' from application menu
- Fill in parameters
- Add results to form







×



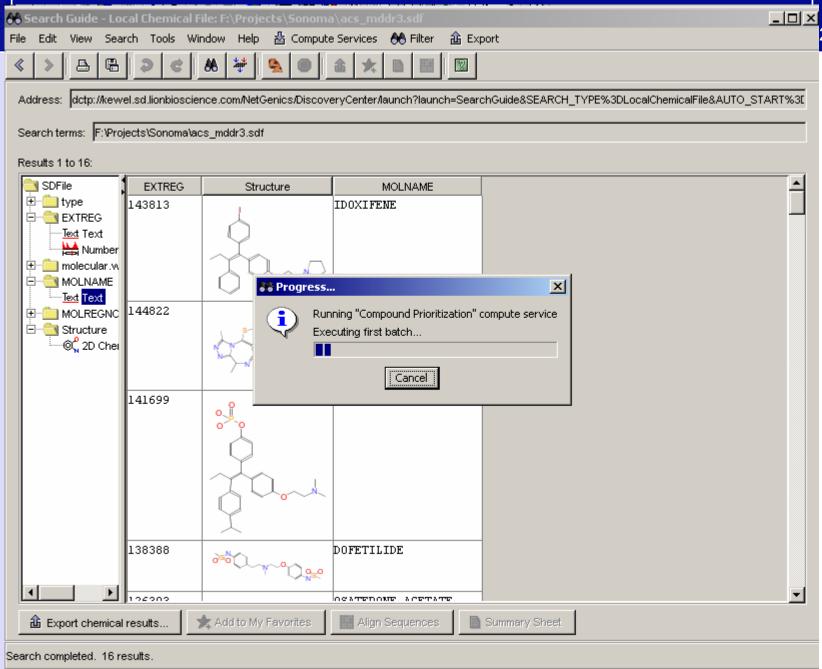
Compound Prioritization

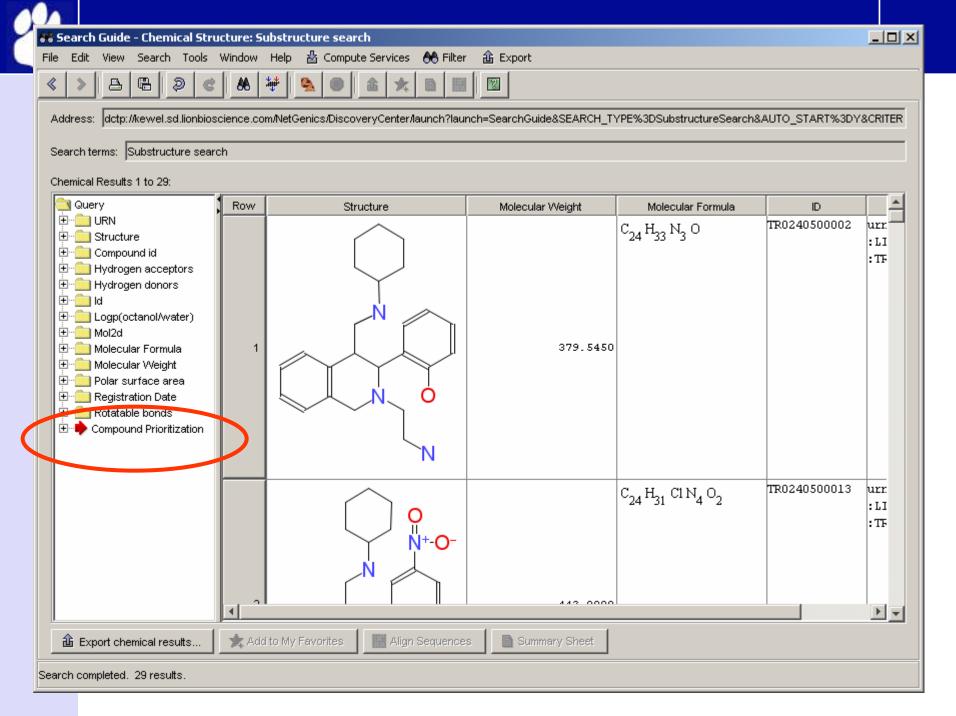


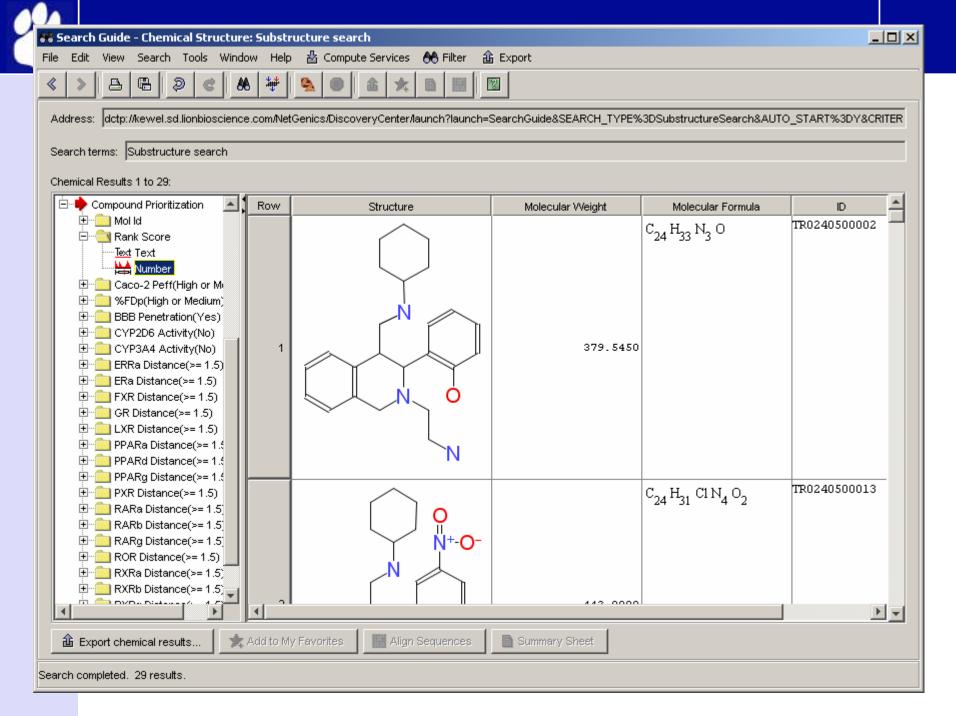
Available Models

☑ Caco-2 Peff	Criterion:	High or Medium 💌	Weight:	2.0			
<b>▽</b> %FDp	Criterion:	High or Medium 💌	Weight:	1.0			
BBB Penetration	Criterion:	No 🔻	Weight:	1.0			
CYP2D6 Activity	Criterion:	No 🔻	Weight:	3.0			
CYP3A4 Activity	Criterion:	No 🔻	Weight:	3.0			
ERRa Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
▼ ERa Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
▼ FXR Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
GR Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
☑ LXR Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
PPARa Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
PPARd Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
PPARg Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
PXR Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
RARa Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
RARb Chemspace	Criterion:	>= 🔻	Distance:	1.5	Weight:	1.0	
RARg Chemspace	Criterion:	>= ▼	Distance:	1.5	Weight:	1.0	
ROR Chemspace	Criterion:	>= ▼	Distance:	1.5	Weight:	1.0	
		ect all models S	elect all mod	lels			

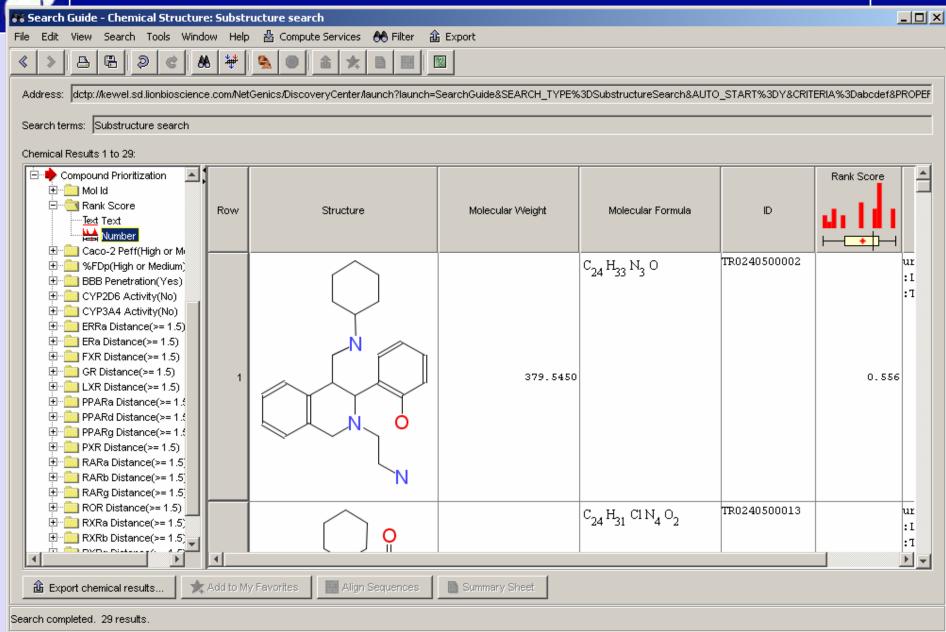


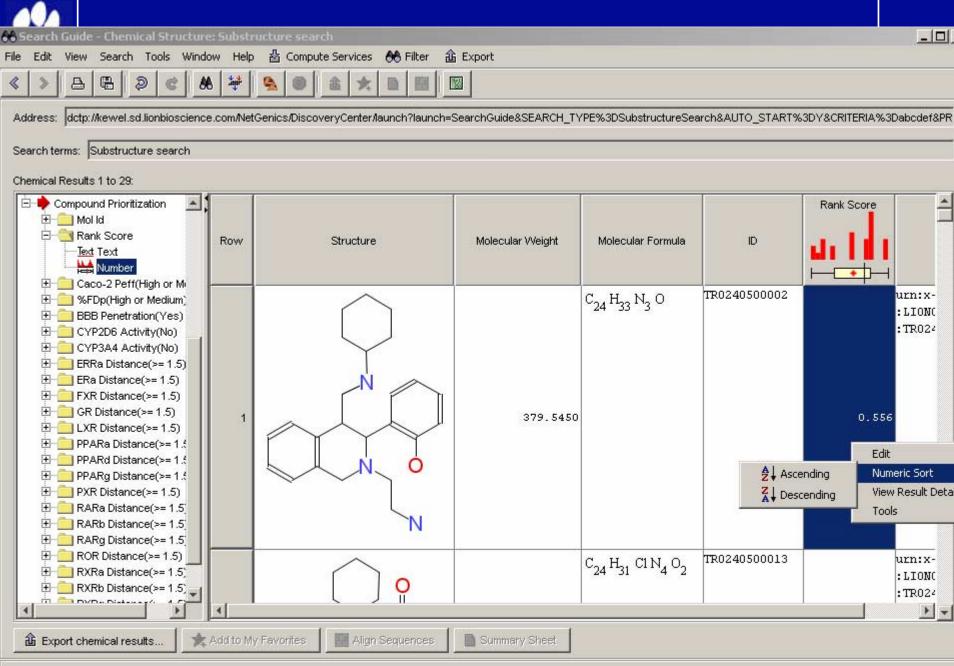




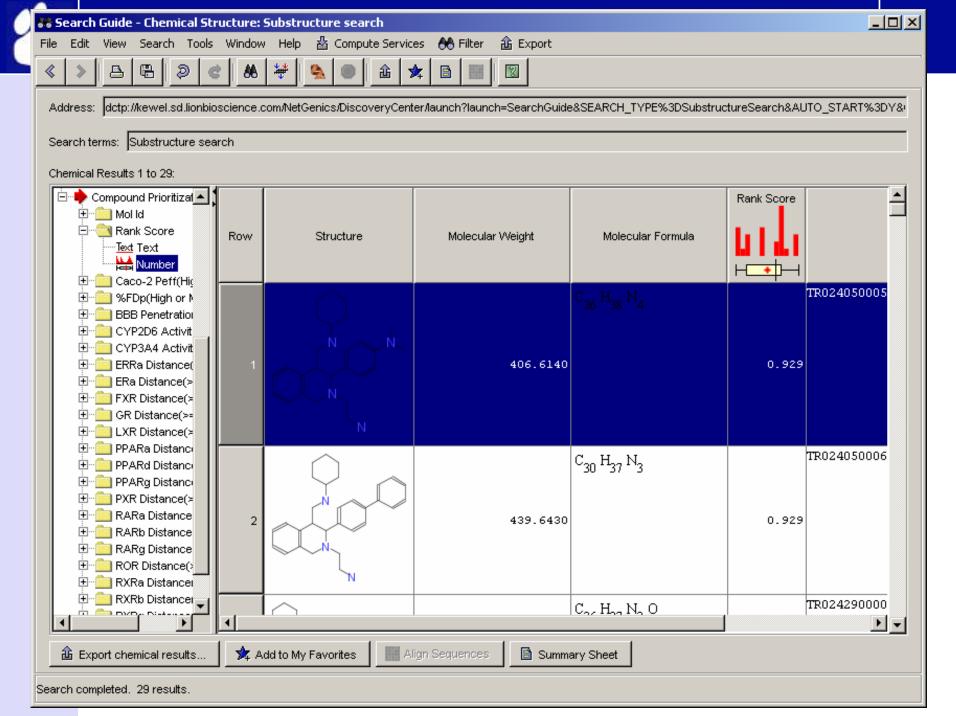


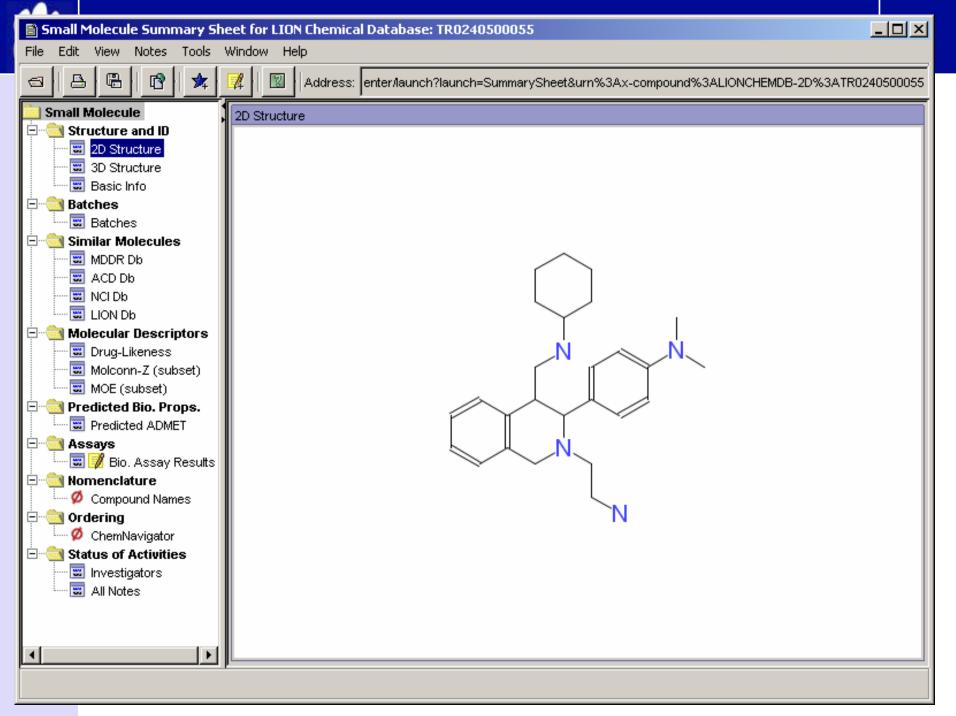






Search completed, 29 results.



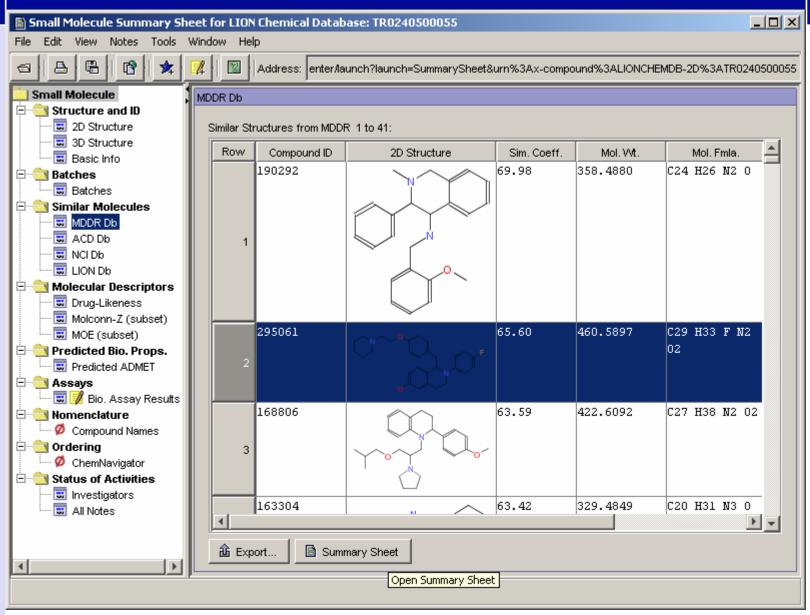




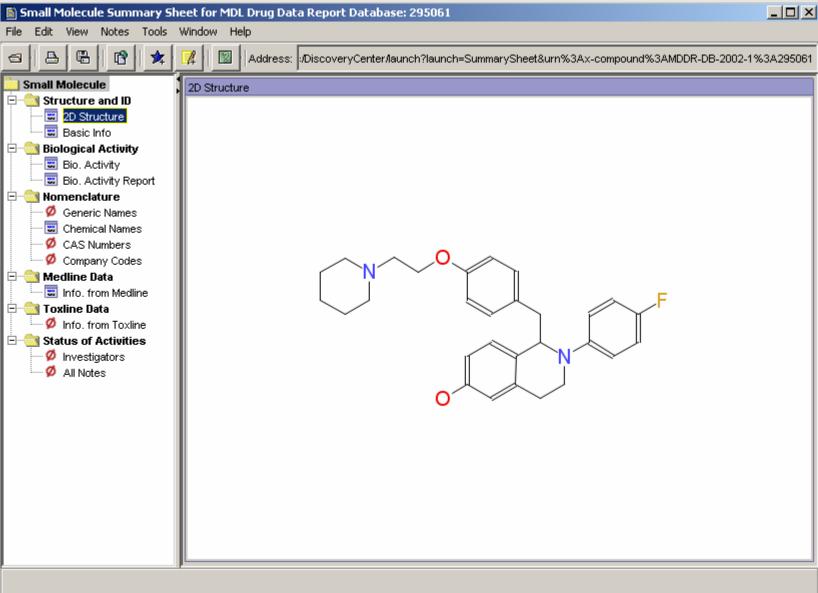
### **Integrating Different Entities by Relationship**

- Starting from one compound of interest, we can proceed in many directions:
  - Find other, similar compounds
  - Find assays of interest

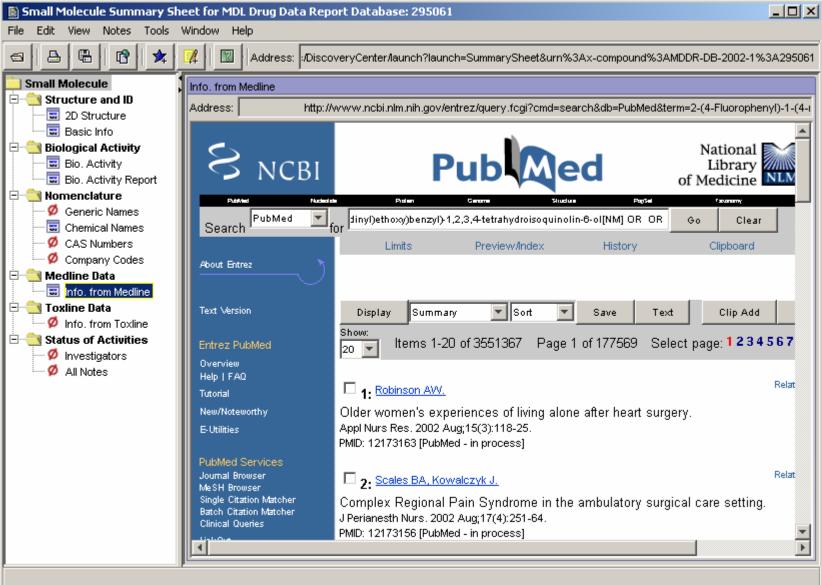


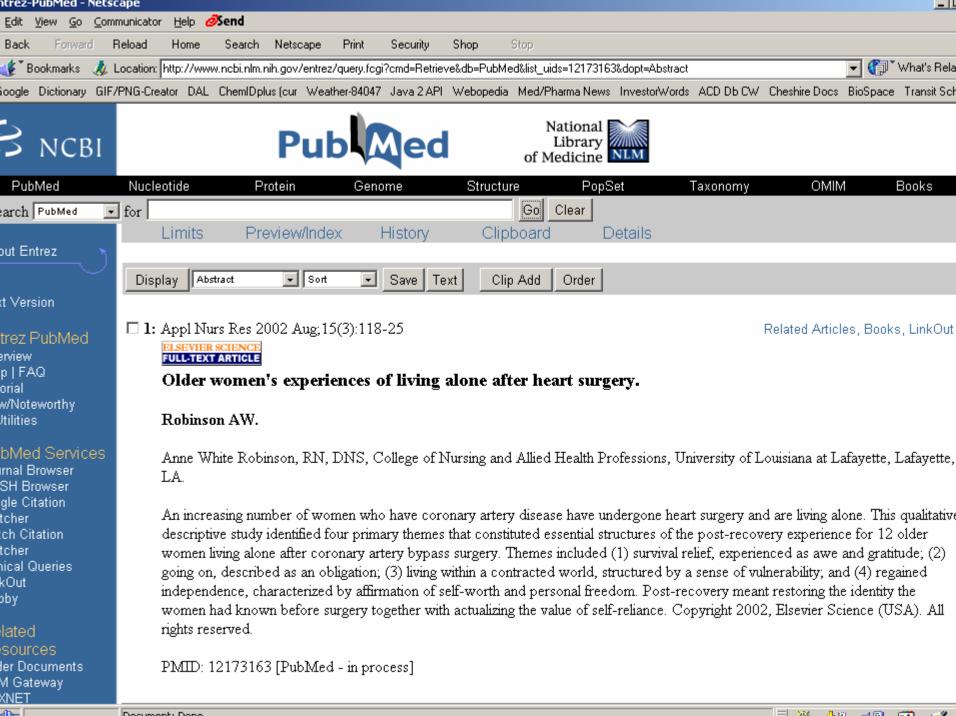




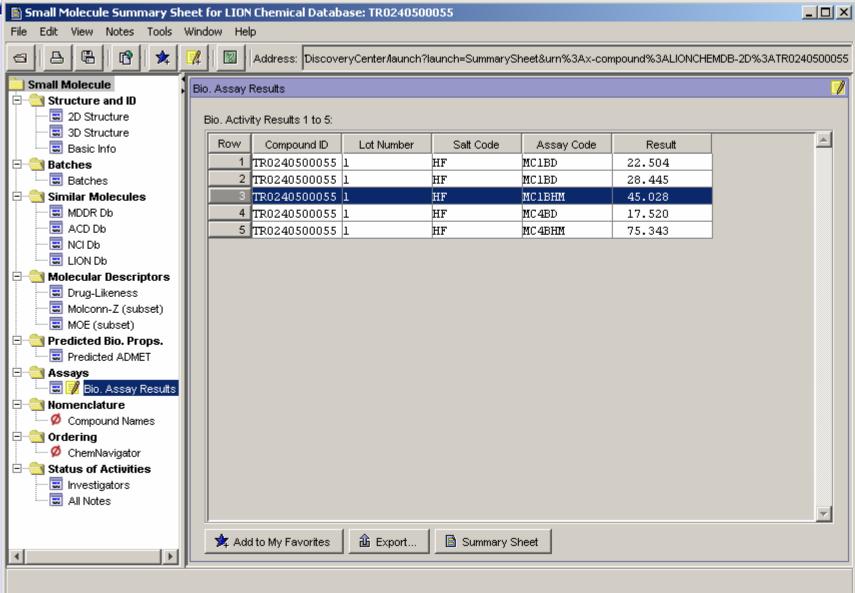




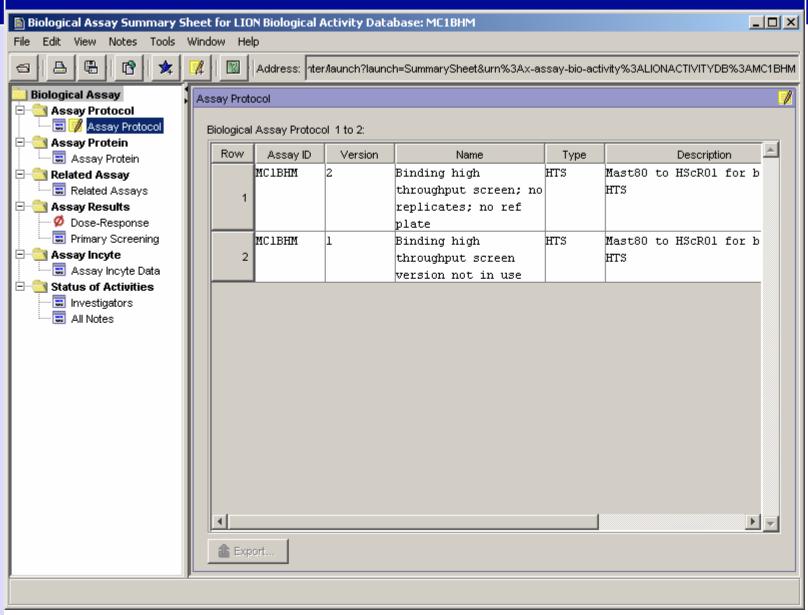




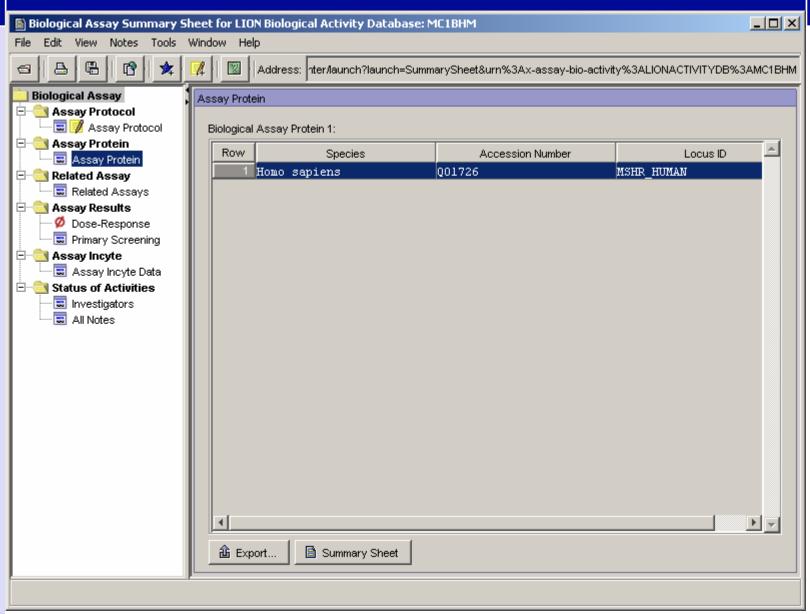




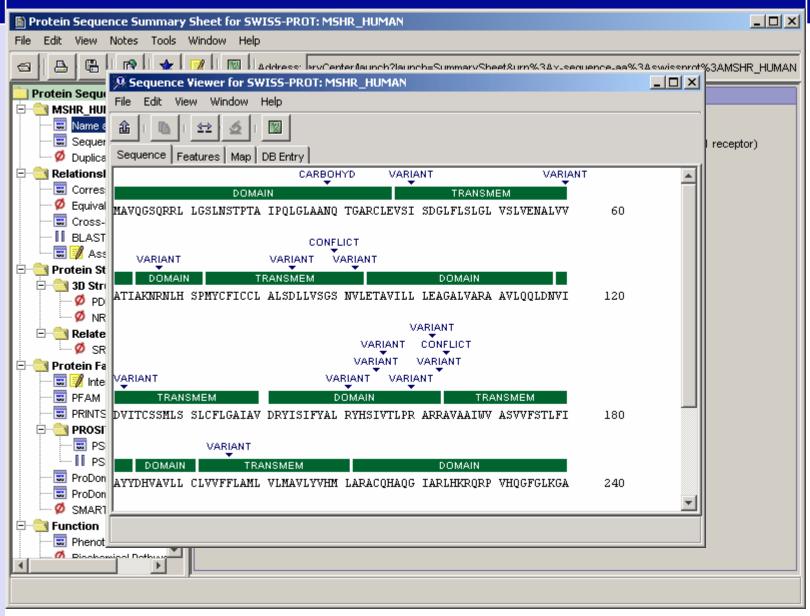




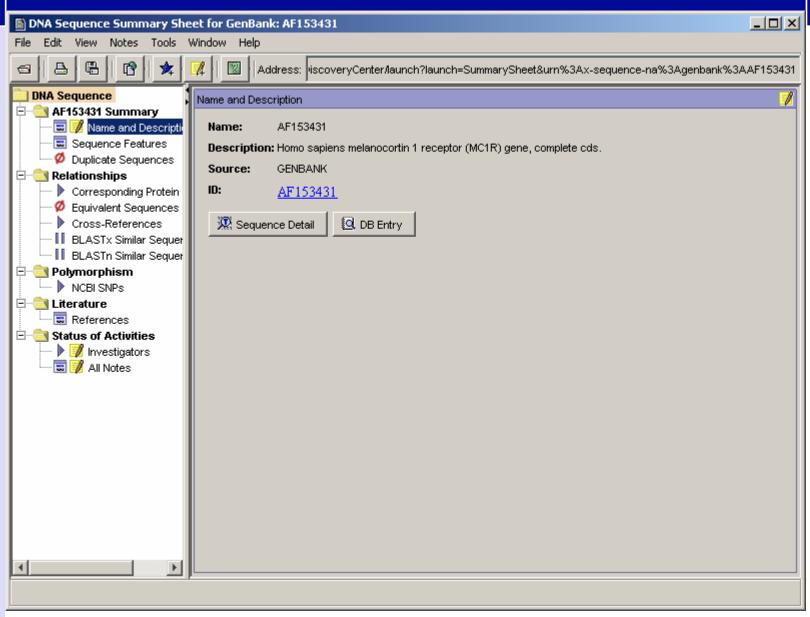














### **Conclusion**

- Integration decreases the learning curve for a set of applications by providing a unified interface and allowing them to share information
- LION DiscoveryCenter provides integration
  - between chemical applications
  - between chemistry and related drug-design disciplines

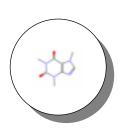


### **Acknowledgements**

- John Harby
- Ty Jacobs
- John Jaeger
- Lewis Jardine
- Chris Knauft
- Anatoli Krassavine

- Darryl Leòn
- Scott Markel
- Stephen Porter
- Adrian Smith
- Manish Sud
- IT Groups





mitchell.miller@lionbioscience.com