

# SEARCHER

The Magazine for Database Professionals



## THE **BETTER** MOUSETRAP

by Edward Vawter

President

QD Information Services LLC

## MDL DiscoveryGate: A Great Chemistry Web-Based Search Product

It's not often that I get excited about a Web-based search product. I've seen my share of them over the years and they are usually lacking many of the features that would make them useful for performing

database research. However, after recently trying out MDL DiscoveryGate [http://www.discoverygate.com], a paid search service from Elsevier, I find myself impressed. Elsevier understands what good Web-based

services can do and has really produced an interface and product worthy of the 21st century.

MDL DiscoveryGate combines databases, major reference works, and xPharm, an integrated collection of pharmacological information. The database function aggregates several databases and allows you to search for chemical information by both structure and/or text in fields across databases. It claims to cover more than 20 million compounds, 11 million reactions, and 500 million observed properties. The databases available include Crossfire Gmelin and Beilstein, MDL Compound Index, MDL Comprehensive Medicinal Chemistry, MDL Drug Data, National Cancer Institute DB, MDL Metabolite, MDL Toxicity, and many more. The databases are divided into several categories, such as bioactivity, patents, and

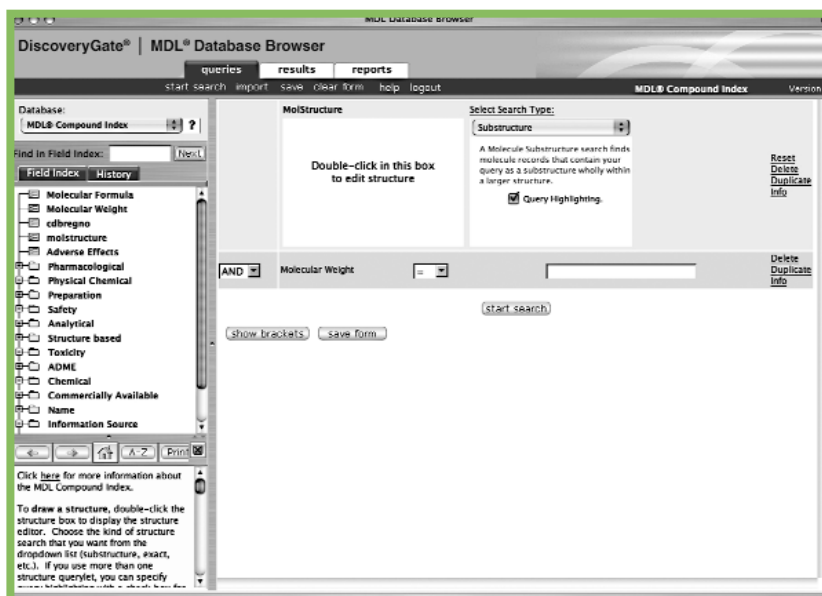


Figure 1.

synthetic methodology. The Help information describes each database and the information it contains.

The Web-based product runs in Java, so it is truly cross-platform compatible. Sometimes Java-based products are slow, but in this case, the creators have done their homework and the browser-based product works well. It's almost as fast as a normal application. As a Mac user, I appreciate the fact that time was taken to ensure the platform independency of the product. The product also advertises that it is Mac compatible. (See Figure 1.)

When you select a database to search on the left, the fields below are updated to reflect those available within that source. You can use a default search form or create your own by double-clicking on any field to add it to the search form. I have played with it quite a bit and found it easy-to-use and I am generally able to find just about any sort of information. (See Figure 2.)

Users do have to develop an idea of what each database contains in order to know where to search. However, once you perform a search and get results, you can easily look for a given compound in other databases. I like this, because it allows the searcher to concentrate on the content they are trying to retrieve, rather than trying to remember archaic commands that allow you to switch between databases. For example, I recently began tracking four first-generation over-the-counter antihistamines: Benadryl, Hydroxyzine, Chlortrimeton, and Tavist. The custom search form I used is shown in Figure 3.

While the field name says "Generic Name," it appears to contain both generic and trade names for compounds. I used the MDL Medicinal Chemistry database to perform the search since it contains information on such drugs. One criticism: The system lacks a progress bar to

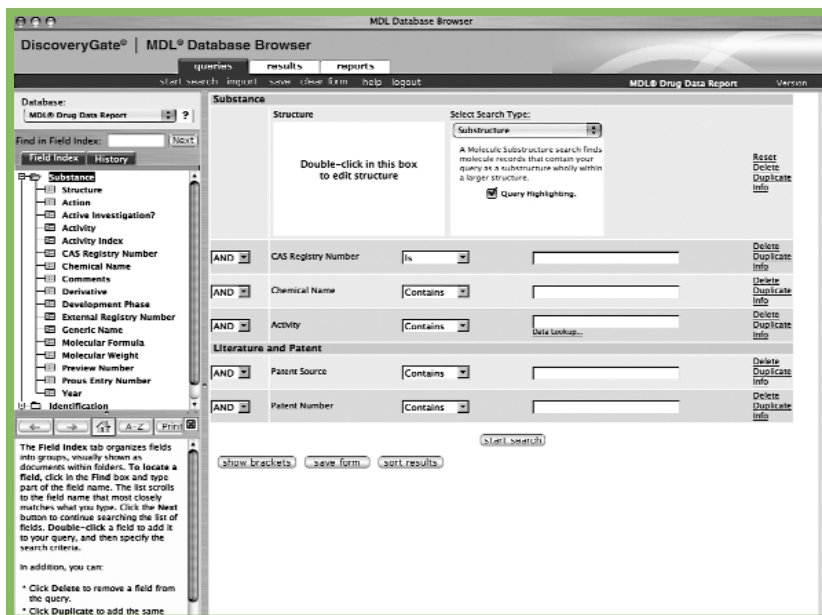


Figure 2.

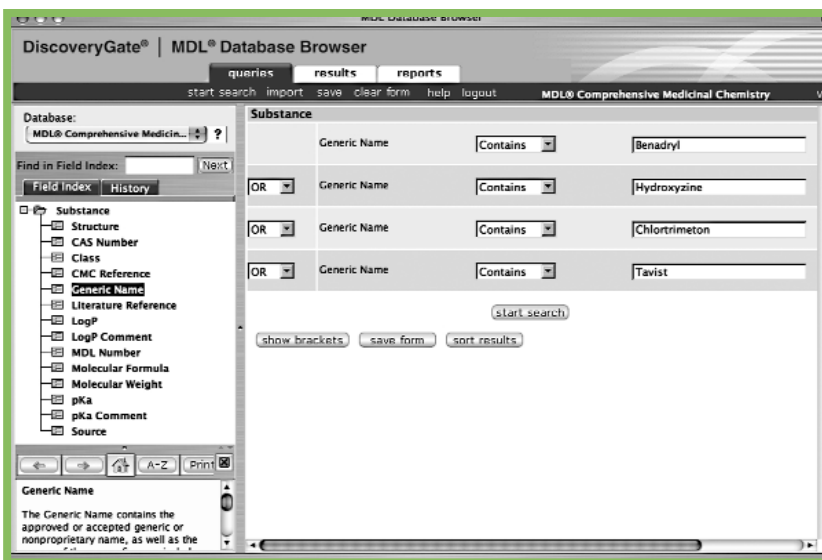


Figure 3.

tell you the system is still searching and hasn't "frozen."

This particular search returned four records, one for each compound. (See Figure 4.)

When you begin exploring the details for each compound, you notice that there are other databases listed across the top (and bottom). These link to each compound

in other databases. Figure 5 shows the detailed result for the first compound.

Information on this compound also appears in 11 other databases. You can easily switch to any of those by simply clicking on that link. The only downside is the use of an abbreviation for each database; you have to learn what each abbreviation stands

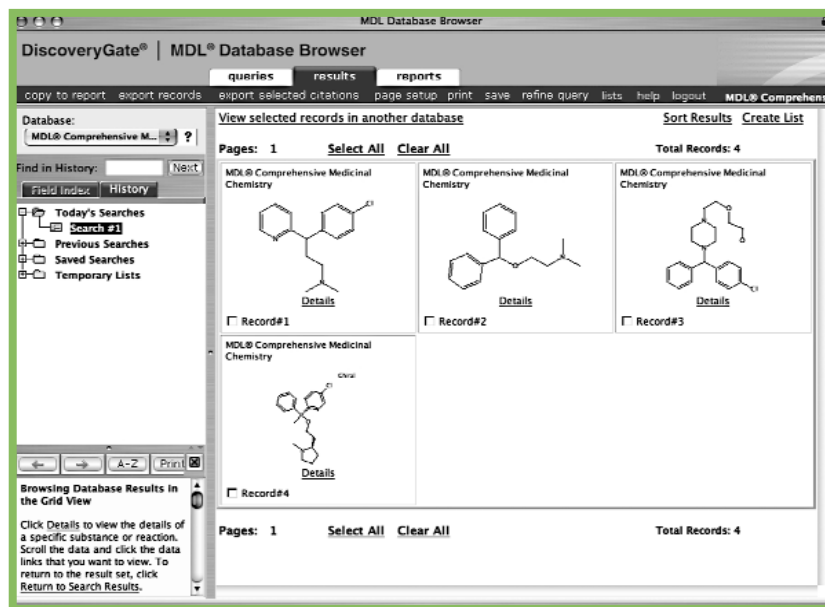


Figure 4.

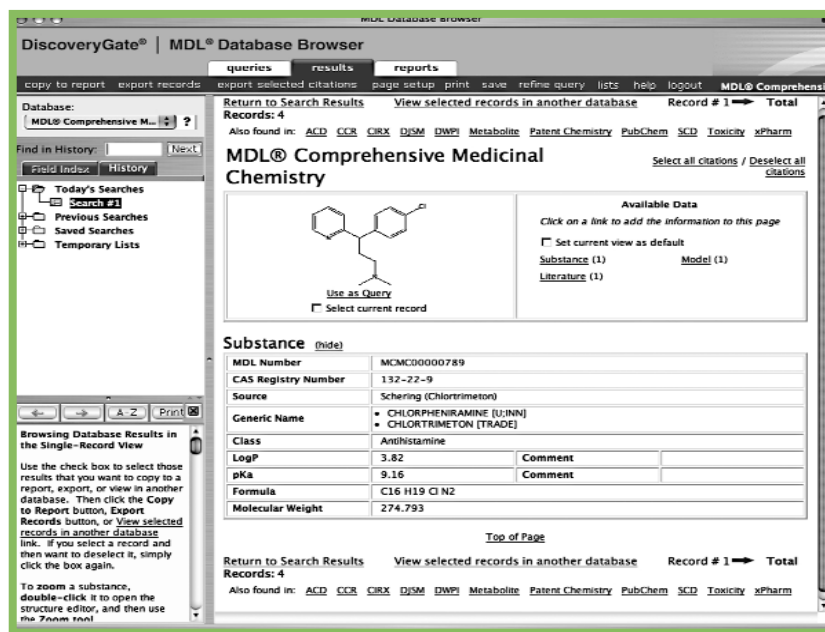


Figure 5.

for. In general, the abbreviations are fairly straightforward; ACD is Available Chemicals Directory and DWPI is Derwent World Patent Index. It would help if there was pop-up information with the full name when you hover your mouse over the link. The ability to quickly switch be-

tween databases allows you to concentrate on finding the information of interest rather than worrying about how to perform a search. In this case, clicking on the ACD took me to the entry in the Available Chemicals Database, where I found suppliers and how much they charge

for various quantities.

One of the strengths of DiscoveryGate is the interconnectedness via links between databases. This even allows you to retrieve the structure based on a name and then do further searching based on that structure. You can go from a name to a structure and then find related compounds by performing a sub-structure search on that result — all with a few simple clicks. It also has a very good chemical drawing program that is easy to learn and use. Anyone who has used Chemdraw or ISIS Draw will find it easy to draw even complex structures with MDL Draw. I did find it easier to retrieve a structure by name or CAS number and then do modifications, rather than drawing a structure from scratch.

DiscoveryGate keeps track of what you have done in three tabs; queries, results, and reports. This allows you to easily look at the data you are working with. If you have forgotten your exact query, you can easily look back and see it. You can also save your results and go back and look at your previous searches, provided you save that information.

Another excellent feature is the ability to create custom reports. You can choose which compounds or data you wish to include from various databases and create a report to reflect exactly the information you want. The pieces of data are called "sections" and you can rearrange them before printing or saving. You can then print out the results or save them locally in HTML format. You can also create citation reports of literature references in either RIS format for use within a bibliographic program, such as Procite or Endnote, or as a simple text file.

The product has many other features that I don't have the space to talk about, but if you want to see what it looks like and what you can do with it, check out the 3- to 6-

minute interactive videos available at <http://www.mdl.com/solutions/videos/index.jsp#discoverygate>. There are a range of tutorials covering searching for molecules, reactions, physical properties and more. I found them very helpful in getting acquainted with the product.

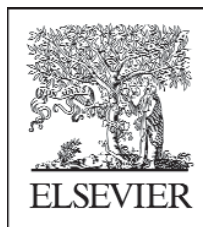
## A Winner, at Last

Overall, I found MDL DiscoveryGate an excellent product, far better than others I have tried. It still has a few issues, such as the lack of a progress bar during searching and the use of initials for database names, but these are minor com-

plaints. While access to these products is not cheap, the interface for searching and browsing the data, as well as for creating reports, is excellent. Each database is individually priced and can be purchased in any combination; the price includes one user with unlimited searches. For example, the MDL Comprehensive Medicinal Chemistry database, one of the cheapest, costs \$969 per year. On the upper end of the price range, the MDL Patent Chemistry database costs \$4,447 per year. The most expensive is the MDL ChemInform Reaction Library at \$9,661 per year. You can find the price list at [https://secure.mdli.com/online\\_order/buy\\_categories.jsp?type=dg](https://secure.mdli.com/online_order/buy_categories.jsp?type=dg). The price

per user goes down if you purchase multiple licenses and site licenses based on IP-authenticated access rather than designated user access are offered. For information on these options, you should contact MDL directly using the contact information on its Web page.

Overall, the product may be too expensive for a small company, but for a company of even moderate size, MDL DiscoveryGate is a first-rate tool that allows the searcher to concentrate on the information they need rather than typing in cryptic commands to perform complex functions. ♦



### ELSEVIER MDL

*Data on demand. Power for decisions.*

2440 Camino Ramon, Suite 300  
San Ramon, CA 94583  
Phone: +1 (925) 543-5400  
Fax: +1 (925) 543-5401  
[www.mdl.com](http://www.mdl.com)

Elsevier MDL provides informatics, database and workflow solutions that accelerate successful R&D by improving the speed and quality of scientists' decision making. Academic and industrial researchers around the world depend on Elsevier MDL for innovative and reliable discovery informatics software solutions augmented by 400 Elsevier chemistry and life sciences journals and related products.