MDL Patent Chemistry Database

Brief Database Description

Date: 3-November-2004

MDL Patent Chemistry Database is a new structure searchable database of reactions and substances from organic chemistry and life science patents since 1976, specifically designed for researchers and information professionals.

MDL Patent Chemistry Database covers chemical reactions, substances and related substance information from about 340,000 World (WO), United States (US) and European (EP) patents from 1976. The database covers patents from the following research areas [International Patent Class in brackets]: Organic Chemistry [C07]; Preparations for Medical, Dental, Toilet Purposes [A61K]; Biocides, i.e. Agrochemicals, Disinfectants [A01N]; Dyes and Related Compounds [C09B]. The database addresses the following business critical issues:

More Effective Synthesis Planning

At release, the MDL Patent Chemistry Database contains 1.5 million structure-searchable reactions from patent documents since 1976. By providing access to patent reaction data in an easy-to-search format interlinked to other important information sources, the MDL Patent Chemistry Database helps you develop effective synthesis plans based on a wider view of known chemistry.

− In addition to standard reaction data fields, almost all reactions have the full **experimental section** from the patent text supplying you with crucial details about the reaction process.
− Reactions are indexed with **InfoChem ClassCodes**, which enables you to cluster similar reactions and find similar reactions in all other MDL reaction databases.
− It is the first patent database displaying **Markush reactions** [1] allowing a better understanding of the claim.
− For reaction products, spectral data (incl. **signals**) and other physical data are given.
− The **location in the patent*** (e.g. the page number) makes it easy to find the reaction in the patent document.

Better Bioactivity Profiling

It is crucial to find the gaps in patent coverage for chemical processes, compounds, and their applications. Especially in the life sciences, researchers must compile full accounts of compound data sets for developing bioactivity profiles, selecting and optimizing leads, or designing combinatorial libraries. By combining structural, text, and numerical patent data with powerful visualization tools, the MDL Patent Chemistry Database helps you explore the compound and bioactivity space. The first release will contain about **1.6 million compounds from patents.** Unique compound related content and features are:

− The database indexes not only defined compounds, but also more than 800,000 **"Prophetic Compounds"** rarely covered elsewhere. Prophetic compounds are legally relevant compounds for which the inventor states that they can be prepared analogously to described methods, but doesn't prove their existence with supporting data (e.g. yield, spectral data).
− The database indexes measured **numerical data** like application and bioactivity* data as well as formulation*, detailed spectral data – e. g. i.e. NMR, IR, MS signals -and other physical data (i.e. LogP).

- Exporting structures and their numerical bioactivity* data into a tabular form supplies a clearly arranged **structure-activity-relationship-table.**
- The result view of a structure search summarizes all property data for a given compound derived from different patent documents.
- You can display a graphical representation of a **Markush structure*** linked to the correlated defined compounds within the patent document.
- The **location in patent*** (most often the page number) allows easy finding of the data in the original patent document.

Easy Relevance Check

Quickly understanding the scope and relevance of patents can save time, effort and money by directing research to avoid patented compounds, reactions and processes. By bringing claims text, main Markush* structures and reactions to the desktop in an easy-to-view format, the MDL Patent Chemistry Database can make it faster and easier to check the relevance of patent documents.

- In addition to the abstract users can also search the complete **claims text** in combination with a structure or reaction search.
- To simplify relevance checking of the citation, the **claims text** is displayed together with all related **Markush structures* and Markush reactions*** cited in the patent [1].

Integration

MDL Patent Chemistry Database is accessible via CrossFire Commander 7.0 (CrossFire Direct or in-house) and **DiscoveryGate** (as of 2005). The database will be updated bi-weekly.

[1] Note: In first database release these features will be available for patents published from December 2003 onwards.