CHEMICAL STRUCTURE SEARCHING QUICK REVIEW -- 2005

Beilstein/Gmelin 2-14

MDL CrossFire V6 2
Getting Started 3
Drawing Chemical Structures 4-5
Pre-defined Generic / User defined Groups & Atom Lists 6
Templates 7
Searching & Displaying … 8-9
Reaction Searching & Search Tips 9
Custom Displays & Printing 10
Autonom & Data (name/property) Searching 11
Beilstein Notes 12
Gmelin Notes 13
Beilstein & Gmelin's Relationship to SciFinder Scholar 13
Crossfire Documentation 14

Combined Chemical Dictionary 15-17

Dictionary of Organic Chemistry
Dictionary of Inorganic and Organometallic Chemistry
Dictionary of Natural Products
Dictionary of Analytical Reagents
Dictionary of Drugs

Text Searching 16
Structure Searching 17

SciFinder Scholar (Chemical Abstracts) 18-20

Databases 18-19
Search Tips 19
SFS Solutions (Web based help files) 20

Beilstein/Gmelin vs CCD vs SciFinder Scholar 21

MDL CROSSFIRE V6

What is CrossFire?

- CrossFire is a search system with a locally loaded client interface (Beilstein Commander) which provides access to the CrossFire server at the University of Wisconsin - Madison. A handout describing both the databases and search techniques is available at: http://library.caltech.edu/publications/beilstein/BeilsteinGmelinV.6.pdf

What is Beilstein?

- A database of over 35M property records for over 9M organic compounds (including salts, mixtures, peptides with <12 amino acids). Polymers are not indexed.
- Covers the scientific literature from 1771 to the present, with over 1.9M articles. Currently indexing 175 journals http://www.mdl.com/products/pdfs/BSJournals.pdf
- Reactions (9.5M) and properties are critically evaluated, with references to original articles.

What is Gmelin?

- ·A database of property records for over 2.1M inorganic and organometallic compounds (including coordination compounds, alloys, ceramics, minerals, etc.).
- Covers the scientific literature from 1772 to the present, with over 1.1M articles. Currently indexing 62 journals http://www.mdl.com/products/pdfs/GMJournals.pdf
- Reactions (1.65M) and properties are critically evaluated, with references to original articles.

How can I access CrossFire V6?

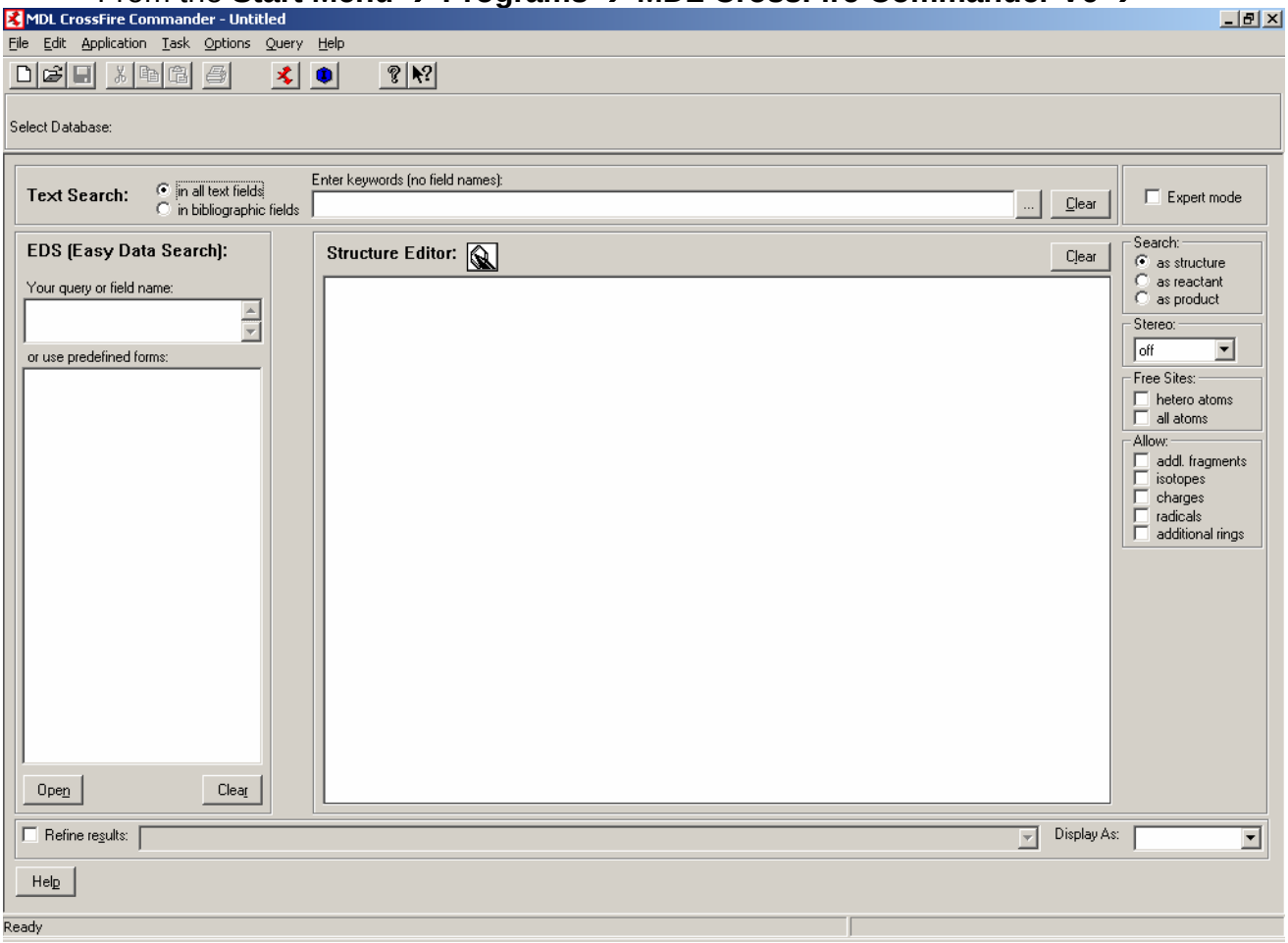
- Workstations with the V6 client are available in the Millikan, SFL, Astro & Geology libraries, under the START MENU (as MDL Crossfire Commander v6).
- Macintosh or MS Windows versions are available for downloading to Caltech workstations from the library web page at: http://library.caltech.edu/publications/beilstein/default.htm

CrossFire features:

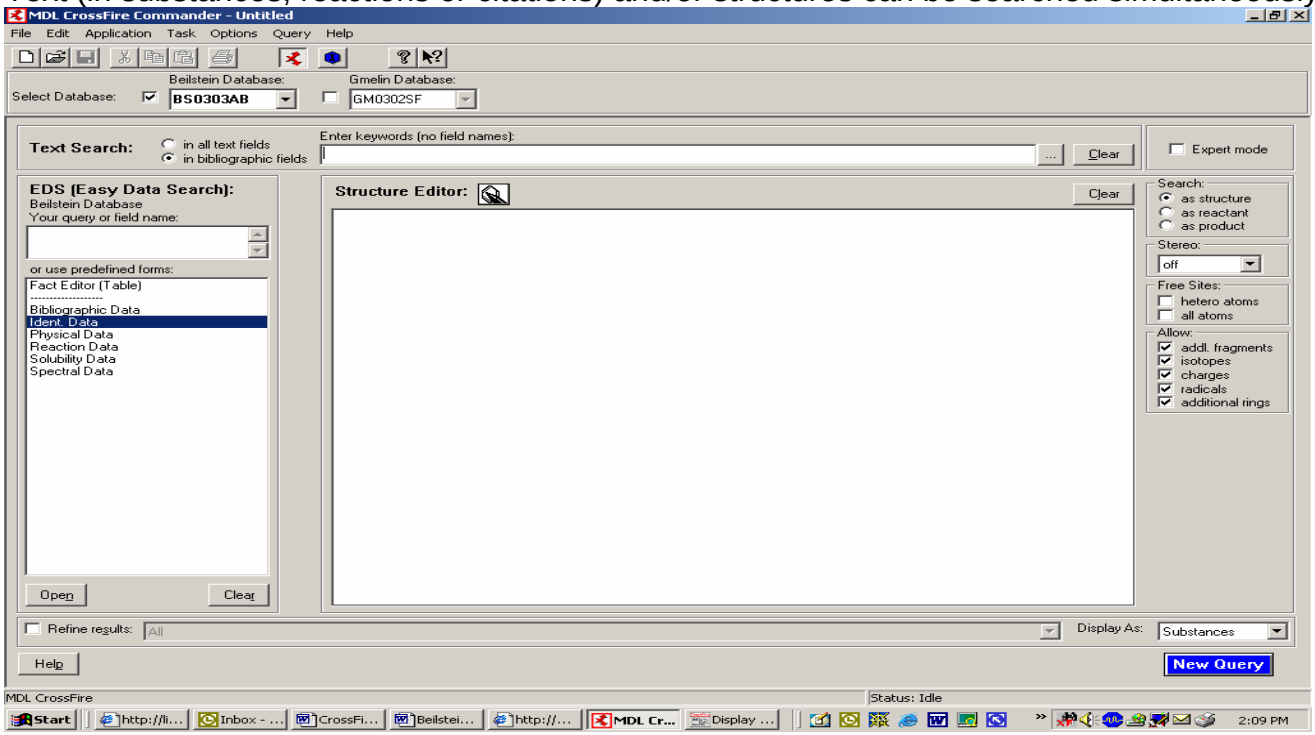
- Autonom - generates an IUPAC name from an organic structure.
- Text Searching - Search text words or author names (article titles/abstracts 1980+)
- Structure Searching - (Beilstein/Gmelin can be searched simultaneously).
  Either draw a structure / sub-structure, modify a (Beilstein) template (under File menu), or search with a name and copy the structure from the 'short display'. Combine with property or text terms (e.g. NMR, malaria). Autosearch may provide additional retrieval.
- Reaction Searching (both Beilstein and Gmelin) by:
  Specify the role (reactant and/or product) of a compound or sub-structure in a reaction.
- EDS (Easy Data Search) - Tables or predefined forms
  Names or portion(s) of chemical names. CAS Registry numbers. Author names.
  Property values and range searching (e.g. melting point between 100 and 102 degrees).
  Reaction data, Solubility, Spectra.
- Combination Searching
  Combine structures / substructures with property terms (e.g. NMR) in B or G (not both).
  Search for text and/or structures in B & G simultaneously.

Getting Started on a Library Workstation

From the **Start Menu → Programs → MDL CrossFire Commander V6 →**



Structure or data searching: Click on the 'Crossed Red Arrows'. After connection, select either Beilstein (organic) or Gmelin (inorganic/organometallic), or both. The default is the last one used. Text (in substances, reactions or citations) and/or structures can be searched simultaneously.

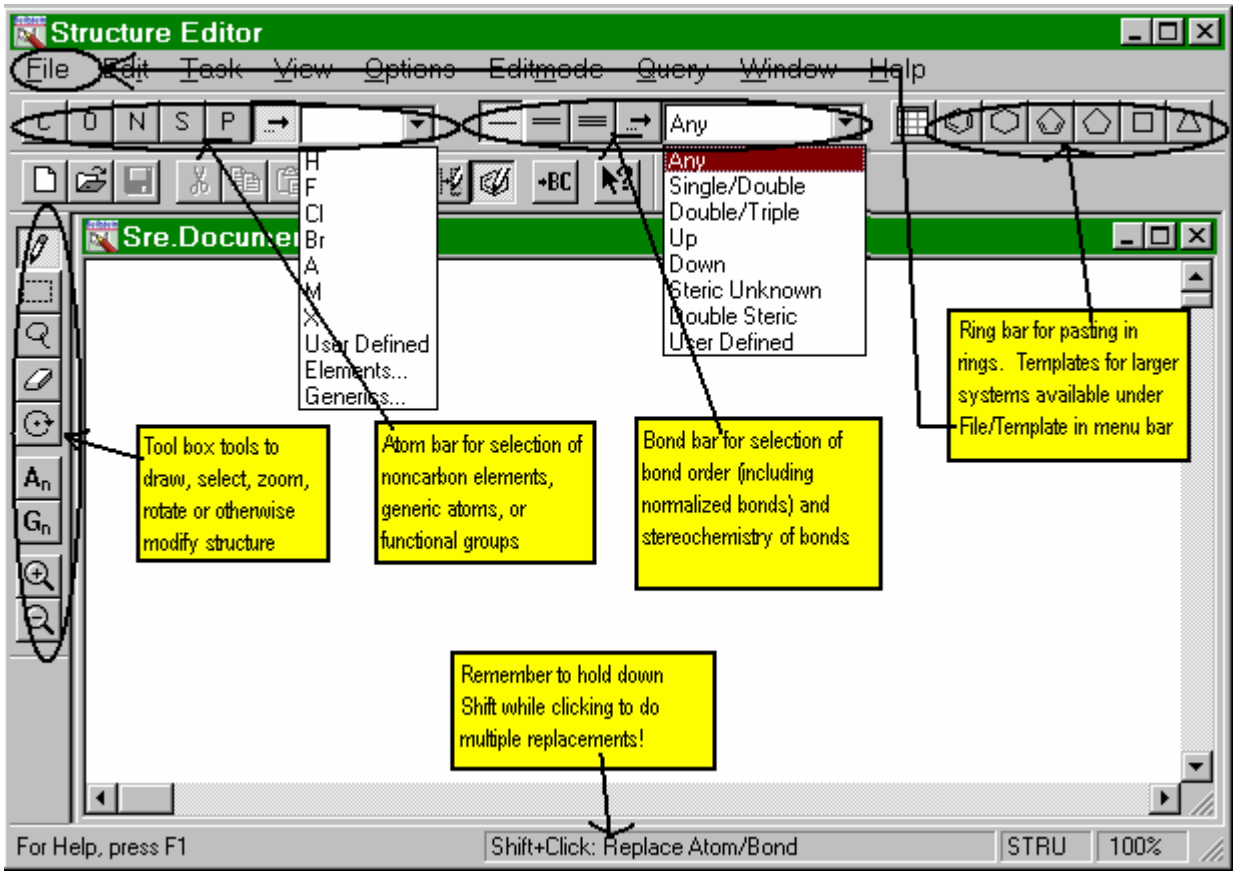


Structure Searching requires double clicking in the Structure Window, or clicking the Structure icon.

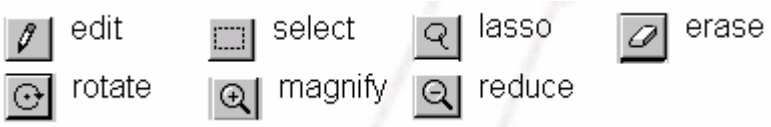
Drawing Chemical Structures -- (Crossfire or ISIS/Draw editors)

Commander screen → Options → Structure Editors → MDL ISIS or MDL Crossfire

The **Structure Editor** screen has a toolbox, and atom / bond / template / ring bars.



The **toolbox** along the left side of the screen provides a set of tools for drawing, selecting, activating, erasing, rotating or zooming structures.



The **atom/bond bars** offer a selection of default atoms and bond types for the **Edit** tool. Atoms or bond types not shown may be specified from either the Atom or Bond pull-down menus, although it is often easier to draw a carbon skeleton and change atoms and bonds as necessary.

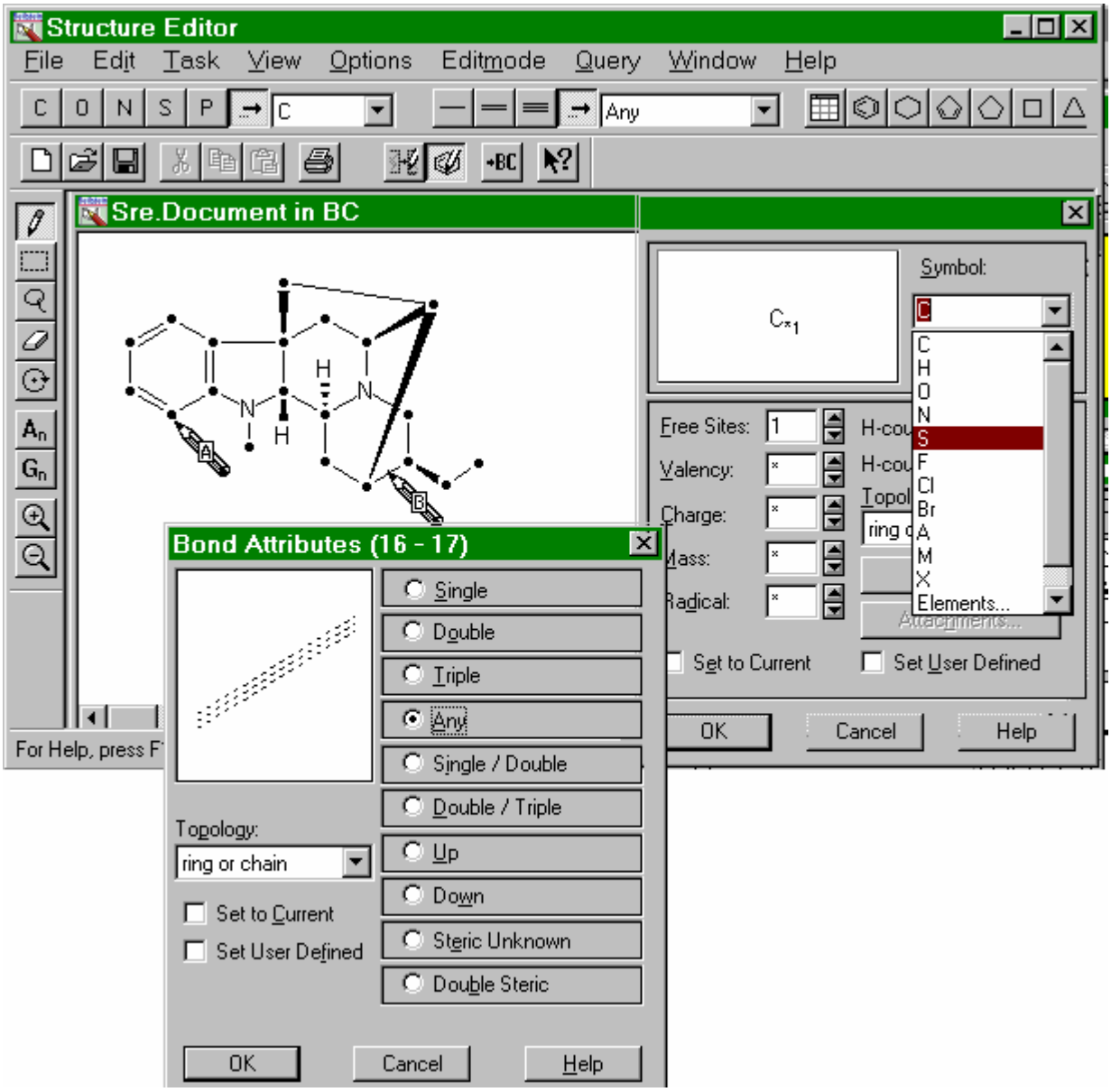
Click on the **Edit** (pencil) **icon** to draw a structure. The default atom and bond type are carbon and single. **Click and hold** down the left mouse button. **Drag** the mouse to the desired end point of the bond; **Release** the mouse button to terminate the bond and place the second atom. Repeat (**Click and hold, Drag, Release**) to continue drawing atoms and bonds.

Atoms and bonds already drawn can be changed by simply changing the default atom or bond (on the Atom Bar) and, while holding down the shift key, click on the atom or bond to be changed. Do **NOT** draw in hydrogen atoms.

To erase a portion of a structure that has been drawn, select the **Eraser** tool from the tool bar. Position the eraser (the tool shows a "B" or "A" when positioned correctly) and click on each atom/bond as needed. To erase a larger portion of a molecule, use the **Lasso** tool to circle the area to be erased and click on the **Scissors Icon**. The screen can be cleared by clicking **Delete All** under the edit menu. **Undo** under the edit menu will undo the last operation.

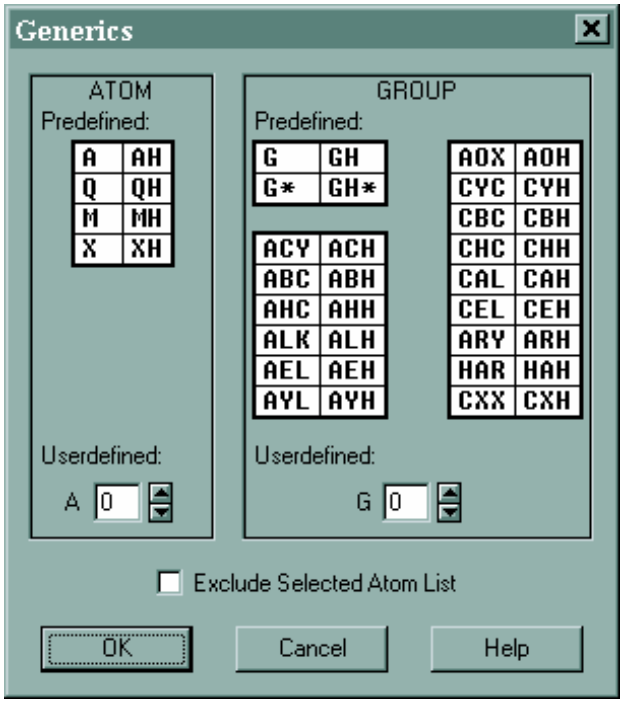
To move or rotate a molecule, first **activate** it (by double clicking one of the atoms or bonds) with the **Select** tool.

To access the full range of options for atoms and bonds, use the **Edit** tool (pencil) to click on a bond or atom (the tool shows a "B" or "A" when positioned correctly). Clicking on an atom displays the **Atom Attributes** box (see figure below, right side), which allows changing the element, free sites, hydrogen counts, etc. Clicking on a bond displays the **Bond Attributes** box (see figure below), which allows changing the bond type, topology, etc.



Pre-defined Generic Groups and Atom Lists

Individual elements, and Generics (either Atoms (e.g. Any Halogen, Any Metal, etc.) or Groups (e.g. Alkyl, Aryl, etc.) may be added to a structure by clicking the 'down arrow' in the atom box on the function bar and selecting either '**Elements**' or '**Generics**' The Generics menu is shown below. Selecting one of the pre-defined atom lists or generic groups, changes the **Edit** tool (pencil) default to that atom list or generic group.



Alternatively, for adding a single group, you can double click an atom with the **Edit** tool, then click the down arrow in the Atom Attributes Symbol Box and select '**Generics…**' from the menu. This also displays the Generics menu shown above. Click on one of the pre-defined atom lists or generic groups. Click OK on the Generics menu and click OK again on the Atom Attributes box. This changes the selected atom.

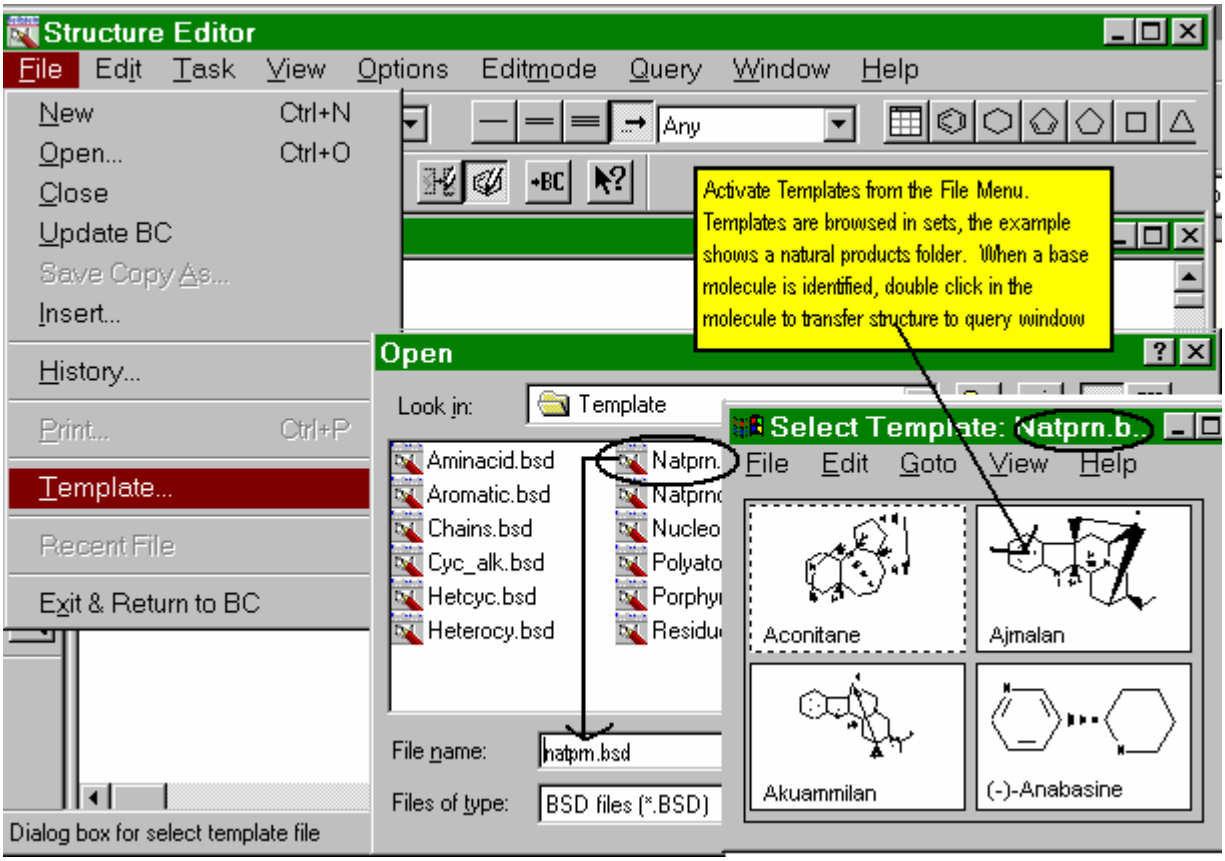
Creating User Defined Groups and Atom Lists

Atoms: Click the atom in the structure, which needs to be variable. Type '**A1**' in the Atom Box and click OK to make the change. Next, click the '**An**' button in the **Tool** Box (left side), and the 'Atom List Number' box will appear. Click OK to display a 'Define Atom List A1' periodic table. Click as many elements or element groups as needed and click OK. A list of the all the selected atoms will appear in the Structure Editor window.

Groups: Click the atom, which will be the variable group in the structure. Type '**G1**' in the Atom Box and click OK to effect the change. Next, draw a group in the Structure Editor window, '**Select**' a group structure (i.e. by double clicking an atom or bond with the **select tool**) and click the '**Gn**' button in the **tool box**. Set **G=1** and click OK. Repeat for additional groups. One atom in each group must be designated as the attachment point. Click on this atom (with the **Edit** tool), to display the 'Atom Attributes box. Click '**Set User Defined**' and then click '**Attachments**'. Click '**1**' in the 'Attachment Points' box and click OK (in that box). Then click OK in the 'Atom Attributes' box.

TEMPLATES (Both Beilstein & Gmelin)

Clicking on a **menu ring**, or clicking on one of the structures in the **Group Templates** folder, transfers an **'active'** structure to the Structure Editor window. This 'active' structure may be moved, resized or rotated. Click outside the 'box' to 'deactivate' the structure. The resulting structure may be further edited (**after clicking on the edit tool**).



Templates for Complex Structures (BC2000 has Group and User Template files)

The following group template files are available:

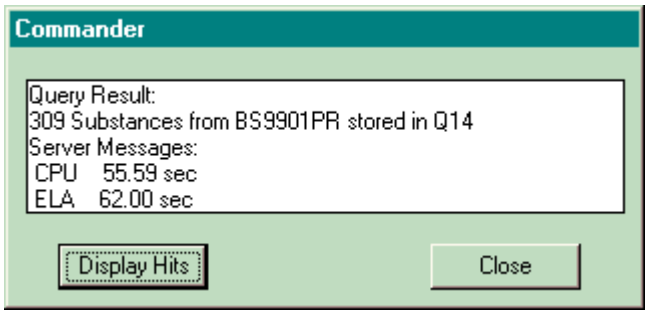
AMINACID.BSD	Amino Acids	POLYATOM.BSD	Coordination Types
AROMATIC.BSD	Aromatics	PORPHYR.BSD	Porphyrins
CHAINS.BSD	Chains	RESIDUE.BSD	-NO2, -COOH, -SO3
CYC_ALK.BSD	Cycloalkanes	STEROIDS.BSD	Steroids
HETEROCY.BSD	Heterocyclics	SUG_OPEN.BSD	Sugars (open)
NATPRN.BSD	Cyclic N-containing	SUG_RING.BSD	Sugars (cyclic)
NATPRNOS.BSD	Cyclic N-, O-, S-containing	TERPEN.BSD	Terpenes
NUCLEOSI.BSD	Nucleosides		

After drawing a structure, click on the Crossed Red Arrows → Beilstein Commander.

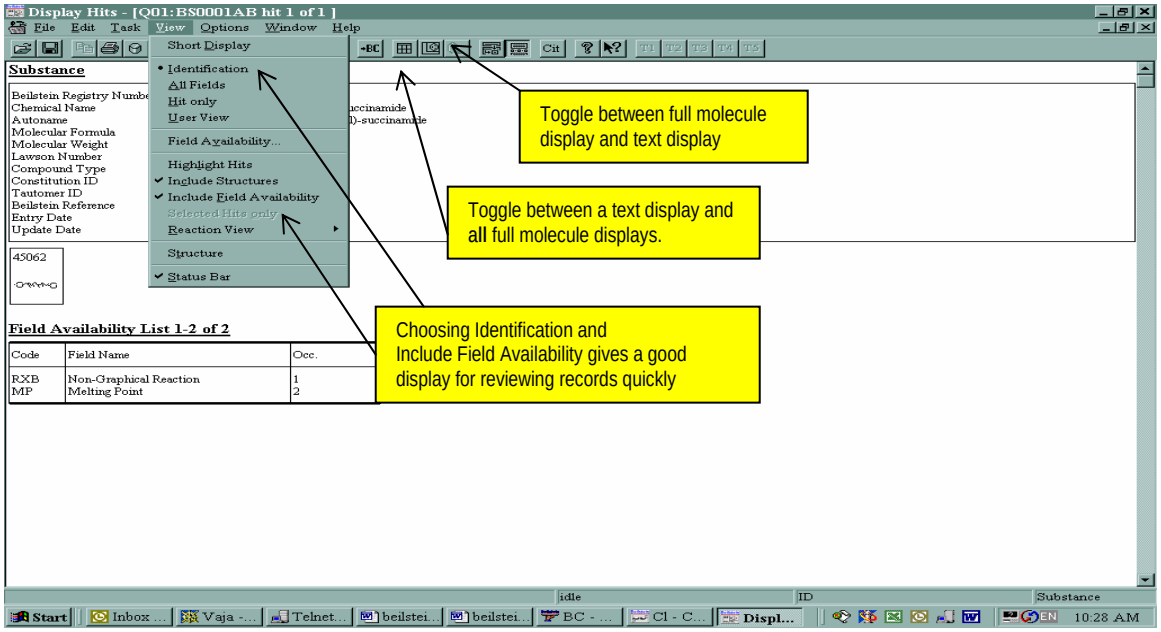
Next, under the **Query** Menu, click 'Structure Query Options'. This gives options for adding isotopic variants, radicals, ions, etc. A structure search can be combined a physical property data search (e.g. click Spectral Data (under Easy Data Search menu) and check NMR), in either Beilstein or Gmelin (not both).

Searching and Displaying …

Click the **Start Search** button. CrossFire will search the database and report how many compounds match the search criteria (i.e., how many "hits").



Search Results Click the **Display Hits** button to review the substances.



Display Hits Screen Basic Information

First (under the **View Menu**) click 'Short Display', to identify compounds of interest among a large set. Clicking the 'Field Availability' icon, or from under the **View Menu**, provides a quick display of available properties.

The 'All Fields'/'Field Availability' mode provides a convenient display for compounds with extensive properties. The 'Field Availability' display provides a scrollable property list that can be clicked to quickly display specific data fields.

To sequentially view the text records, use the center two arrows on the menu bar at the top of the screen. Use the outer two arrow buttons (the ones pointing to a vertical line) to move to either the first or the last hit.

Underlined numbers in various colors seen in the text records are hyperlinks to the Beilstein record for another compound, reaction or citation. To move to the linked record, click on the highlighted number. Moving back and forth between the original and linked records is done with the back and forward buttons (shown below).

'Reaction View' under the **View Menu** provides for display of 'substance as a reactant' or 'substance as a product'. To quickly search for preparations of a specific organic compound, do a compound search, go to the 'View' menu and click on 'Reaction View' and 'Substance as a Product'.

Reaction Searching

There are currently about 8.7 million structure searchable reactions in Beilstein and 1.5M reactions in Gmelin allowing you to build complete reactions or half reactions and search for them.

First draw the reactant and/or product molecules, in the Structure Editor. Then switch to 'Reaction Editmode' by clicking the '**Reaction Attr**.' Button on the tool bar.

The function bar then changes to allow reaction roles to be set. '**Select**' the appropriate molecule (with the select tool) and click either the Reactant or Product button to define that molecule's role in the reaction. Half-reactions can also be defined by drawing and defining only a reactant or a product.

The structure of a molecule can be changed by switching back to the Structure Editor, after '**Selecting**' the molecule, and using the '**Undefine**' button, to delete product and reactant assignments.

Click on the **BC** button to transfer the structure to Beilstein Commander, and click **'Start Search'**.

Search Tips:

Preparative reactions can be quickly displayed, after a compound search, from the 'View' menu → Reaction View → Substance as a Product.

Structures can also be copied into the Structure Editor from the **Short Display** under the **View** menu. First search for a model compound by name (e.g. Ferrocene). Highlight the 'short display' structure by clicking once and then from the **Edit** menu, select 'copy structure to SE'.

To move a structure -- **'Select'** the structure / drag the small black square
To expand or contract a structure -- '**Select**' the structure / drag the side of the 'box'.
To copy a structure -- select structure / shift key / drag

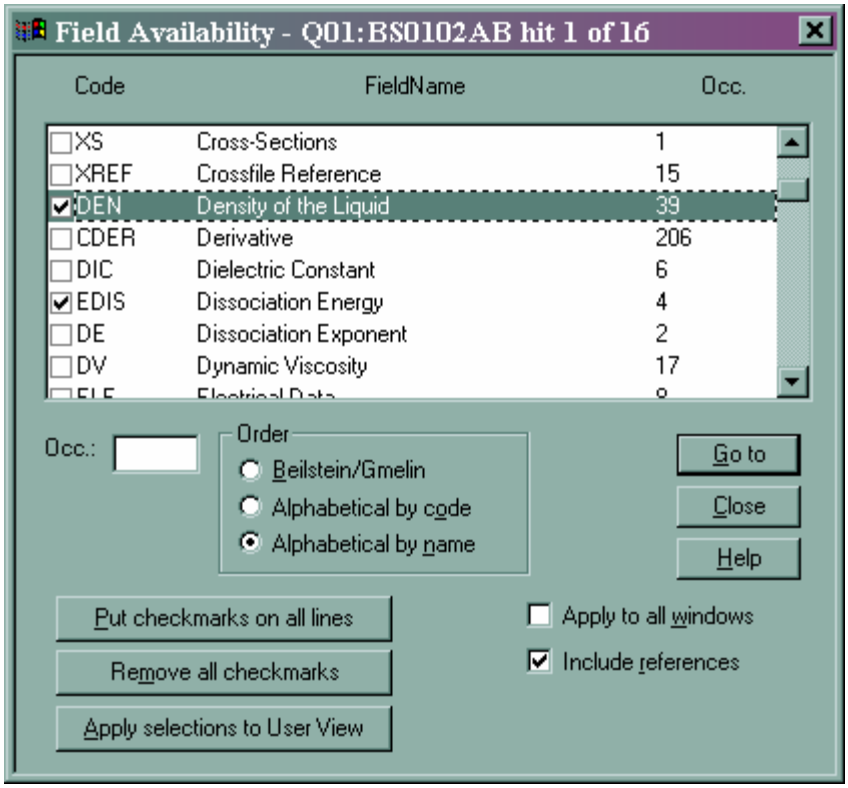
Beilstein Lawson Numbers define a structural similarity. After finding the Lawson Number for a model compound and then using it as a search term, results are restricted to substance classes without actually drawing a structure and searching for it.

CUSTOM DISPLAYS AND PRINTING

Records in the database can be hundreds of pages long. To avoid printing unwanted information, you can:

1. Use the **check boxes** to designate specific 'fields' in a record, and click on the **print icon**

2. Under the **View Menu** click both **All Fields** and **Field Availability**. The **Field Availability** list can be alphabetized (in Gmelin, however, the order is case sensitive). The default order corresponds to the print volumes (chemical properties, physical properties, etc.). For printing selected fields, click **Remove all checkmarks** and check off fields (as shown) and click **Apply selections to User View**

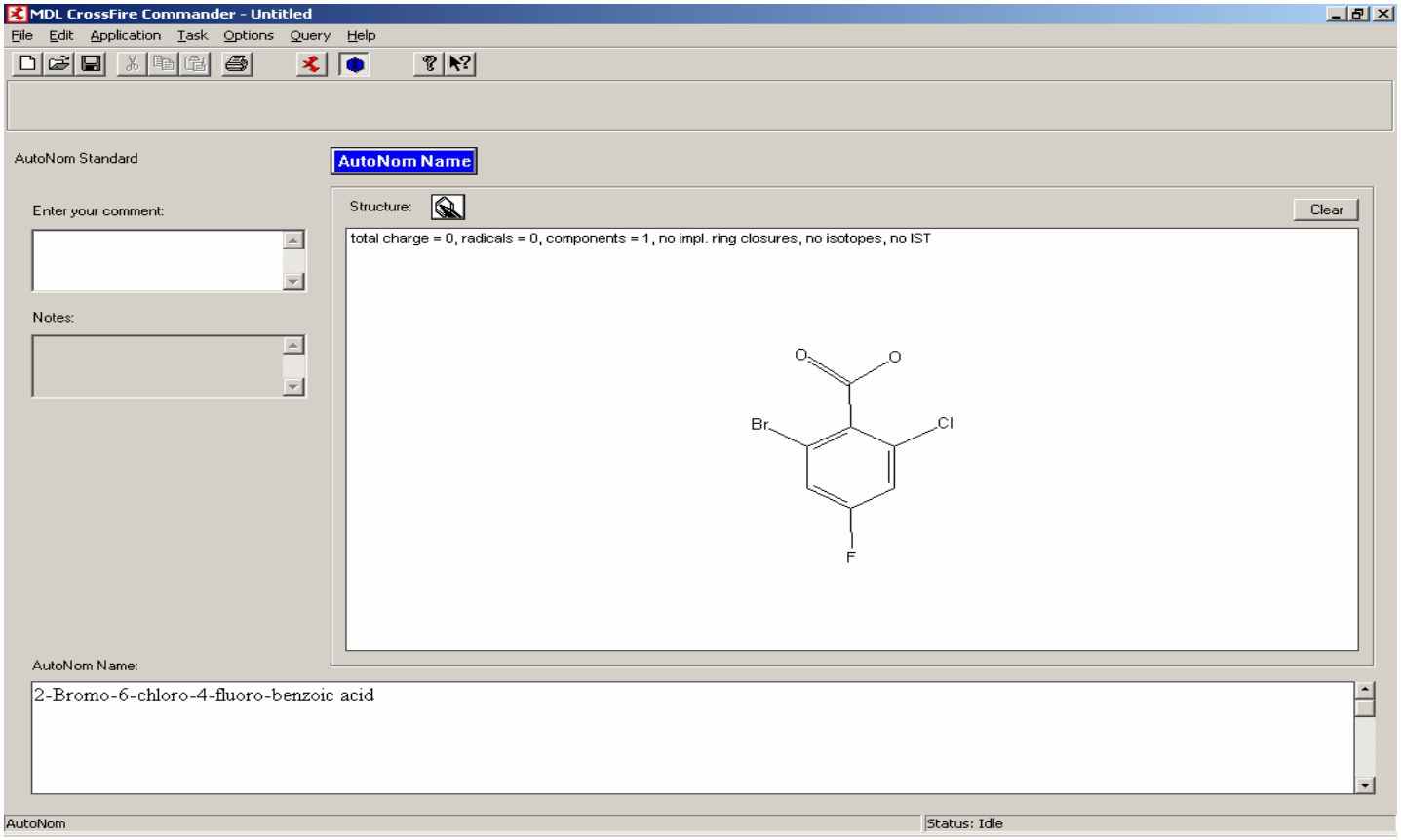


This limits the Display Hits window to just the desired fields, which can then be printed. The full display is regenerated by clicking **All Fields** under the **View Menu.**

3. Text can easily be copied and pasted into e-mail messages, MS Word and Excel documents. Structure images must be first transferred to the **Structure Editor**, and from the **Edit Menu**, click **Copy All** and in MS Word, paste special **Picture**.

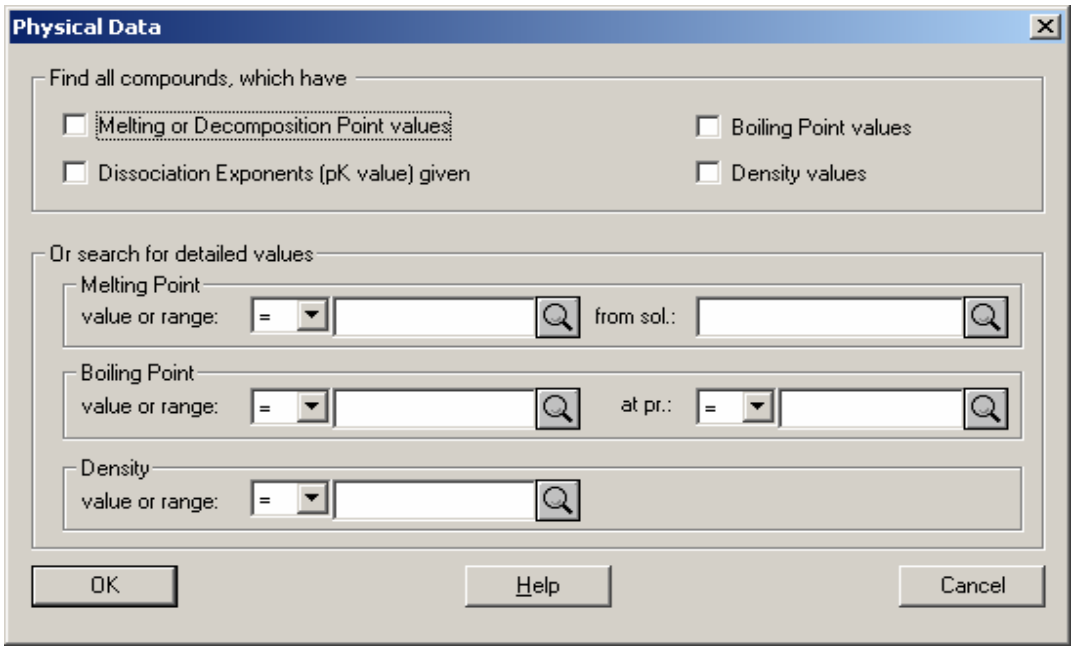
AUTONOM (Organic structure → IUPAC Name)

Clicking the blue 'benzene' ring on the 'Commander Screen' opens Autonom. Open the Structure Editor, draw a structure and return to the Commander Screen. Then click 'AutoNom Name'.



DATA (NAME / PROPERTY) SEARCHING

Data searching requires double clicking the Fact Editor or one of the quick search options (e.g. Physical Data). Beilstein display is shown. Gmelin offers additional options. The magnifier button opens a dictionary to check spelling, etc.



BEILSTEIN NOTES:

Some of the original print database was not translated into English. The 'Other Conditions' line in the Reaction Field, for example, is in German. A Beilstein German-English dictionary is available at:

http://www-sul.stanford.edu/depts/swain/beilstein/bedict1.html

In the creation of Beilstein CrossFire, the algorithm used to update the very cryptic abbreviations, used in the early years, assigned the current abbreviation to journals that have changed title over the years. For example, 'chem.ber.' is given as the abbreviation for Ber. Dtsch. Chem. Ges.(1868-1946)

Beilstein discontinued coverage of patents in 1980. Although a few scattered 19th and early 20th century patents were covered, the strength of Beilstein's patent coverage begins in 1925. About 2K patents were indexed in 1930, steadily increasing to 5000 in 1956 and to 15K in 1966. Coverage began to slowly decrease dropping to about 10K patents in 1979.

Chemical names in Beilstein generally reflect usage in the article. Since 1992, Chemical Abstracts Registry Numbers are not assigned, unless they appear in the article. It is advisable to rename a compound according to IUPAC rules (e.g. tert-butyl-methyl ether → 2-methoxy-2-methyl-propane) or, better yet, do a structure search.

Simply drawing a cis or trans structure will NOT define it for a structure search. The results will include both cis and trans compounds (the bonds are labeled E and Z in the Short Display). The wavy bonds refer to mixtures of E & Z isomers or a compound of unknown configuration.

Range searching is allowed. For example, you can search for compounds with melting points greater than 500 degrees (mp>500 in a Tabular Query). Searching for a melting point of 134 degrees will retrieve compounds with a melting point listed as 133-135 degrees.

Coverage of the literature is somewhat inconsistent in Beilstein. For example, Organic Syntheses has not been covered since 1979 (since it is not a journal). Coverage of SynLett did not begin until 1997, and coverage of Organic Letters did not begin until mid-2000. In general, articles are not indexed until about one year after publication.

In the mid-seventies, there was a new definition of "inorganic" and "organic" compounds. Beilstein currently indexes compounds that contain carbon and the yellow-shaded elements in the periodic table at: http://www.library.wisc.edu/libraries/chemistry/beilstein/beilsteinelements.htm The remainder are indexed by Gmelin.

The following are NO LONGER Beilstein compounds: CO, CS, CO2, CS2, COS, C3O2, C3S2, H2CO3, HCN, HOCN -- their sulfur analogs and metal salts/complexes, Dicyanogen (C2N2), Phosgene (CCl2O), Metal Carbides, Metal salts of formic, acetic and oxalic acid, Fullerene and Carboranes. 'Gmelin' compounds included in Beilstein are organometallic compounds known before 1980, such as Ferrocene, Dimethyl Mercury, etc. Searching both databases is strongly recommended for complete retrieval.

GMELIN NOTES:

1. For all compounds with molecular units (e.g. small molecules like H2O and coordination compounds) a structure search or substructure search is recommended. If the connectivity of ligands is not unique, coordination compounds may only be searched as fragments (e.g. a substructure of the ligand in combination with complexed metal).

2. For compounds without molecular units, e.g. solid state materials, salts (like NaCl), formula searching is the preferred. Note that H2O is included in component MF for hydrates. To reduce the hits, restrictions can be set, e.g. number of components. nc=1 is a single phase compound -- K4Fe(CN)6. number of fragments. nfrag=2 -- K and Fe(CN)6. number of elements. ne=4 -- K, Fe, C, and N.

3. Searching for chemical names is another possibility. However, not every compound has a systematic name, since only names given in the article text are entered. Furthermore, since chemical names of inorganic compounds are less standardized than organic ones, structure or formula searching is recommended. Searching with chemical names can be useful if a ligand is very complex to draw. Search for the ligand by name (with asterisks, e.g. *ferrocene*) to retrieve compounds containing this sub-structure. The desired structure can then be loaded into the structure editor, and altered as needed.

4. In contrast to Beilstein, Gmelin indexes polymeric compounds. These may be searched as structures, with the structure editor. For example, enter a Ge atom (with max free sites). Return to Commander and (in the fact editor) enter [type=polymer]. This will retrieve all Germanium polymers. Alternatively, you can attach two generic groups (e.g. Aryl or Acyl) to retrieve, for example, compounds with a [Aryl--Ge--Aryl] repeating structure (with a Ge backbone).

5. Polymers can also be searched by monomer molecular formula (e.g. C12H10Ge will retrieve polydiphenylgermane. In SFS, this would be searched as (C12H10Ge)x.

6. More than 99% of the original Gmelin Handbook thru 1975 is indexed in CrossFire Gmelin. Some material was added to Beilstein (a matter of re-definition) and some volumes (e.g. Geology of Iron) were omitted. Handbook volumes published after 1975 were not indexed for CrossFire Gmelin.

Beginning in 1976, Gmelin began indexing the 120 most important journals in inorganic and organometallic chemistry. There is currently a major effort to bring the database up to date, by adding material from 1995-2001. The GDCh is responsible for excerpting articles from 1995-2000 by the end of 2004. Coverage of new journals is problematic. Inorg. Chem. Commun. (1998+) will be indexed from 2003+ . Both Z. Kristallogr. New Cryst. Struct. and Acta Crystallogr. Sect. E are under evaluation.

BEILSTEIN & GMELIN'S RELATIONSHIP TO SCIFINDER SCHOLAR

SciFinder Scholar (Chemical Abstracts plus) and Beilstein/Gmelin have different criteria for indexing articles that describe preparative reactions and their products. Beilstein & Gmelin, while currently indexing a much smaller number of journals, are much more comprehensive in their coverage of the synthetic chemistry literature, since their primary focus is on chemical compounds, their physical properties and reactions. CAS, per its stated policy, focuses on 'new' information and the main point of the article.

CrossFire Documentation and Help:

MDL CrossFire Commander Version 6 (2002) (Caltech Splash Page)

MDL CrossFire Commander V6 -- Quick Reference Guide

CrossFire Beilstein Brochure

- CrossFire Beilstein Data Fields Reference Guide (v6)
- List of … journals covered in the Beilstein Database

CrossFire Gmelin Brochure

- CrossFire Gmelin Database Reference Guide
- List of … journals covered in the Gmelin Database

Autonom Brochure

For searching tips, consult these guides:

Beilstein and Gmelin Databases (UW-Madison)

- Quick Guide, Exporting, Fact Searching, Practice Exercises
- Reaction Searching, Saving Property Data from a Personal Computer
- Structure Searching, Beilstein/Gmelin CrossFire Workshop Outline

University of Chicago Web Guides for Beilstein CrossFire

- Version 6 (2003)
  - Guide to Beilstein CrossFire Fact Searching (PDF -- 2 Mb)
  - Guide to Beilstein CrossFire Structure Searching (PDF -- 2.2 Mb)

SciFinder Scholar Software Support

- EndNote

The Combined Chemical Dictionary

http://www.chemnetbase.com/scripts/ccdweb.exe

includes all compounds contained in the:

Dictionary of Organic Compounds (266,000 records)

Dictionary of Natural Products (181,000 records)

Dictionary of Inorganic and Organometallic Compounds (102,000 records)

Dictionary of Drugs (45,000)

Dictionary of Analytical Reagents (14,000 records)

The compounds indexed in the CCD can be generally characterized as follows:

- Fundamental **organic and inorganic compounds** of simple structure, including the elements, inorganic binary and ternary compounds (hydrides, halides, oxides, sulfides);
- Virtually every known **natural product** including those of unknown structure;
- All **currently marketed drugs**, including all those listed in generic name compilations (US Adopted Names, International Nonproprietary Names, British Approved Names, Japanese Accepted Names), as well as those undergoing clinical trials;
- Compounds with an established use such as **catalysts, solvents, starting materials, synthetic reagents, analytical reagents;**
- Important **co-ordination compounds**, e.g., amines, phosphines, alkoxy complexes, and major well-characterized bioinorganics;
- **Organometallic compounds** representative of all important structural types (in the case of ligands with organic substituents, typically the parent member of each series, where known, together with a selection of homologues);
- Important **biochemicals and minerals**;
- Other compounds of particular interest because of their chemical, structural or biological properties, including many **newly synthesized compounds of active research interest**.

The CCD is unique because it provides a wider variety of references, for the compounds it indexes, than Beilstein/Gmelin or SciFinder Scholar (e.g. Aldrich spectra catalogs; Fieser & Fieser's Reagents…; Ullmann's and Kirk-Othmer; Extra Pharmacopoeia; Bretherick's Handbook of Reactive Chemical Hazards; RSC's Hazards in the Laboratory; Sax's Dangerous Properties …; Browning's Toxicity and Metabolism …; Organic Syntheses (which Beilstein hasn't indexed since 1980); Encyclopedia of Reagents for Chemical Synthesis and references to book chapters).

Because of its currency and the indexing practices of Beilstein/Gmelin & SciFinder Scholar, it will also provide unique journal article references.

A web based tour is available at: http://208.254.79.30/chemnetbase/tours/ccd/index.html

Text Searching the Combined Chemical Dictionary

After entering the database, search with a compound name, molecular formula, physical property, etc. or preferably browse for search terms.

Browse Index is very strongly recommended, as it will give a perspective on the variety of index terms.

Click on the 'browse index' button, then type the first part of the required search term in the 'Index Stem' box and click 'Go To'. Index terms with that stem will be displayed. Alternatively, you can click on the alphabetical links above the Index Stem box and page thru the terms.

Each entry in the 'browse index' lists is hotlinked to the search box. Click on an term, go to the top of the page and click 'return to search'. The term will now be in the correct search box. Click on 'submit search' and then click on the compound name to display the record. Clicking on the 'Benzene' ring next to the compound name (Netscape) will display the structure.

The **Molecular Formula** 'Browse Index' first displays 'ALL' search terms for all compounds with a given element (e.g. -ALL-Ag is the search term for retrieving the 621 compounds containing Silver). Click on 'Next Entry' to display the next page which gives 'ONLY' search terms for compounds with only certain combinations of elements (e.g. -ONLY-C H P X is the search term for retrieving the 820 compounds that contain only Carbon, Hydrogen, Phosphorus and Halogen). These two unique Molecular Formula search term families (-ALL- and -ONLY- are followed by the molecular formulas of individual compounds in Hill order. Simple MF searches require Hill Order input (e.g. F4Xe).

Range searching for physical property data is allowed (e.g. For example, to search the Melting Point field for a melting point within the range of 47 to 51 degrees Celsius, key in 47 - 51. You can also use the following additional symbols:

> greater than < less than >= greater than or equal to <= less than or equal to

Truncation symbol is * , with ? (for one character).

References are searchable by author, patent assignee, journal title, etc.

Type of Compound is useful in searching for specific structural types of inorganic compounds, very general classes of organic compounds, or classes of natural products.

Chemical Name contains all chemical names and synonyms for each compound in the database. It includes trade names, generic names for pharmaceuticals, CAS names, trivial names and semisystematic names.

The **CCD Help File**: http://www.chemnetbase.com/help/ccdhelp/ccdindex.htm

Library workstations will have both the plugin & viewer installed.

STRUCTURE SEARCHING IN THE CCD

Select either Substructure or Exact Match, and click on **Draw Query**. This will display the **Structure Drawing Screen**.

DRAWING TOOLS (Left margin)

Pencil Tool	draw atoms and bonds
Ring Tool	draw 3-15 membered rings and
Chain Tool	draw chains of 1-30 nodes
Eraser Tool	deletes atoms and bonds one by one
Selection Tool	highlight single nodes and bonds, before using the QueryDef menu
Lasso Tool	selects part of a structure and drags it
Text Tool	add text to structure (not searchable)
Reaction Tools	add reaction arrows, reaction roles and sites (not searchable)

SPECIAL TOOLS (Along the top)

Short Cuts	functional groups (-NO2, -CF3, -CHO, etc.)
Variable Selection	include or exclude variable atoms (Metal, Halogen, etc.)
Query Attributes	bond type, hydrogen counts, charge, etc.
Preferences	chemistry (show hydrogens, etc.), colors, display, etc.
Help File	?

PULL DOWN MENUS (which expand on the special tools above)

You can draw double bonds, or common atoms other than carbon, by clicking on the atoms or bonds at the bottom of the screen. A single click is for one-time use, a double click for multiple uses (until it is double clicked back again). The DRAW pull down menu gives access to the full range of bonds and atoms.

After drawing a structure, click on the 'green arrow' to return to the Searching Screen.

Ignore the 'Query not defined' message and Click on Submit Search.

On the Summary Display Screen, click on the compound name to display the textual data.

Note the variety of references.

SCIFINDER SCHOLAR (Chem Abstracts)

- ## CAplus

Chemical Abstracts 1907+ -- plus cover-to-cover indexing of 1,350 key chemical journals since Oct. 1994+, -- plus bibliographic information from both 'electronic' articles and from US, EP, WO patents (including machine translated abstracts for JP and DE patents), prior to CAS indexing. Records from 1907-1966 do NOT include CA Index Terms or Registry Numbers. 'Research Topic' searching allows both keyword and Registry Number searching of Chemical Abstracts records. Keyword and 'substance identifier' searching is often faster than structure searching and avoids problems of trying to limit the retrieval from a structure search, which is often plagued by CAS' 'maximum retrieval' philosophy.

- ## Registry File

Chemical Abstracts Registry File 1957+. Currently lists nearly 19M organic and inorganic compounds and over 16M biosequences, and provides links to 'Commercial Availability'.

- ## ChemCats

Commercially available chemical compounds and suppliers. Requires a search of the Registry File, which provides the link to ChemCats/ChemList. First search 'Chemical Substance or Reaction' (it will be much faster if you have a CAS Registry Number) although names or molecular formulas can also be used. Clicking on the 'Commercial Sources' hotlink provides a list of suppliers, catalog date, order number, product name and description. Suppliers must be contacted directly for current prices.

- ## CA Patent Family Information

Chemical Abstracts Patent Concordance/Index (1967+). This database is only available after searching for a compound, 'getting' the references and clicking on the 'microscope icon' for a patent record and then clicking on the 'Patent Family Information' hotlink.

Since CA generally abstracts European, German or WO applications first, this is a good way to find U.S. equivalents, as is the Esp@cenet Worldwide database at: http://ep.espacenet.com/

It is not currently possible to directly access full text patents from the Patent Family record.

- ## CASReact

Reaction information from the organic/organometallic sections of Chemical Abstracts. Journal articles 1985+, patents 1991+).

This database is only available while performing a structure search. Searching is based on the concept of 'maximizing retrieval', which often makes it impossible to limit a search to a specific compound.

- ## Regulated Chemicals Listing (CHEMLIST)

CHEMLIST was originally built from data in the 1985 TSCA Inventory and supplements, and the Federal Register for 1978-1987. Major sources for the TSCA updates are listed below.

Other sources are the European Inventory of Existing Commercial Chemical Substances (EINECS), the European List of Notified Chemical Substances (ELINCS), the Canadian Domestic Substances List (DSL) and Non-Domestic Substances List (NDSL), the 1996 Korean Existing Chemicals List (ECL), and the Japanese Existing and New Chemical Substances list (ENCS. The Federal Register (scanned daily), Chemical Regulation Reporter, EPA Chemicals-in-Progress Bulletin, NTIS TSCA Inventory Tape (semi-annual updates), Pesticide Toxic Chem. News (from 1989) are scanned for new data.

The file covers 1979 to the present, with more than 214,715 records (4/2000). It is updated weekly with more than 50 additions to existing records or new substances

- **ChemPort Links**

Many of the patent links are to the esp@cenet database, which first displays the front-page data and abstract. A second click is required to view the full text of the patent (if available).

Journal article links to ACS journals retrieve the full text directly.

SCIFINDER SCHOLAR SEARCH TIPS

Structure Searching -

The SciFinder Scholar ‘structure searching’ software is designed to maximize retrieval, making it often difficult to quickly retrieve information on a specific compound. Searching for Phenyllithium, for example as an exact match, retrieves 69 substances. Many are addition compounds or salts, requiring you to page thru the structures to find the ‘real’ Phenyllithium.

If your search retrieves more than about 200 compounds, you can click on 'Analyze Substances' and then change the 'radio button' to 'Precision' and click 'OK'. This gives a histogram of the of the types of substances. Click on 'Conventional …' and 'Get Substances'. This will reduce the display to the compound, its isotopic variants and its multicomponent substances. Another option, for well known compounds, is to click ‘Refine Substances’ and limit to 'commercial availability'.

In contrast to Beilstein, the default structure in SFS is open (i.e. allows maximum substitution). Use the 'Lock Out' tools in the left side tool box to prevent further substitution on selected atoms or ring fusion …

Medline Duplicates - to eliminate duplicates (default search is a combined CAplus/Medline search), go to the view menu and select title order. Duplicate references will appear next to each other, however the list is now in alphabetical order by title.

Chemical Name Searching - CA Index Names (and some synonyms) can be used as search terms (and can sometimes be truncated). For example, a search for Isopropanol labeled with O18 requires either the CA Index Name: 2-Propanol-18O or Isopropyl alcohol-18O.

SciFinder SCHOLAR Solutions

http://www.cas.org/SCIFINDER/SCHOLAR/SOLUTIONS/index.html

SciFinder Scholar Solutions gives tips for:

Designing structures for substructure searching,
Finding compounds containing specific structure fragments,
Copying, pasting, and saving structures,
Importing data from SciFinder Scholar to EndNote, etc.

Finding information on Regulated Chemicals or Commercial Sources:

http://www.cas.org/SCIFINDER/SCHOLAR/chemlist.html

- Substructure Searching Overview
  - Controlling Substitution
  - Controlling Ring Formation
  - (Substructure) Decision Chart
- Searching for Synthetic Pathways Overview
  - Full Reaction Searching (organic structural transformations)
  - Searching for Preparations
  - Functional Group Searching
  - Searching for Reaction Mechanisms
- Polymer Searching Overview
  - Polymer Formula Searching
  - Formatting Polymer Formulas
  - Structure Searching for Polymers
  - Searching for Polymer Classes

The Challenges with Substance Databases and Structure Search Engines
Helen Cooke, Damon D. Ridley
Australian Journal of Chemistry, 2004, 57(5), 387-392
http://www.publish.csiro.au/?act=view_file&file_id=CH03315.pdf

An explanation of the problems caused by the variety of structural representations for non-valence bonded substances in SciFinder, Beilstein, Gmelin, etc. Including: alloys, catenanes, polymers, and salts. There are also issues of definition (such as when a substance is a co-ordination compound, or hydrate, or salt), of bonding types (resonance, donor complexes and pi-complexes) and further problems with: allotropes, isotopic forms, physical states, etc.

Beilstein/Gmelin vs CCD vs SciFinder Scholar

SciFinder Scholar (Chemical Abstracts Plus) has completely different criteria for indexing articles than Beilstein/Gmelin or the Combined Chemical Dictionary.

Chemical Abstracts, per its stated policy, covers the full range of the 'chemical' literature, focusing on 'new' information and the main points of the article or patent, etc. This differs significantly from B/G and the CCD which index a small subset of chemical journals, focusing on the synthetic chemistry literature. Their primary focus is on compounds, physical properties and reactions.

Thus, the indexing of documents (SFS) vs. indexing compounds (B/G & CCD) obviously leads to significant differences in retrieval. In addition, the extensive coverage of the patent literature in SFS (which Beilstein & Gmelin dropped after 1979, and the CCD covers very selectively) strongly suggest that compound searches be performed in all three databases.

SciFinder Scholar's Registry File contains over 43M substances, while Beilstein/Gmelin have indexed 9.2M organic and 2.1M inorganic/organometallic compounds respectively. The CCD only lists 486K organic compounds and natural products and 102K inorganic/organometallic compounds. A comparison of results from the four databases highlights the need to search all three databases. For example:

Lithium, [2-(dimethylamino)phenyl]-(SFS) [2-(Dimethylaminophenyl)] lithium (SFS,CCD)
2-Dimethylaminophenyllithium (B) 2-(dimethylamino)phenyllithium (G)
C8H10LiN -- 22608-37-3

Beilstein: 5 references (4 as a reactant and 1 preparation) from 1977-1997.
The unique Beilstein 'preparation' reference is actually a synthesis 'prediction'.

Gmelin: 15 references as a reactant (including a German patent) from 1943-1990

CCD: 6 references (5 preparation and 1 use) from 1966-1990, including a book chapter.
3 of the preparation references are unique to the CCD.

SFS(CA): 36 references, from 1969-2004, of which 9 describe preparation. The Beilstein 'synthesis' reference is not included in SFS, since the main point of the article was the reactivity of the starting materials. 5 of the preparation references are unique to SFS.

Brian M. Stoltz (as an author)

Beilstein: 28 references (including 2 with Stoltz misspelled as Stolz. The most recent article is JACS (2004), 126, 24-25. Beilstein did not index (TL 1999, 40, 2061).

Gmelin: 15 references (all in Beilstein). The most recent being JACS (2004), 126, 24-25

SFS(CA): 73 references, of which 44 were synthesis articles (including all the references in Beilstein and Gmelin). In addition, there are references to a thesis, a patent, ACS Abstracts -- as well as 7 articles more recent than JACS (2004), 126(1), 24-25 (e.g. JACS (2004), 126(46), 15044-15045.

Predefined Atom Lists

A	Any atom
AH	Any atom or Hydrogen
Q	Any atom except Carbon or Hydrogen
QH	Any atom except Carbon
M	Any metal atom
MH	Any metal atom or Hydrogen
X	Any Halogen
XH	Any Halogen or Hydrogen