

## CHAPTER

### 1

## SEARCHING THE LITERATURE

It is a good idea to review some of the experimental principles and techniques that were covered in the organic teaching laboratory before undertaking the more sophisticated and less pre-defined activities of the research lab. One of the best resources for this type of information is *The Organic Chem Lab Survival Manual* (Zubrick, 2003). Awareness and observation of all of the best chemical safety practices is also essential. An excellent resource for this type of information is *Prudent Practices in the Laboratory* (Committee, 1995). Specific comments concerning chemical safety are made at some places in this book, but safety always must be foremost in the mind of the experimenter.

When aiming to obtain a particular molecule, a good appreciation of how it has been prepared in the past is essential. Electronic data retrieval tools are ideal for finding this information. The simplest way to obtain a compound is to buy it, of course.

### 1.1 COMMERCIAL AVAILABILITY

The SciFinder® program is a good tool for comprehensive searching of commercially available compounds. It includes the information that is published in *Chemical Abstracts*, but makes it much more accessible. This browser-like application permits searching by chemical structure, either as an exact match or based on substructure. While a comprehensive description of how to use SciFinder is not intended here, a few points should be kept in mind. Structures can be drawn in SciFinder itself, or in another chemical drawing program and pasted into the structure window in SciFinder. To avoid an unmanageable number of hits when searching by substructure, it may be necessary to limit the structure based on atoms that may be further substituted or rings that may be added. These parameters

*The Synthetic Organic Chemist's Companion*, by Michael C. Pirrung  
Copyright © 2007 John Wiley & Sons, Inc.

must be set within SciFinder itself, with the {Lock out substitution} and {Lock out ring fusion} commands. If a compound identified in a structure search is commercially available, a small button with an orange flask will appear at the top of the compound window. A refinement of the hits from the initial structure search can also be based on their commercial availability.

For compounds of interest that show the small orange flask button, clicking it will open another window that lists the companies selling the compound and the quantities available. This last point is very important. Quite a number of compounds have been registered with *Chemical Abstracts* by small companies aiming to sell compound libraries for biological screening. These compounds, while strictly meeting the definition of commercial availability, are not in the same sense articles of commerce as are compounds sold by major reagent suppliers like Aldrich and Fluka. Information on less well known suppliers (phone, address, URL) is available by clicking a link next to each supplier listing.

Another way to search for commercially available compounds is with the ChemACX® program running on a MSWindows-based computer. A CD with a current compilation of suppliers' catalogs is needed for this application, for which periodic updates can be purchased. ChemDraw® is used to draw structures for this application. Further information on this program is available from the ChemACX manual.

It is always dangerous to refer to a Web-based resource in a book that should have long-term value, since Web pages seem to come and go at a rapid pace. However, some equivalent resources may arise to take their place. The case at hand involves the Web site of the Sigma-Aldrich Company, which is an excellent free resource to use in locating commercially available compounds. From the main page, a few clicks brings the browser to a search page. Searching by structure is among the options there. Chemical structure drawing on Web pages works fine, provided that the plug-ins are installed in your browser, but can be clunky. Better in this case is to enter the structure as a SMILES string. Many approaches have been taken to manipulate chemical structures electronically, with one basic machine-readable nomenclature being the SMILES format. SMILES stands for Simplified Molecular Input Line Entry Specification. The SMILES description of the structure sought can be obtained from

a drawn structure in a program like ChemDraw using the command {Copy as SMILES}. Upon pasting the SMILES string into the Sigma-Aldrich search page and loading it, the structure appears. It can be searched for using several options available on the Web site. A nice aspect of this search is that it encompasses all of the Aldrich sister companies, including Sigma and Fluka. Aldrich is a company that has always provided excellent service to chemistry, and many chemists keep an Aldrich catalog at their desks because of the wealth of data it includes on all of the compounds listed.

## 1.2 LITERATURE PREPARATIONS

If a compound cannot be purchased, there may be a known preparation of it. It is almost always preferable to at least begin with a method that others have used to prepare a compound, rather than trying to invent one from scratch. Identifying the specific publications in which a compound has been *prepared* among many more papers in which it has been mentioned in some way is straightforward when using SciFinder. The {Refine} command allows those sources to be selected. If a particular method used to prepare a compound is sought, it may prove useful to search by reaction rather than by compound. This is done simply—multiple structures are entered into the chemical structure search window, and a reaction arrow is drawn from the starting material(s) to the product. SciFinder will then label the molecules that it understands to be the reactant(s) and product(s). Searching for this reaction can next be initiated. The number of hits is often quite large, but adding qualifiers can reduce them to a manageable number. For example, reactions can be selected by applying the {Refine} command based on yield, or the number of steps, or the general type of chemical transformation. Another option that often reduces the possibilities is to identify which atoms of the reactant correspond to the same atoms in the product. This may seem obvious to chemists, but it is not obvious to this software.

A preparation of a compound that was conducted in one's own laboratory is most often preferred as a starting point for today's preparation. In some cases the chemist who actually did the prep may be available for consultation, or at least his or her research

notebook pages can be consulted. A properly kept research notebook (see Chapter 7) is almost always more useful than any literature preparation because it gives details that never appear in a publication, such as actual chromatograms and spectra, drawings or pictures of apparatus, and properties of chromatographic or distillation fraction.

The best literature methods will be found in the compilation *Organic Syntheses*. These procedures are generally aimed at synthesizing compounds on a fairly large scale (grams or larger, rather than milligrams). They are distinguished from essentially all other literature preparations by having been checked in the laboratory of another experienced synthetic chemistry group. The little details that can be important to success can be incorporated into the published procedure through the experience of the “checkers” in doing the experiment targeted for today, reproduce someone else’s preparation. The only drawback of *Org. Syn.* preps is that so few of the known compounds one might need have been through its rigorous review process. Even if the exact target compound is not in *Org. Syn.*, a prep of a related compound might provide a good starting point if today’s target is not too different structurally. Structure, keyword, title, author, registry number, molecular formula, and chemical name searching of the compilation are available on the Web site for *Org. Syn.* at <http://www.orgsyn.org/>.

Another source that is focused on preparative chemical procedures is the Synthetic Pages Web site. It also provides information not usually found in journal articles, such as troubleshooting tips, frequently encountered problems, and known variations with scale. It offers the virtue that it is interactive, with continual updates and comments from users. Its main limitation is that as of October 2006 there were just over 220 procedures in the database. This free database is found at <http://www.syntheticpages.org>.

Next preferred for the preparation of a desired compound is one that has been described in a “full paper” (one that gives experimental details, as in *J. Org. Chem.*). These experimental details are increasingly being relegated to the electronic/Web version of the journal, often called Supporting Material or Supplementary Information. This material is often available to all, even those without electronic access to the main journal. The techniques for searching for this literature include *Chemical Abstracts* as described above, as

well as specialized sites maintained by each publisher for their own journals.

A particularly useful Web site to locate information concerning specific reactions is the *Encyclopedia of Reagents for Organic Synthesis*, in its online form, e-EROS. It allows searching of a database of more than 50,000 reactions and more than 3000 of the most commonly used reagents by chemical structure and substructure, reagents, conditions, and reaction type. The URL is <http://www3.interscience.wiley.com/cgi-bin/mrwhome/104554785/>.

Least informative concerning the preparation of a target compound are the so-called communication or letter publications, which do not include experimental details. In some cases experimental descriptions may be provided in Supplementary Information. If not, the author might provide some detail about reaction conditions in the written narrative or on the reaction schemes. It is always worth contacting the senior author to find out if experimental details might be obtained. It is often easy to find and contact people electronically. Still, the novice experimenter may have little to go on, and reproducing results from these types of papers is widely regarded as difficult at best.

Finally, it should be recalled that the electronic versions of journals are in some cases only available for relatively recent editions. Electronic coverage of compounds and reactions earlier than that time cannot be assured. Thus some traditional book-based literature search methods may be required if the reaction of interest has a long history.

