Current Comments®

EUGENE GARFIELD
INSTITUTE FOR SCIENTIFIC INFORMATION®
3501 MARKET ST., PHILADELPHIA, PA 19104

New Compounds and Reactions Are As Close As Your PC with the *Index Chemicus* and *Current Chemical Reactions Personal Databases*

Number 9 March 2, 1987

Over the past decade I have described microcomputer software systems, such as *Sci-Mate®*, that have solved the problems of managing the storage and retrieval of bibliographic or clinical case histories in a personal database.[1,2] Whether it is reprint files, laboratory notebook entries, or other texts that need to be stored, they are personal in that you the user created the information or obtained the data from existing databases such as *SciSearch®* or *Medline*.

While I have encouraged *Current Contents®* readers to jump on the pc bandwagon, I have cautioned that it could involve some culture shock. If that is true for the "simple" problem of storing and retrieving bibliographic or clinical records, imagine the problems involved in graphically displaying and manipulating complex chemical structures and reactions.

Software systems have recently become available that solve these difficult problems. To take advantage of these sophisticated systems, ISI® has developed *Index Chemicus®* (*IC®*) and *Current Chemical Reactions®* (*CCR®*) *Personal Databases*, a series of diskettes covering topics useful for teaching and research in both academic and industrial environments. Derived from the extensive files contained in ISI's *Index Chemicus* and *Current Chemical Reactions*, these topical databases provide you with direct pc access to recently synthesized chemical compounds or new synthetic methods that have been developed in laboratories around the world. In addition, you can manipulate the information in our databases to meet your individual information needs.

IC Personal Databases are searchable by two chemical database management systems, Molecular Design Limited's *ChemBase* and Scott Gould's *ChemSmart*, while *CCR Personal Databases* are readable only with *ChemBase*. The choice of software systems is dependent upon your needs. *ChemBase* has more industrial applications, while *ChemSmart* matches the academic's needs. Using these software systems, you can use our databases to build chemical structure or reaction files as well as search for specific structures, substructures, and text from the files.

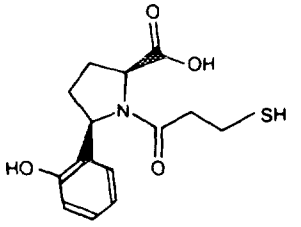
Index Chemicus Personal Databases

With over seven million chemical compounds reported in the literature, the potential chemical structure and synthesis data are staggering. Fortunately, *IC* and *CCR Personal Databases* help search for active compounds and viable synthetic pathways.

ISI has indexed over four million organic compounds since 1962. About 200,000 compounds reported as new in chemistry journals are indexed each year. With the new *IC Personal Databases*, ISI provides 16 subsets of these extensive files. Each database topic, produced on a floppy diskette, contains 200 selective compounds derived from the most recent literature, providing the best overview of the classes and types of compounds in each topic area.

Table 1 lists the 16 database topics currently developed for *IC Personal Databases*. ISI provides quarterly updates for *ChemBase* and annual updates for *ChemSmart* of selected topics so that you are informed of the most recent compounds developed in your specific field of interest.

Figure 1: Sample record of a compound from the *Index Chemicus® Personal Databases.*



ID:	1
ICN:	387095-001
	(C) 1986 ISI
	C14 H17 N O4 S
MW:	295.36

ACN:	4A	TN:	ANTIHYPERTENSIVES/HYPOTENSIVES
NAME:	PROLINE, 5-(2-HYDROXYPHENYL)-1-(3-MERCAPTO-1-OXOPROPYL)-		
AU:	KATO E, YAMAMOTO K, KAWASHIMA Y, WATANABE T, OYA M, ISO T, IWAO J		
SO:	CHEM PHARM BULL TOKYO 33(11),4836-46 (1985)		
NOTE:	ANTIHYPERTENSIVE (TESTED)		

Each compound record contains the compound name, molecular formula, molecular weight, biological activity (if indicated), journal source, and other information. Figure 1 shows a typical record of a compound from the *IC Personal Databases*.

Current Chemical Reactions Personal Databases

Updates on recently developed chemical compounds are only part of a chemist's needs. The pathway used to produce a compound is often more informative than the end product itself. *CCR* was developed as a monthly indexing service to alert you to new or modified synthetic methods. The *CCR Personal Databases* provide six topics, listed in Table 1, covering recent synthesis methods or modifications of particular reactions. Each topic contains 100 reactions, providing a broad overview of that field. An individual record describes the reaction type, catalysts, solvents, reagents, experimental conditions, yields if indicated, journal source, and other information. Quarterly updates are available for selected topics. Figure 2 shows a sample *CCR Personal Databases* record.

Table 1: Topics for *Index Chemicus®* and *Current Chemical Reactions® Personal Databases.*

Index Chemicus Personal Databases

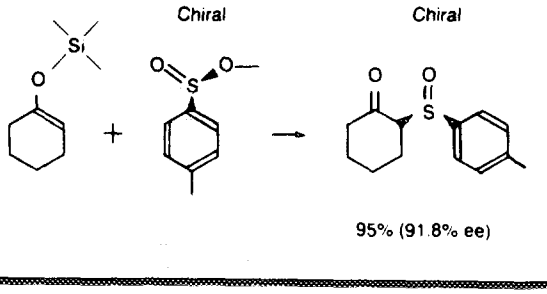
- Amino Acids/Peptides
- Antibiotics
- Herbicidal Compounds
- DNA, RNA, Protein Synthesis Inhibitors*
- Anti-inflammatory Agents
- Antihypertensives/Hypotensives
- Heterocycles
- Macrocyclic Compounds
- Natural Products and Derivatives
- Introduction to Organic Chemistry*
- Unisolated Intermediates
- Nucleosides/Nucleotides
- Organoboron Compounds
- Organometallic/Organosilicon Compounds
- Prostaglandins
- Steroids/Antifertility Compounds

Current Chemical Reactions Personal Databases

- General and Name Reactions
- Asymmetric Synthesis*
- Catalysis
- Synthesis of Natural Products and Derivatives
- New Reagents*
- Protecting Groups*

* Quarterly updates for 1987 are not available for these topics.

Figure 2: Sample record of a chemical reaction from the *Current Chemical Reactions® Personal Databases*.



ID:	1
CCRN:	32169
	(C) 1986 ISI
RCS:	BF3, ET2O, THF
CND:	12-14 HR, 0 DEG C
TN:	ASYMMETRIC SYNTHESIS
AU:	HIROI K., MATSUYAMA N
SO:	CHEM LETT (1),65-8(1986)
NOTE:	EXPLOSIVE RXN

Advantages

The printed versions for *IC* and *CCR* are still necessary for browsing and current-awareness purposes. Libraries will also store them for archival purposes. But tabulating lists of compound names and associated compound activities does not highlight the relationship between a compound's structure and its activity. By collecting the information from the *Personal Databases*, however, you can develop an updated, more comprehensive, or personalized database. These structures in turn can be used as templates to develop more complex molecules.

Once you have created a database, you can use certain computer-driven techniques to analyze the data. One method is to sort the database file for activity and then examine the structures found by the search. Another method is called substructure searching—a technique that allows you to identify those compounds with particular fragments of a structure. Reviewing and comparing the activity of these substructures from different compounds helps you to visualize new active compounds or envision unique synthetic pathways. Daniel Meyer, manager, Online Development, Chemical Information Division, ISI, and Peter Cohan, project manager, Molecular Design Limited, San Leandro, California, note that "the data manipulation and display facilities of the database thus highlight a possible significant correlation, which was not readily apparent in the original journal article tables."[3]

The easy-to-read format of the compound and reaction records of the *Personal Databases* makes them ideal for pc browsing. They are excellent tools for you to scan the most recent compounds or reactions to ensure that you are not duplicating work already done in another laboratory. Furthermore, the *Personal Databases* offer the advantages of an online chemical database without the cost of associated hardware and service charges. And data stored on diskette offer both flexibility and greater security of sensitive information.

Conclusion

Our Chemical Information Division has utilized new computer systems and technology to offer important data on new organic compounds and new synthetic methods in a variety of formats—print, microfilm, microcomputer, and mainframe computer—to meet the information needs of the chemical community. We discuss some of

these products briefly in the appendix. In the near future, we will devote an essay to our mainframe products—*Current Chemical Reactions In-House Database* and the *Index Chemicus In-House Databases*.

The advent of the computer has revolutionized the methods used for data analysis, storage and retrieval, and manipulation of chemical information. The new sophisticated technology can only enhance the progress and pace of chemical synthesis research. ISI's *Personal Databases* provide a convenient and easy-to-use means for harnessing that portion of new chemical information that reflects your particular research needs.

* * * * *

My thanks to Lisa Holland and Daniel Meyer for their help in the preparation of this essay.

REFERENCES

1. **Garfield E.** The integrated *Sci-Mate Software System.* Part 1. Combining file management, online access, and searching with manuscript editing. *Essays of an information scientist: ghostwriting and other essays.* Philadelphia: ISI Press, 1986. Vol. 8. p. 360-7.
2. ----------. The integrated *Sci-Mate Software System.* Part 2. The *Editor* slashes the Gordian knot to conflicting reference styles. *Current Contents* (11):3-10, 17 March 1986.
3. **Meyer D & Cohan P.** Designing new compounds with PC databases. *Amer. Clin. Prod. Rev.* 5(12):16-9, 1986.