Environmental Uses of the NIH-EPA Chemical Information System

G. W. A. Milne, C. L. Fisk
Stephen R. Heller, Rudolph Potenzone, Jr.

Summary. A publicly accessible computer system for chemical information has been developed jointly by a number of agencies of the U.S. government. The system contains spectroscopic, crystallographic, toxicological, and regulatory data for more than 200,000 chemicals. The entire data base may be searched for a particular chemical structure or substructure, whose properties may then be retrieved. Alternatively, searching with numeric properties data is possible, permitting the identification of chemicals. Access is by local telephone call, and the system is used on a fee-for-service basis by organizations in over 20 countries. An important application of the system is to problems of chemical pollution.

With continuing reports of chemical spills, explosions, and illegal dumping of toxic wastes, access to reliable information on chemicals is essential. Similarly, research workers must be able to rapidly retrieve desired data from the mass of chemical information. This article describes a publicly available information system that has been relied on by federal, state, and local governments and the chemical industry for the past several years.

This system, the NIH-EPA Chemical Information System (CIS) (*1*), can be used to identify chemical substances and to obtain physical, chemical, and regulatory information about them. It is used by both the research community (*2*) and by those engaged in environmental work (*3*) and is supported jointly by the National Institutes of Health (NIH) and the Environmental Protection Agency (EPA) in collaboration with the National Bureau of Standards (NBS), the Food and Drug Administration (FDA), the National Institute for Occupational Safety and Health (NIOSH), and the United Kingdom's Department of Industry. More than 60 percent of the over 500 organizations using CIS are industrial chemical corporations.

Structure of CIS

The CIS, which has been under continuous development since 1971, consists of a collection of disk-stored data bases and a battery of computer programs for interactive searching through them (*4*). In addition, CIS has data referral capability (*5*) and a data analysis software system (*6*). It can be thought of as having five main areas.

1) *Numeric chemical and physical properties data bases.* These deal mainly with spectroscopic or diffraction data. They can be used as sources of data when given the identity of a compound. A more significant use of these files, however, is for identifying unknown compounds from experimentally measured data. The programs in CIS permit either of these activities.

The mass spectral search system (MSSS) (*7*) contains over 38,800 evaluated mass spectra. It is accompanied by smaller files of carbon-13 nuclear magnetic resonance (NMR) data, containing about 10,000 spectra (*8*) and infrared spectral data. The infrared spectral data base contains some 3000 spectra measured in the gas phase with Fourier transform techniques. The Cambridge Crystal data base (*9*) has full atomic coordinate data on all organic molecules that have been so analyzed—some 25,000 entries. As a complement to this, CIS carries the NBS single crystal file, also prepared by the Cambridge group. This contains the crystal lattice parameters for some 65,000 molecules and is useful for identifying a compound from its crystalline cell dimensions. There is also a file of 34,000 powder diffraction patterns (*10*). This data base, maintained by the Joint Committee on Powder Diffraction Standards, is very useful for identifying organic and inorganic materials. A particular strength of this system is its ability to identify the components of a mixture from the aggregate diffraction pattern. A data base of thermodynamic properties, collected and evaluated by NBS, is being installed in the system, and there is also a pilot version of a data base of nucleic acid sequences.

2) *Regulatory and toxicological data bases.* The major regulatory data base in CIS is the *Federal Register* search system (*11*). This contains all notices dealing with chemicals which have appeared in the *Federal Register* since the beginning of 1978. The data base is updated on a weekly basis and is useful for identifying actual and potential regulatory actions that apply to specific chemicals or groups of chemicals.

The registry of toxic effects of chemical substances (RTECS), which is maintained by NIOSH, can be searched in CIS (*12*). It is used to retrieve toxicity data for specific compounds or to find chemicals with particular patterns of toxicity. There are several files related to RTECS in the system. For example, a data base (*13*) containing the composition of readily available consumer products, such as household chemicals, can be searched on the basis of chemical ingredient, trade name, manufacturer's name, and so on. The toxicity of individual chemical ingredients is also provided.

The CIS contains a file of information on the production site and volume for the approximately 30,000 nonconfidential entries in the commercial chemicals inventory compiled by EPA under the Toxic Substances Control Act. Another file developed and maintained by EPA is the oil and hazardous materials file (OHM-TADS), which provides data useful in responding to chemical spills. EPA has also assembled a data base on chemicals that have been identified as pollutants in U.S. and foreign water systems (*14*). The file contains information from 10,593 reports involving about 1,700 chemicals.

3) *Structure and nomenclature search*

G. W. A. Milne is a research chemist at the National Cancer Institute, Bethesda, Maryland 20205. C. L. Fisk is a staff fellow at the National Institutes of Health, Bethesda, Maryland 20205. Stephen R. Heller is a physical scientist and Rudolph Potenzone, Jr., is a systems analyst at the Management Information Data Systems Division, Environmental Protection Agency, Washington, D.C. 20460.

0036-8075/82/0122-0371$01.00/0

system (SANSS). This system contains the names (including English and foreign-language synonyms and trade names), structure, and formula of some 200,000 different chemicals. Thus the user has a variety of ways in which to search for a chemical (*5*). With SANSS it is possible to find all compounds in the data base which meet certain criteria, such as association with a particular name, molecular formula, or structural or substructural fragment. Also, the user is informed of every numeric data base in CIS which carries information on a given compound. The user may then refer to the files that are identified to retrieve and examine the appropriate numeric data.

4) *Data base referral*. Every chemical in SANSS is identified in terms of its Chemical Abstracts Service (CAS) registry number. Once a compound has been retrieved by SANSS the user is informed of the CIS files that have data pertaining to this compound. With the registry number, however, SANSS can also pass on to the user the names of files external to CIS which contain information on the compound of interest. The user may learn then that a specific chemical, in addition to being represented in MSSS, is also in the *Merck Index* and on EPA's list of registered pesticides. These "secondary" sources are not carried on-line by CIS; they are merely offered to the user, who chooses whether or not to consult them.

The SANSS can compare the chemical content of a file, such as the *Merck Index*, with that of another file, such as the Toxic Substances Control Act inventory. The system can readily report the number and identity of chemicals common to both files. This sort of file manipulation has important applications in data base management. For example, when EPA began to measure mass spectra for inclusion in MSSS (*15*), it was possible to focus on commercially important, high-volume chemicals whose mass spectra were not already on file.

Much information can be obtained through this chemical locator function of SANSS. The system acts as a directory to over 70 data bases, ten of which are part of CIS. A much more powerful capability of SANSS, however, is its ability to "package" and "ship" an entire group of registry numbers to other data retrieval systems, where the numbers can be used as search terms. For example, suppose that a substructure search in SANSS results in 400 retrieved compounds. The 400 registry numbers can be stored in a memory terminal in a format acceptable as input to Lockheed's DIALOG, System Development Corporation's ORBIT, or the National Library of Medicine's CHEMLINE. Once the numbers are stored, the user detaches from CIS, ties in to one of these other systems, and sends the group of numbers to the new system, where a full search is conducted. In this way, all the literature citations for each of the 400 chemicals can be obtained on demand. This approach combines the structural search capability of CIS with the bibliographic retrieval ability of the other system.

5) *Data analysis*. The analytical programs of CIS include a family of statistical analysis and mathematical modeling algorithms. These permit linear and nonlinear regression and curve fitting as well as differential and integral calculus (*16*). The mathematical modeling laboratory (MLAB) has associated with it a clustering laboratory (CLAB) that permits generalized pattern-recognition experiments on the user's data.

Second-order analysis of complex NMR spectra, often difficult to accomplish manually, can be completed with GINA (graphical interactive NMR analysis) (*17*). This CIS program can adjust a set of chemical shifts and coupling constants in an iterative fashion until the theoretical spectrum based on these parameters is identical to the measured spectrum. In this way, the correct shifts and coupling constants for a molecule can be determined.

The chemical modeling laboratory (CHEMLAB) is a multifaceted program for modeling chemical structures in three dimensions (*6*). Given a two-dimensional structure, it can calculate the energetically most favored three-dimensional structure. The program has many other estimative capabilities, including calculation of solubilities and partition coefficients, and is used in studies of the biological disposition of chemicals for which few physical properties are known.

Data Evaluation

The accuracy and reliability of the data in a numeric data base are of great importance. This subject, which does not arise in the handling of bibliographic data bases, has required some attention during the development of CIS.

The quality of the data in CIS is mixed, ranging from the evaluated data in the x-ray crystallographic and mass spectral files to the relatively unevaluated entries in RTECS or OHM-TADS. Numerous checks for internal consistency can be made on the atomic coordinate data that result from x-ray crystallographic studies. These checks are arithmetic and geometric in nature and are performed in England by the staff of the Cambridge Crystal Data Centre. This, coupled with the fact that this data base contains almost every reported organic structure, results in an authoritative file of relatively well evaluated data.

Mass spectral data are less susceptible to internal checking, but the CIS staff does examine every new spectrum and, using a derivative of an algorithm proposed by Speck *et al.* (*18*), calculates quality index (QI) factor (a number between 0 and 1000) for the spectrum. Next, a check of the data base is made to see whether the compound is already represented there. If it is not, the new spectrum and its QI factor are added to the file. If the compound is already represented in the file, only the spectrum with the higher QI value is retained. This procedure does not guarantee that all the entries in the data base are of a particular quality, only that there is one spectrum per compound and that a QI factor is available for each entry.

Thus, whereas the x-ray data are quality controlled, the mass spectral data are merely "evaluated." In a program at EPA (*15*), mass spectra are being measured for commercially important chemicals that are not represented in the data base. Quality control is possible in this case (since spectra are measured) and is being exercised. As a result, all these new spectra are receiving high QI factors as they are added to the working file.

Efforts are in progress to extend the approaches in data evaluation and quality control to the other numeric data bases of CIS. Such a process is relatively straightforward when applied to carbon-13 NMR spectroscopy or infrared spectroscopy, but becomes more difficult in a field such as toxicology. The RTECS data base is compiled by NIOSH on a continuing basis. The only control over the quality of the data stems from the fact that NIOSH collects data that have been published in refereed journals. Since the refereeing process is far from certain to detect errors in numeric data, this is not a stringent control; consequently, the toxicology data in RTECS should not be relied on in any quantitative sense. Such data appear almost anecdotal compared to the toxicology data demanded by FDA in connection with any investigational new drug.

The problem of evaluating and controlling the quality of toxicology data is a

difficult one. Internal consistency checks generally are not possible and remeasurement of the data is not feasible. Consequently, a "caveat emptor" approach is taken by NIOSH and, in turn, by CIS. As users enter RTECS, they are greeted with a short statement concerning the quality of the data. Such statements are provided for every numeric data component of CIS, and, while they do not solve the problem, they do apprise the user of possible deficiencies in the data.

Access to CIS

The entire CIS is resident on computers in the private sector (*19*). It operates on a cost-recovery basis through an annual subscription fee of $300, from which educational institutions are exempt, and an hourly charge of $55 or $85 (*20, 21*). The $55 hourly rate is charged for data relating to mass spectra, hazardous materials, acute toxicity, carbon-13 NMR spectra, x-ray structure, NMR literature, and x-ray single-crystal analysis. The $85 rate is for data relating to chemical structure and nomenclature, x-ray powder diffraction, the *Federal Register*, conformational analysis, math modeling, water pollutants, and toxicity of consumer products. CIS uses the GTE-Telenet network and thus is available throughout the United States, Canada, Mexico, Brazil, Argentina, Venezuela, Bermuda, most of Western Europe, Japan, Israel, Australia, New Zealand, Hong Kong, and Singapore through a local telephone call. In the United States and some other countries both 300- and 1200-baud services are available over any standard terminal.

With the CAS registry number, one can transfer between CIS data bases to obtain information pertaining to a given compound. This linking capability can be used to retrieve data corresponding to a specific structure or to identify structures that are consistent with particular data. For example, the SANSS component could be used to retrieve and display the structure, molecular formula, registry number, and name and synonyms for pyridine. With the registry number the user could then retrieve over 50 citations to pyridine published in the *Federal Register*. Retrieval of such specific information is possible with a single command once the registry number of a compound has been ascertained. As another example, SANSS could be used to retrieve the structure of 2-ethylthiopyridine, its registry number, its molecular formula, some additional names, and the fact that the compound is represented in MSSS and several other files. The registry number would then be used by the system as a link to go to MSSS and retrieve the mass spectrum of the compound in conventional bar plot form.

Such linking can be between different CIS components or to data systems outside CIS. For example, SANSS can transmit a file of registry numbers (*22*) from CIS through the GTE-Telenet network to Lockheed's DIALOG system, where CAS bibliographic citations for each registry number can be retrieved.

Because of the modular nature of CIS, it is possible to build individual components elsewhere for subsequent inclusion in the system, so long as the registry number is available for transferring between different components. Also, since the registry number is used within the SANSS component, CIS is an excellent starting point in searches for a particular chemical or class of chemicals.

Use of CIS in Environmental Emergencies

The CIS is a key information resource for EPA's emergency response teams. These teams carry portable computer terminals to chemical spill sites. Using various components of the system to aid in identifying the materials involved, they obtain information on how to approach the site and how to deal with the spill, any resulting fires in the area, and the surrounding population of people, flora, and fauna. OHM-TADS is queried when direct examination of the site provides information on the properties of the spilled substance. Samples are analyzed, typically by gas chromatography–mass spectrometry (GC-MS), and identification is completed with the help of MSSS. Additional information concerning the identity of the chemical is then obtained from SANSS, acute toxicity data are retrieved from RTECS, and regulatory information is obtained from the *Federal Register* search system. These data are used by the agency in conjunction with those provided by OHM-TADS.

In October 1979 some abandoned storage tanks in the town of Sharptown, Maryland, were found to contain a total of 40,000 gallons of unknown chemicals. The tanks were 27 years old and were deteriorating. There appeared to be considerable potential risk to the local population and the environment because the chemicals were thought to be flammable and heavily contaminated with polychlorinated biphenyls. Using GC-MS in conjunction with MSSS, EPA teams were able to identify 76 chemicals, including a number of polychlorinated biphenyls. CIS promptly provided information (*23*) that was used in the removal of the chemicals to a secure site in Curtis Bay, Maryland, and in the closing of the Sharptown tank farm. The Sharptown case was one in which CIS was used to provide many different sorts of information, and the utility of CIS is examined here for one specific chemical, isophorone, found at the Sharptown dump.

From the mass spectrum of the material, obtained by GC-MS, only three peaks (at mass-to-charge ratios 82, 39, and 138) were necessary for the unambiguous identification of the compound as isophorone (3,5,5-trimethyl-2-cylohexen-1-one), whose registry number was stored in a temporary results file. Comparison of the experimental spectrum with the library spectrum was used for confirmation.

The registry number for isophorone was then used in one CIS component after another to obtain and list information concerning the compound. The SANSS record for the compound revealed its chemical identity (structure, molecular formula, and names), and, in its chemical locator capacity, SANSS identified six CIS data bases that contain information on the substance and 13 non-CIS files that also carry relevant data.

The acute toxicity record for isophorone, retrieved from RTECS, suggested that the compound, although not highly toxic, is a serious irritant. The same record revealed that there is an Occupational Safety and Health Administration standard for the compound and that it was being tested by the National Cancer Institute for carcinogenicity. The OHM-TADS record for the compound showed that it is flammable. The water distribution register of organic pollutants (WDROP) component of CIS revealed that the substance has been found as a water pollutant in locations as disparate as Jefferson Parish, Louisiana, and Lake Tahoe, California. Finally, the *Federal Register* search system revealed that isophorone, in addition to being regulated by EPA, is identified by the National Institute of Environmental Health Sciences in its National Toxicology Plan. Identification of isophorone from its mass spectrum and subsequent retrieval and listing of all the information described here required 13 minutes of CIS connect time, for a total cost of approximately $15.

In August 1979 an abandoned mine in

Pittston, Pennsylvania, was found to contain illegally dumped chemicals. During the subsequent year, samples were taken by EPA, analyzed by GC-MS, and identified by MSSS. Substances identified included dichlorobenzenes and some cyanamides. More detailed information on the chemicals was obtained from OHM-TADS, SANSS, and TOXLINE. The EPA based its response on the data it retrieved from these systems.

The May 1980 rioting in Dade County, Florida, led to many fires, several of which involved chemical storage facilities. The regional EPA spill response team used CIS to identify the chemicals involved and to delineate the appropriate procedures for combating the chemical fires (*24*). In addition, RTECS was sampled to check on the toxicity of the chemicals.

Use of CIS in Pollution Monitoring

Most of EPA's environmental monitoring laboratories use GC-MS as their primary method for identifying chemical pollutants, so heavy use is made of MSSS (*25*). A routine search of this data base is made with the measured spectra of unknown compounds, and, if an identification results, the remainder of CIS is searched for further information. In many cases, however, the unknown substance is not represented in CIS. Recently, CIS programs have been developed (*26*) which can analyze mass spectra and determine whether or not specific substructural features are present in a compound. This promising approach may be expanded in an effort to deal with compounds whose spectra are not in CIS. (At the same time, a vigorous effort is being made to enlarge CIS.)

In a variation on this approach, when a positive identification cannot be made, some insight into the structure of the unknown substance can often be gained from close matches or from other information available. The chemist, using an intuitive approach, can access SANSS and look for substances that are structurally similar to the unknown. If a relevant chemical structure comes to light, the laboratory tries to obtain a sample and measures its mass spectrum. If the spectra match, the identification is complete and the compound and its spectrum can be entered into MSSS. Procedures of this sort are used at EPA monitoring facilities across the country and have been used in routine monitoring as well as special cases, such as Love Canal.

The WDROP data base is a unique source of data on the pollution of water systems by organic chemicals. This CIS component is routinely queried by EPA's Office of Pesticides and Toxic Substances in efforts to determine the exposure factors required for risk assessment. The EPA laboratory in Athens, Georgia, uses WDROP to respond to many requests from EPA's Office of Drinking Water for information concerning specific organic pollutants in the nation's drinking water supplies. These requests, which can now be honored in minutes, often require review of the data reported between particular dates and for specific concentrations—all search terms accommodated by WDROP. In a recent study of the feasibility of reusing potable water supplies, successful searches were made in WDROP for data on the incidence of pollutants in treated waste water, river water, surface water, and ground water.

The EPA has been studying the toxic effects of arsenic-containing chemicals. In response to a special request from the Office of Pesticides and Toxic Substances (*27*), SANSS, RTECS, and MLAB were used to prepare a histogram showing the number of organic and of inorganic arsenic-containing compounds with median lethal doses in the ranges 0 to 1, 1 to 10, 10 to 50, 50 to 100, and more than 100 milligrams per kilogram. The data were retrieved in under half an hour at a total cost of $16, and a report was prepared and delivered on the day of the request.

Another project from the Office of Pesticides and Toxic Substances involves the analysis of studies of toxic chemicals in other countries. CIS contains a number of foreign lists of chemicals and the office routinely uses SANSS to retrieve them, making full use of the Boolean logic combinatorial functions.

Use of CIS for Public Awareness and Information

Several of the CIS components play an important role in the area of public awareness. Foremost among these is the *Federal Register* search system, which is used by the many commercial organizations that must keep abreast of federal regulatory activities. The CIS component that is based on *The Clinical Toxicology of Commercial Products* (*13*) is an invaluable source of information on the composition and toxicity of products such as cleaners, polishes, and over-the-counter drugs. Since such materials are frequently involved in domestic poisoning, the data base is of value to hospitals and poison control centers.

Use of CIS in Research

A fundamental purpose of CIS is to provide numerical data to research chemists (*28*). It is for this reason that much of the system's financial support is provided by NIH. A single example of the use of CIS as a research tool follows.

A knowledge of the composition of the ash left from combustion of coal is valuable in efforts to increase the efficiency of the combustion process. The chemicals found in coal ash include metal oxides and more complex materials, such as silicates. Identification of such substances can be at least partially effected by fluorescence analysis, but a more generally applicable method is x-ray powder diffraction. With the relevant CIS component, rapid identification of such materials from powder diffraction patterns is routine. In this way, monitoring of both efficiency and emission characteristics of industrial furnaces is carried out by heavy industries.

Expansion of CIS

Development of CIS falls into two categories: augmentation of data bases currently searched by CIS and addition of new ones. The content of some of the data bases, such as the x-ray structure file, is not the responsibility of CIS, which simply leases the file. In contrast, MSSS is the joint responsibility of NIH, EPA, and NBS and is an example of the way in which an interagency effort can provide updating of a CIS data base.

In 1978, in response to a request from EPA, the NBS reviewed the over 30,000 spectra in MSSS and, using SANSS, found that of all the pesticide chemicals of concern to the agency, only 23 percent were represented in the file. Accordingly, samples of many of the missing pesticides were obtained by NBS and their mass spectra were measured by an NBS contractor and added to the file. Over 1200 spectra were added, and the proportion of hazardous pesticide chemicals represented in the file rose to 64 percent (*29*).

The EPA's Environmental Research Laboratory in Cincinnati has a major responsibility for identifying pollutant chemicals. In 1979 this laboratory, using SANSS, demonstrated that too few commercial chemicals were represented in MSSS. Accordingly, the laboratory awarded a 3-year contract (*15*) to measure the appropriate spectra to the specifications used by NBS, EPA, and NIH. The laboratory used SANSS to prepare a list of all chemicals present in the Toxic

Substances Control Act inventory but absent from MSSS. This list of 39,217 chemicals (91 percent of the inventory) was then arranged in descending order of domestic production volume. Thus prioritized, the list was provided to the contractor, which acquires each of the chemicals, checks it for purity, purifies it if necessary, and then measures its mass spectrum. Since so little material is needed for mass spectral measurements, it is possible to assemble a collection of purified reference standards for those chemicals simultaneously. This is being done, and plans are being made to measure the gas phase infrared spectra of the compounds for inclusion in the infrared data base of CIS.

Use of CIS in Chemical Modeling

The CIS is proving useful for modeling environmentally significant properties of certain chemicals. For example, a SANSS-derived list of pesticides was assembled with their soil retention times, the latter thought to be a valid measure of bioaccumulation. These chemicals were then examined structurally with SANSS. They were also examined by conformational analysis using CHEMLAB, and several structurally relevant, linear, free energy–derived parameters were calculated. These included such quantities as partial atomic charge and polarizability (*30*). Experimental partition coefficients and solubilities were also provided to the program, and all these data were then evaluated by multiregression analysis with MLAB. The resulting models showed good promise for the use of these types of parameters for modeling chemical bioaccumulation, although the partition coefficient and solubility remain the best predictive measures.

Conclusion

Since 1971 the CIS approach to handling chemical information has continued to prove effective. CIS has been particularly useful to that sector of the chemical community beset by questions of an environmental nature. As more data bases are added to CIS and as additional links are established between CIS and commercial bibliographic systems, the utility of CIS can only be enhanced.

References and Notes

1. S. R. Heller, G. W. A. Milne, R. J. Feldmann, *Science* **195**, 253 (1977).
2. S. R. Heller and G. W. A. Milne, *Am. Chem. Soc. Symp. Ser.* **84**, 144 (1978).
3. ———, *ibid.* **94**, 255 (1978).
4. ———, *Environ. Sci. Technol.* **13**, 798 (1979).
5. R. J. Feldmann, G. W. A. Milne, S. R. Heller, A. Fein, J. A. Miller, B. Koch, *J. Chem. Inf. Comput. Sci.* **17**, 157 (1977); G. W. A. Milne, S. R. Heller, A. E. Fein, E. Frees, R. G. Marquart, J. A. McGill, J. A. Miller, D. S. Spiers, *ibid.* **18**, 181 (1978).
6. R. Potenzone, Jr., E. Cavicchi, H. J. R. Weintraub, A. J. Hopfinger, *Comput. Chem.* **1**, 187 (1977).
7. S. R. Heller, *Anal. Chem.* **44**, 1951 (1972); ———, H. M. Fales, G. W. A. Milne, *Org. Mass Spectrom.* **7**, 107 (1973); ———, *J. Chem. Educ.* **49**, 725 (1973); S. R. Heller, D. A. Koniver, H. M. Fales, G. W. A. Milne, *Anal. Chem.* **46**, 947 (1974); S. R. Heller, R. J. Feldmann, H. M. Fales, G. W. A. Milne, *J. Chem. Doc.* **13**, 130 (1973); S. R. Heller, H. M. Fales, G. W. A. Milne, R. J. Feldmann, N. R. Daly, D. C. Maxwell, A. McCormick, *Adv. Mass Spectrom.* **6**, 1037 (1975); R. S. Heller, G. W. A. Milne, R. J. Feldmann, S. R. Heller, *J. Chem. Inf. Comput. Sci.* **16**, 176 (1976); S. R. Heller, R. S. Heller, A. McCormick, D. C. Maxwell, G. W. A. Milne, *Adv. Mass Spectrom.* **7B**, 985 (1977); G. W. A. Milne, S. R. Heller, R. S. Heller, D. P. Martinsen, *ibid.* **8**, 1578 (1979).
8. D. L. Dalrymple, G. W. A. Milne, C. L. Wilkins, S. R. Heller, *Org. Magn. Reson.* **11**, 535 (1978); G. W. A. Milne, J. Zupan, S. R. Heller, J. A. Miller, *ibid.* **12**, 289 (1979); J. Zupan, S. R. Heller, G. W. A. Milne, J. A. Miller, *Anal. Chim. Acta* **103**, 141 (1978).
9. R. J. Feldmann, G. W. A. Milne, S. R. Heller, *Trans. Am. Crystallogr. Assoc.* **12**, 75 (1976).
10. R. G. Marquart, I. Katsnelson, G. W. A. Milne, S. R. Heller, G. G. Johnson, Jr., R. Jenkins, *J. Appl. Crystallogr.* **12**, 629 (1979).
11. R. G. Marquart, L. Marquart, S. Mintz, J. R. McGill, J. R. McDaniel, S. R. Heller, G. W. A. Milne, *Online* (1980), p. 45.
12. J. R. McGill, S. R. Heller, G. W. A. Milne, *J. Environ. Pathol. Toxicol.* **2**, 539 (1978).
13. R. E. Gosselin, H. C. Hodge, R. P. Smith, M. N. Gleason, *The Clinical Toxicology of Commercial Products* (Williams & Wilkins, Baltimore, 1975).
14. B. L. Carson, J. E. Going, V. Lopez-Avila, J. L. McCann, C. J. Cole, L. J. Holt, paper presented at the National Meeting of the American Chemical Society, Las Vegas, Nev., 26 August 1980.
15. Contract 68-03-2879, EPA Environmental Monitoring Support Laboratory, Cincinnati, Ohio 45219.
16. G. D. Knott, *Comput. Programs Biomed.* **10**, 271 (1979).
17. S. R. Heller and A. E. Jacobson, *Anal. Chem.* **44**, 2219 (1972); R. B. Johanneson, J. A. Ferretti, R. K. Harris, *J. Magn. Reson.* **3**, 84 (1970).
18. D. D. Speck, R. Venkataraghavan, F. W. McLafferty, *Org. Mass Spectrom.* **13**, 209 (1978).
19. This policy was adopted for CIS in 1972 and is consistent with *Circular A-76* of the Office of Management and Budget. As a result, CIS resides in the private sector and is subject to many of the free-market forces there. The system has access to the most modern computer facilities, since the often lengthy process of government procurement of computers is obviated.
20. The charges do not include communications costs, which average about $5 per hour, irrespective of location. Users outside the United States must pay an additional local telecommunications charge to their own government.
21. To receive further information concerning access to CIS, contact Charles Sullivan, CIS Project, 2135 Wisconsin Ave., NW, Washington, D.C. 20007.
22. S. R. Heller and G. W. A. Milne, *J. Chem. Inf. Comput. Sci.* **20**, 204 (1980).
23. J. Wright, T. Massey, P. Retallick, personal communications.
24. J. Wright and G. Moein, personal communications.
25. R. A. Hites and V. Lopezavila, *Anal. Chem.* **51**, 1452 (1979); R. A. Hites, G. A. Jungclaus, V. Lopezavila, L. S. Sheldon, *Am. Chem. Soc. Symp. Ser.* **94**, 63 (1979); L. H. Keith, *J. Chromatogr. Sci.* **17**, 48 (1979); H. Knoppel, *Ann. Chim. (Rome)* **67**, 649 (1977); G. A. Jungclaus, V. Lopezavila, R. A. Hites, *Environ. Sci. Technol.* **12**, 88 (1978).
26. W. Meisel, M. Jolley, S. R. Heller, G. W. A. Milne, *Anal. Chim. Acta* **112**, 407 (1979).
27. D. Sterling, personal communication.
28. S. R. Wilson and J. C. Huffman, *J. Org. Chem.* **45**, 560 (1980).
29. Mass spectra are available for only 46 percent of all pesticides. This proportion has decreased since the NBS contract was completed because over 1600 new pesticides have been added by EPA to the lists that were originally considered. About half of the pesticides for which mass spectra are unavailable are inorganic compounds, which probably do not have useful mass spectra.
30. J. Dragun, R. Potenzone, Jr., C. S. Fowler, C. S. Helling, *Proceedings of the 53rd Annual Water Pollution Control Federation Conference, Las Vegas, Nevada, 23 September 1980* (Water Pollution Control Federation, Washington, D.C., 1980).

31. We would like to acknowledge the willing assistance provided by the EPA scientific and administrative staff members who provided background material for the examples cited.