Chemical Structure Information Systems

Interfaces, Communication, and Standards

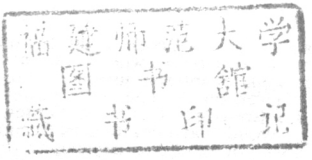
ACS SYMPOSIUM SERIES 400

Chemical Structure Information Systems

Interfaces, Communication, and Standards

Wendy A. Warr, E
ICI Pharmaceutic

Developed from a symposium sponsored
by the Division of Chemical Information
at the 196th Meeting
of the American Chemical Society,
Los Angeles, California,
September 25–30, 1988



American Chemical Society, Washington, DC 1989

Library of Congress Cataloging-in-Publication Data

Chemical structure information systems.

(ACS Symposium Series, ISSN 0097–6156; 400).

"Developed from a symposium sponsored by the Division of Chemical Information at the 196th Meeting of the American Chemical Society, Los Angeles, California, September 25–30, 1988."

Bibliography: p.

Includes index.

1. Chemical structure—Data processing—Congresses. 2. Information storage and retrieval systems—Chemical structure —Congresses.

I. Warr, Wendy A., 1945– . II. American Chemical Society. Division of Chemical Information. III. American Chemical Society. Meeting (196th: 1988: Los Angeles, Calif.). IV. Series

QD471.C517 1989 541.2'2'0285 89–14988
ISBN 0–8412–1664–9

ACS Symposium Series

M. Joan Comstock, *Series Editor*

1989 ACS Books Advisory Board

Foreword

The ACS SYMPOSIUM SERIES was founded in 1974 to provide a medium for publishing symposia quickly in book form. The format of the Series parallels that of the continuing ADVANCES IN CHEMISTRY SERIES except that, in order to save time, the papers are not typeset but are reproduced as they are submitted by the authors in camera-ready form. Papers are reviewed under the supervision of the Editors with the assistance of the Series Advisory Board and are selected to maintain the integrity of the symposia; however, verbatim reproductions of previously published papers are not accepted. Both reviews and reports of research are acceptable, because symposia may embrace both types of presentation.

Preface

In the last few years many computer systems have appeared which allow graphics entry of chemical structures. At first, structure entry required an expensive graphics terminal. As cheaper graphics terminals came on the market and there were more sales of suitable software, more and more users became familiar with drawing chemical structures on a screen. Once microcomputers became widely available there was a great increase in the number of software packages on offer to research workers and information professionals. The healthy competition fostered by the advent of microcomputer packages has, on the whole, been beneficial to the end-user. However, the multiplicity of available systems does lead to one major problem: the end-user is not prepared to learn and use several different methods of drawing chemical structures. He, or she, wants to be able to use the same drawing method to prepare scientific reports and to access public, corporate and personal data bases. Not all end-users would agree on the preferred drawing method, or "front-end", and it is undesirable for any vendor to have a monopoly, but there is obviously much scope for building interfaces between the most popular drawing methods and the major sources of chemical structure information.

In the ideal world, neither commercial pressures nor technical complications should prevent a user from using his or her preferred interface to access multiple public, corporate and personal chemical structure data collections.

The "universal interface" is thus an exciting objective, but, related to it, and equally important, is the need for standards which allow the transfer of data between many different chemical structure handling systems.

This book directly addresses the problems of interfaces, communication, and standards. It is authored by vendors and users of a variety of chemical information systems, for example, for scientific document production, substructure searching, reaction indexing and molecular modeling.

My own, introductory, chapter deals with the development of standards for chemical structure representation. I have given copious literature references which should be helpful to readers of those chapters which have fewer references.

The reader who is particularly concerned with information technology, and software and hardware standards, would be well advised to read Chapter 3.

The book does not attempt to address the issue of standards for electronic publishing. However, the reader may be interested to know the electronic methodology used in producing the book. All but two of the chapters were submitted to me on floppy disks. The remaining two were downloaded from the Chemical Abstracts Service electronic mail system. Only one floppy disk could not be converted to ASCII and reprocessed. That chapter had to be word processed from scratch by my secretary. All the other chapters were sent electronically to her using the DEC All-in-1 office automation system in use at ICI. The earliest chapters unfortunately required careful checking because the terminal emulation software used did not have an error checking protocol and occasional clauses were

duplicated or omitted. After this experience, we used the Kermit protocol to transmit later chapters.

My secretary edited all the chapters (according to my instructions on a hard copy) on a Wordplex machine using the WPS+ word processor. A Wordplex floppy disk of each chapter was sent for typesetting (with a hard copy indicating "problems" such as special fonts, umlauts etc.). The floppy disks were re-converted to ASCII and typeset on a Linotron 202.

Unfortunately we were not able to process "compound documents". Only the text of the book was reprocessed and typeset electronically. All diagrams and figures had to be resized from hard copy and inserted.

Finally, I sent a photocopy of the actual printed version to each author. However, I recognize that I myself must bear responsibility for any errors that may have crept in during the processing of the camera-ready "manuscript" sent to ACS.

Acknowledgments

I would like to thank all the people in the ACS Division of Chemical Information who helped to make the original symposium possible and, in particular, Joanne Witiak, the chairperson of the program committee. I am also grateful to all the authors in this book for their efforts in writing, and rewriting, the chapters within the imposed deadlines. Three people in Information Services Section at ICI Pharmaceuticals deserve special mention: Frank Loftus who translated and transmitted electronic files for me and also redrafted some diagrams; Madeline Gray who traced and checked numerous literature references; and my secretary, Mary Burgess, who has put more effort into the book than any of us. I would also like to thank Hope Services (Abingdon) Ltd., for their efficient typesetting and helpful advice.

WENDY A. WARR
April 1989

Other ACS Books

Chemical Structure Software for Personal Computers
Edited by Daniel E. Meyer, Wendy A. Warr, and Richard A. Love
ACS Professional Reference Book; 107 pp;
clothbound, ISBN 0–8412–1538–3; paperback, ISBN 0–8412–1539–1

Personal Computers for Scientists: A Byte at a Time
By Glenn I. Ouchi
276 pp; clothbound, ISBN 0–8412–1000–4; paperback, ISBN 0–8412–1001–2

Biotechnology and Materials Science: Chemistry for the Future
Edited by Mary L. Good
160 pp; clothbound, ISBN 0–8412–1472–7; paperback, ISBN 0–8412–1473–5

Polymeric Materials: Chemistry for the Future
By Joseph Alper and Gordon L. Nelson
110 pp; clothbound, ISBN 0–8412–1622–3; paperback, ISBN 0–8412–1613–4

The Language of Biotechnology: A Dictionary of Terms
By John M. Walker and Michael Cox
ACS Professional Reference Book; 256 pp;
clothbound, ISBN 0–8412–1489–1; paperback, ISBN 0–8412–1490–5

Cancer: The Outlaw Cell, Second Edition
Edited by Richard E. LaFond
274 pp; clothbound, ISBN 0–8412–1419–0; paperback, ISBN 0–8412–1420–4

Practical Statistics for the Physical Sciences
By Larry L. Havlicek
ACS Professional Reference Book; 198 pp; clothbound; ISBN 0–8412–1453–0

The Basics of Technical Communicating
By B. Edward Cain
ACS Professional Reference Book; 198 pp;
clothbound, ISBN 0–8412–1451–4; paperback, ISBN 0–8412–1452–2

The ACS Style Guide: A Manual for Authors and Editors
Edited by Janet S. Dodd
264 pp; clothbound, ISBN 0–8412–0917–0; paperback, ISBN 0–8412–0943–X

Chemistry and Crime: From Sherlock Holmes to Today's Courtroom
Edited by Samuel M. Gerber
135 pp; clothbound, ISBN 0–8412–0784–4; paperback, ISBN 0–8412–0785–2

For further information and a free catalog of ACS books, contact:
American Chemical Society
Distribution Office, Department 225
1155 16th Street, NW, Washington, DC 20036
Telephone 800–227–5558

Contents

Chapter 1

Introduction

Wendy A. Warr
ICI Pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG, England

Over the past twenty years there has been a trend away from user-hostile chemical information systems towards systems with user-friendly, graphics interfaces, which allow the scientist to use his preferred method of communication, that is, the chemical structure diagram. The abundance of chemical structure interfaces nowadays actually causes a new problem: there are too many systems for the user to learn and there are too few standards.

Chemical structure handling by computer in the 1960s required very specialized expertise.

During the 1970s there was a boom in Information Technology and great strides were made in chemical information systems, but such systems were usually beyond the reach of the end-user.

The advent of computer graphics (*1*,*2*), the proliferation of personal computers (*3*) and the development of relatively user-friendly software has brought chemical data base building, structure and substructure searching, and chemical report production within the reach of the average chemist.

The problem is no longer the lack of systems and data bases but rather the proliferation of systems which cannot be linked in a seamless manner. Drawing the same input structure into more than one software interface is a waste of time. The user will also no longer tolerate "sneakernet" (the information is put onto a tape or floppy disk and someone in sneakers runs down the corridor with it to the next machine). The systems expert has no time to write multiple conversion routines. There is quite obviously a need for seamless interfaces and good communication. Behind these lies the need for standards.

Early Attempts at Standardization and Interfacing

As far as the chemist is concerned the standard, and preferred way, of representing and communicating chemical structure information is the two-dimensional chemical structure diagram. In the early days, the hardware and software available could not handle such diagrams. If we overlook fragment codes and punch card technology (since such methods did not represent the full topology of a molecule), the earliest methods for handling chemical structures involved chemical nomenclature or line notations.

Many line notations have been suggested but only three gained significance: the

0097–6156/89/0400–0001$06.00/0

SMILES notation (Simplified Molecular Input Line Entry System) used in the Pomona College MedChem project (*4*);the Dyson notation (*5*) adopted as a standard by the International Union for Pure and Applied Chemistry (IUPAC); and Wiswesser Line Notation (WLN), which became the actual standard used in the chemical industry (*6*). The rules of WLN were "standardized" by the Chemical Notation Association (CNA), but the two major WLN data bases, Index Chemicus Registry System, ICRS, from the Institute for Scientific Information and the Commercially Available Organic Chemicals Index, CAOCI (*7*) used somewhat different conventions.

The ICI CROSSBOW (Computerised Retrieval of Organic Structures Based on Wiswesser) system (*8*) had a large number of users worldwide. It used not only WLN but also fragment codes and connection tables derived from WLN. The CROSSBOW connection table was bond-implicit (*9*) as opposed to the bond-explicit connection table used by Chemical Abstracts Service (CAS). The CAS Registry System (*10*) was (and still is) based on nomenclature, registry numbers, and connection tables rather than line notations. A number of European companies bought Chemical Abstracts data for use in-house and were able to access both internal (corporate) and literature information with proprietary systems (*11*). ICI converted the ICRS data base into a CROSSBOW-compatible form and was thus able to use the same technology for searching a literature data base as well as ICI in-house chemical structures (*12*). However, the Chemical Abstracts data base was not accessible this way, despite experiments in interconversion of CAS and CROSSBOW connection tables (*13*).

The Emergence of Graphics Systems for Chemical Structure Searching

The earliest system which allowed substructure searching involving chemical connectivity input and structure display, was the National Institutes of Health/Environmental Protection Agency Chemical Information System, the NIH/EPA CIS (*14*). Its Structure and Nomenclature Search System (SANSS) eventually allowed access to a large range of public data bases. The system allowed access from teletype terminals and was not graphics-based.

The Chemical Abstracts Registry File became substructure searchable online in the early 1980s, first by means of the DARC (Description, Acquisition, Retrieval and Correlation) system (*15*) as implemented by Télésystèmes and, soon afterwards, in the CAS ONLINE Service (*16*).

The evolution of molecular graphics (*1*) is described in an earlier ACS Symposium Series book (*17*) which acts as an interesting precursor to this present volume. Chemical reaction systems such as LHASA (Logic and Heuristics Applied to Synthetic Analysis) (*18*) and SECS (Simulation and Evaluation of Chemical Synthesis) (*19*) had long used graphics but it was some time before the first in-house, proprietary system appeared, attracting much interest in the chemical and pharmaceutical industries. This was Upjohn's Compound Information System, COUSIN (*20–21*).

Neither COUSIN nor the CAS ONLINE Messenger software were portable or commercially available. From the early 1980s there was a big demand for user-friendly, interactive access to in-house, chemical structure data bases. The market leaders became MACCS (*22–23*) marketed by Molecular Design Ltd (MDL) and

DARC in-house (*15,23*) marketed by Télésystèmes. OSAC (Organic Structures Accessed by Computer) from the ex-Leeds University team ORAC Ltd, appeared rather later (*24*).

Many organizations acquiring these new systems already had files of structures on chemical typewriters or encoded in WLN and wanted to generate MACCS data bases or DARC connection tables automatically. Interconversion thus became the rage of the mid-1980s. The Chemical Structure Association's first publication was the proceedings of the CNA(UK) seminar on interconversion held at the University of Loughborough in March 1982. Elder's DARING program to convert WLNs to connection tables is described therein. This program has been widely used but further software is required to convert the DARING connection tables to MACCS, or other, versions, and to generate structure coordinates needed for the actual graphics display of the structures (*25*).

Some companies have written proprietary algorithms to allow structures drawn in MACCS to be entered to the Pomona College MedChem system for structure activity relationships.

Integration of Chemical Structure Data with Property Data

Once many in-house, chemical structure data bases had been built, the users began to realize that it was more efficient to use commercially-available structure handling software for structures alone (or structures and a minimal amount of related property data) and to take advantage of data base management systems to handle property data. Molecular Design Limited (MDL) and Télésystèmes modified their MACCS and DARC software, respectively, to allow for the appropriate interfacing of structures and data.

Chapters 5 and 10 in this book discuss such interfaces (and related topics).

The era of integration of structures with text (*17*) was hardly beginning.

Proliferation of Incompatible Systems in the 1980s

By the mid-1980s the advent of the microcomputer had started to make a big impact on the world of chemical information (*3*). Chemical structure drawing packages such as ChemDraw (*26–27*); the Wisconsin Interactive Molecular Processor, WIMP (*26*); and Molecular Presentation Graphics, MPG (*26*) could be used for document production. Multipurpose, connection table based programs such as PSIDOM (Hampden Data Services' Professional Structure Image Database on Microcomputers) (*26,28*) and CPSS (Molecular Design Limited's Chemists' Personal Software Series) (*26,28–29*) were available for document production, data base building and substructure searching. The structure drawing technology in some of these microcomputer-based packages was much more sophisticated and user-friendly than in previous graphics interfaces such as that to CAS ONLINE. The advantages of using a microcomputer package as a "front-end" to the online systems CAS ONLINE, DARC and CIS were discussed (3).

The first such front-end to appear was Fein-Marquart's SuperStructure (*30*), which allowed graphics chemical structure input for accessing the CIS data bases.

Meanwhile, a multiplicity of chemical and pharmaceutical companies involved in the Molecular Design Limited Software Users Group as users of MACCS and

CPSS started to demand facilities for downloading structures from CAS ONLINE into MACCS in-house data bases. In an unpublished survey carried out in 1987, the author of this chapter established that members of the Molecular Design Limited Software Users Group were even more interested in uploading i.e., drawing a structure query using MACCS methodology and sending it up the line to search the CAS ONLINE data base. Interestingly the users had a preference for MACCS over CPSS for this process. Their views on the user-friendliness of structure input packages were heavily colored by their *familiarity* with MACCS and CAS ONLINE. Indeed, there was an abysmal ignorance of the wealth of structure drawing facilities available both in the United States and in Europe. It is hoped that a recent ACS Professional Reference Book (*26*) will rectify this situation. In that book about 70 different software packages for personal computers are described.

In this author's company, by the end of 1987, information scientists were faced with mastering a large range of methods for drawing chemical structures both on graphics terminals and on personal computers. Examples were:

1. MACCS (for in-house chemical data bases).
2. CAS ONLINE and DARC (for searching the chemical literature).
3. ORAC (Organic Reactions Accessed by Computer, the reaction indexing software written at Leeds University and marketed by ORAC Ltd.) (*31*).
4. PsiGen (the structure drawing module of Hampden Data Services' PSIDOM software) (*26*).
5. PsiORAC (the PsiGen interface to ORAC).
6. ChemDraw (the Macintosh software package from Cambridge Scientific Computing which is the ICI chemist's preferred method for drawing high-quality structures for pasting into documents) (*26–27*).
7. SANDRA (the Structure and Reference Analyzer marketed by Springer Verlag for help with use of the Beilstein Handbook) (*26,32*), and
8. TOPFRAG (the Topological Fragment Code Generator program for chemical structure access to the Derwent patents data base) (*26*).

Other companies might cite different examples but their listings would be equally long.

It is obvious that the end-user will not be prepared to learn this multiplicity of ways of inputting a chemical structure.

In the ideal world, neither commercial pressures, nor technical complications, should prevent a user from drawing a chemical structure the way he or she wants to and accessing any personal, corporate or public file with that structure (or substructure). In practice, both commercial and technical factors cause severe limitations. This is the problem at which this book is aimed.

Front-End Software

Since the book was conceived new front-ends such as STN Express (*26,33*), MOLKICK (*26,34*) and DARC CHEMLINK (*26*) have appeared on the market. All are IBM PC based graphics packages which allow offline query formulation followed by connection to a host computer and uploading of the query for online searching. Offline query formulation means that an end-user can avoid the so-

called "taxi-meter syndrome" that is a consequence of incurring online charges if formulating a search strategy while connected to the host computer.

DARC CHEMLINK (from Télésystèmes) is a program which allows offline formulation of queries for submission to DARC data bases. Its structure drawing interface is intentionally very similar to that of the DARC system online.

MOLKICK (sold by Springer Verlag), which is described in Chapter 8 of this book, is a memory-resident query editor which converts a structure to a string suitable for uploading to Beilstein/Softron, CAS or DARC data bases. Its chemical structure input is very similar to Beilstein's MOLMOUSE software (*35*).

STN Express (the Chemical Abstracts Service approved front-end to CAS ONLINE) is touched upon in Chapters 6 and 7. It, also, allows the user to formulate queries offline and provides help with special difficulties such as tautomerism. In addition it has a Guided Search module for novices in the STN Command Language and Boolean logic. Structure drawing is compatible with PsiGen.

MOLKICK allows the searcher to use his own preferred communication software. STN Express incorporates its own communications package, but in so removing the user's freedom of choice, it does impose an error-checking protocol which has considerable advantages.

With in-house data bases, the financial advantages of offline query formulation are less obvious, but there is still a business need for uploading structures from a PC to a corporate data base or downloading from the mainframe to the PC. MDL's CPSS package supplies such a front-end to data bases under MDL's MACCS software.

Télésystèmes have devised a way of capturing DARC structure vectors (not connection tables) for display with the popular ChemDraw package on a Macintosh, but (as yet) ChemDraw substructure searchable data bases are not a possibility.

Unfortunately downloading is not yet possible with STN Express. There is a package called CASKit (*26*), unsupported by both MDL and CAS, which captures the vectors for CAS ONLINE output structures, converts them into a graphics metafile and then converts the metafiles into MACCS-compatible connection tables plus structure coordinates.

The writer, or user, of a PC package which is an unsupported interface to software from a major vendor, faces obvious dangers. The vendor of the host system may, maliciously or unknowingly, change minor features of his file structure or command language, rendering the PC interface inoperable.

Commercial Considerations

The expression "major vendor" was used advisedly in the above. In most of the collaborative ventures seen so far a major vendor uses the products of a small software house or develops software himself. From the viewpoint of the users, it is unfortunate that major vendors cannot cooperate with each other. The prospects for a supported MACCS interface to the CA Registry File are still not good.

Of the PC-based packages, PSIDOM is one that has been aimed especially at collaborative software developments. There are PsiGen interfaces to the CA Registry File (PsiCAS in STN Express) to ORAC (PsiORAC) to Derwent data

bases (in the programs TOPFRAG and TORC which convert structures to Derwent fragment codes) and to DARC in-house data bases. We will leave the reader to speculate whether it is really technical difficulties that prevent the appearance of "PsiMACCS".

It is commercial considerations that have led to the abundance of IBM PC based software for chemical structure handling and the limited number of packages for use on the Apple Macintosh. However, the user base for the Macintosh is increasing and this situation will change.

Standardization in the Macintosh environment is such that the user of programs such as ChemDraw could easily adapt to other well-designed Macintosh structure drawing packages. The appearance of "ChemDraw-like" front-ends in the near future is a certainty. (There is already a Macintosh-based front-end to the CA Registry File, called ChemConnection, marketed by Softshell of Henrietta, New York, but not supported by CAS.)

Technical Considerations

Computer graphics standards are discussed in an earlier ACS Symposium Series Book (*1*,*2*). In the present book, the reader is particularly referred to Chapter 3, by Smith, for a detailed exposition of trends and standards in hardware, operating systems and environments, and applications software. Smith also examines the implications for chemical information.

Standard File Structures

One could state with wry humour that the good thing about standards is the number of them that there are. This is as true in the chemical structure representation field as in other fields.

Some "standard" ways of storing and transferring chemical structures are proprietary (e.g., MDL's Molfile); others such as the JCAMP-CS format, published by the Joint Committee on Atomic and Molecular Physics, are in the public domain. Barnard (*36*) refers to some of them in a paper that deals with recent developments in improving the Standard Molecular Data (SMD) file format and work towards establishing it as the one standard for transfer of chemical structure information between systems. In Chapter 11 of this book, Donner et al. describe the SMD format in more detail. Garavelli, in Chapter 12, also discusses SMD, but concentrates on existing "standards" for molecular modeling systems.

Summary of Chapters

In Chapter 2 Richard Hong discusses the wide variety of uses of chemical structure information and the need for flexibility in a file format for free exchange of chemical data.

The next chapter by Dennis Smith of MDL describes hardware and software standards in detail. The reader who is particularly concerned with information technology would be well advised to read this. The chapter is a technical and technocommercial one. It is not intended simply to explain MDL's commercial position, any more than Chapter 6, by Chemical Abstracts Service authors, is

supposed to state simply a CAS position. However, these contributions (and others) were invited partly because of the significance of the vendors involved. The chapter by Chemical Abstracts authors is followed by Bill Town's contribution since his company, Hampden Data Services, collaborated in the STN Express front-end to CAS ONLINE.

Sandy Lawson's contribution on chemical structure browsing (Chapter 4) may seem peripheral to the main topic of the symposium, but the algorithm he employs could, he hopes, be used for data bases other than Beilstein in future.

Other developments at the Beilstein Institute, and in particular the so-called ROSDAL string for transferring chemical structure information from PC to host computer, are described in Chapter 8.

Chapter 5 deals with integration and standards both for a reaction indexing systems (ORAC) and for structure management software (OSAC).

Some useful interfaces to the DARC system are described in Chapter 10.

Chapter 9 is concerned with Polygen's CENTRUM system as an integration tool for the various components of scientific documents.

As described earlier, Chapters 11 and 12 concern standard molecular description files.

Conclusion

The present state of the art in interfacing, communication and standards in the field of chemical structure information is rather confused. Many problems and issues are being aired but there are few answers, let alone standards. Software, some of it very useful, is nevertheless appearing and the vendors cannot afford to wait for standards to be laid down. Standards committees are notoriously slow in their deliberations. It remains to be seen whether the SMD movement will establish a standard in a reasonable time or whether a *de facto* standard will become established before then.

Literature Cited

1. Wipke, W.T. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington D.C., 1987; pp 1–7.
2. Sanderson, J.M.; Dayton, D.L. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington D.C., 1987; pp 128–142.
3. Town, W.G. In *Chemical Structures: The International Language of Chemistry*; Warr, W.A., Ed.; Springer Verlag. Heidelberg, 1988; pp 243–249.
4. Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.
5. Dyson, G.M. In *Chemical Information Systems*; Ash, J.E.; Hyde, E., Eds.; Ellis Horwood: Chichester, 1974; pp 130–155.
6. Baker, P.A.; Palmer, G.; Nichols, P.W.L. In *Chemical Information Systems*; Ash, J.E.; Hyde, E., Eds.; Ellis Horwood: Chichester, 1974; pp 97–129.
7. Walker, S.B. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 3–5.
8. Eakin, D.R. In *Chemical Information Systems*; Ash, J.E.; Hyde, E., Eds.; Ellis Horwood: Chichester, 1974; pp 227–242.

9. Ash, J.E. In *Chemical Information Systems*; Ash, J.E.; Hyde, E., Eds.; Ellis Horwood: Chichester, 1974; pp 156–176.
10. Dittmar, P.G., Stobaugh, R.E.; Watson, C.E. *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 111–121.
11. Graf, W.; Kaindl, H.K.; Kniess, H.; Schmidt, B.; Warszawski, R. *J. Chem. Inf. Comput. Sci.* **1977**, *19*, 51–55.
12. Warr, W.A. In *Proc. CNA(UK) Seminar on Integrated Data Bases for Chemical Systems*; Chemical Notation Association (UK), 1979; pp 103–113.
13. Campey, L.H.; Hyde, E.; Haygarth Jackson, A.R. *Chem. Br.* **1970**, *6*, 427–430.
14. Heller, S.R.; Milne, G.W.A.; Feldmann, R.J. *Science* **1977**, *195*, 253–259.
15. Attias, R. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 102–108.
16. Farmer, N.A.; O'Hara, M.P. *Database* **1980**, *3*, 10–25.
17. *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington D.C., 1987.
18. Corey, E.J.; Wipke, W.T. *Science* **1969**, *166*, 178–192.
19. Wipke, W.T. In *Computer Representation and Manipulation of Chemical Information*; Wipke, W.T.; Heller, S.R.; Feldmann, R.J.; Hyde, E., Eds.; John Wiley and Sons: New York, 1974; pp 147–174.
20. Howe, W.J.; Hagadone, T.R. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 8–15.
21. Howe, W.J.; Hagadone, T.R. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 182–186.
22. Anderson, S. *J. Mol. Graphics* **1984**, *2*, 83–90.
23. *Communication, Storage and Retrieval of Chemical Information*; Ash, J.E.; Chubb, P.A.; Ward, S.E.; Welford, S.M.; Willett, P., Eds.; Ellis Horwood: Chichester, 1985; Chapter 7.
24. Magrill, D.S. In *Chemical Structures: The International Language of Chemistry*; Warr, W.A., Ed.; Springer Verlag: Heidelberg, 1988; pp 53–62.
25. Warr, W.A. *J. Mol. Graphics* **1986**, *4*, 165–169.
26. *Chemical Structure Software for Personal Computers*; Meyer, D.E.; Warr, W.A.; Love, R.A., Eds.; ACS Professional Reference Book; American Chemical Society: Washington, D.C., 1988.
27. Johns, T.M. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington, D.C., 1987; pp 18–28.
28. Meyer, D.E. In *Chemical Structures: The International Language of Chemistry*; Warr, W.A., Ed.; Springer Verlag: Heidelberg, 1988, pp 251–259.
29. del Rey, D. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington, D.C., 1987; pp 48–61.
30. McDaniel, J.R.; Fein, A.E. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington, D.C., 1987; pp 62–79.
31. Johnson, A.P.; Cook, A.P. In *Modern Approaches to Chemical Reaction Searching*; Willett, P., Ed.; Gower: Aldershot, 1986; pp 184–193.
32. Lawson, A.J. In *Graphics for Chemical Structures: Integration with Text and Data*; Warr, W.A., Ed.; ACS Symposium Series 341; American Chemical Society: Washington, D.C., 1987; pp 80–87.