

NOTES

Databases in Chemistry

K. V. MURALIKRISHNA

*Computer Centre,**Sri Venkateswara University, Tirupati-517 502, India.*

This paper reports the significant applications of database packages in different branches of chemistry such as physical, inorganic, analytical and organic chemistry. Parallel to presenting such major applications of databases in chemistry, a few other applications in general nature have also been dealt with in this paper.

The term 'database' was an organized integrated and often inter-related collection of Computer-based data, records, files or information. The databases be broadly divided into reference and source databases. Reference databases are bibliographic in nature. Source databases contain original source data, the full text of original source of information. It databases can be classified as numeric, textual numeric and full text. Numeric databases consist of numeric information presented. Full text databases provide access to the complete text of documents. The advantage of this type of databases was the facility to perform free-text search. Legal databases which contain the full text of statutes and law reports provide excellent examples of this type. The advantage of database was hundreds of records in a subject can be searched with in minutes which is impossible in a manual system.

From the scientific literature it has been understood that the database systems have deeply penetrated into several other fields of science besides their important role in chemistry. With the emergence of database packages drafted for the astronomical research results: the data concerning astronomic, photometric and spectroscopic data about stars interstellar objects and other satellites could have therefore been systematically analysed with a greater precision¹. Goryanin *et al*² have developed a package towards analysis of databank on enzymes and metabolic pathways (DBEMP) for its used in PDP-11 and IBM PC computers. Sugawara³ has described the current status and future perspectives of the following database. Proteins and nucleic acids of biochemistry and the biosystems like viruses, microorganisms and other kinds of living biological systems, a workable data base has been in use in recent times. It is equally interesting to make a mention of the usage of the database packages in forensic investigations⁴ and it consists of the following three types. (i) Scientific literature, (ii) Analytical data and (iii) Casuistic data. In respect of pharmaceutical research also, the database systems are found to be playing an important role towards the statistical informations on the drugs⁵.

Databases in chemistry

A fortran 77 based software was developed to work on the mini computers towards the creation of database for undertaking computational manipulations and the data search operations and finally data printing process of the structural details of the compounds through the analysis of heat capacity⁶.

A more convenient database system was developed in order to recommend the solubility value in terms of its weighing factor that could be obtained provided we know the information about the solute temperature, equilibrium, purity, data accuracy and analysis. Thus the package would become an effective software tool in estimating the solubility of the compounds in the recommended solvents to overcome the difficult of the excess mixing in the solvents⁷. A simple database was brought forward for the benefit of chemists to understand the thermo physical properties of materials synthesized in the laboratories⁸.

Togasi and Ohno⁹ have reported an integrated Quantum Chemistry database (QCDB) that consists of four databases, concerning quantum chemistry literature, numerical, programme and estimations. A program was developed by Wang and Leshan¹⁰ would search out the chemical formula in aqueous solution thermodynamic database and inorganic thermochemistry database automatically depending on the input informations such as the elements or radicals of such chemical systems. With this programme atomic coefficients matrix can be processed.

A well defined database was formulated to understand the crystalline structures of certain synthesized solid materials at the laboratories¹¹. This is complete, up to date-data, accurate versatile and user friendly. A new database abbreviated as LABBASE 200 was described for the purpose of chromatographic analysis of a personal desk top computers. This package provides an opportunity to read chromatographic data directly and its automatic feeding into database for searching, sorting, plotting and finally data interpretation. Under this package, the examples considered are the files of drugs in pharmaceuticals and plant extraction in chemistry laboratory¹². A description was made through a database for radio caarbon dates¹³.

Now-a-days the NMR spectral analysis have become a vital optical method to investigate the structural details of the organic compounds. A computer program was described for the identification of organic compounds for structure and substructure searching and for simulation of new compounds through ¹³C NMR spectra. This can be used for the study of chemical structure property correlations¹⁴ in spectroscopic data other than ¹³C NMR. Noeth and Striedl¹⁵ described the factual database, B-Base. It incorporates approximately 3,300 document units concerning ¹¹B-NMR data. This is designed for Personal Computer and is interactive *via* Chem Base software. Marshall¹⁶ described chemical database CHEMBASE. The program allows the storage, manipulation and searching of a database in terms of chemical structures and sub structures, as well as in terms of complete reactions.

A data base dealing especially with designs, model and developmental facilities, a micro computer base package was described. This package consists of a useful

information on the bioactive trends in certain compounds defining their structures, biological activities and trade names. Thus, this package will become a source of information to analyse several other bioactive materials¹⁷.

Conclusions

The Databases in Chemistry are abundantly useful and very stimulating, to cater the needs of the students and researchers in Chemistry in finding solutions to their problems. Therefore, the users should an idea about the different databases available to enable them to solve their problems in Chemistry in no time.

ACKNOWLEDGEMENTS

My heartfelt gratitude to Dr B Rangamannar, Reader in Chemistry, S.V. University for his excellent cooperation and help in the present work.

REFERENCES

1. W. Bronkalla and E. Sehilbach *Sterne*, **64**, 149 (1988).
2. I.I. Goryanin, E.L. Shaevelev and I.A. Yunus, *Stud. Biophys*, **129**, 165 (1989).
3. Hideaki Sugawara, *Bio Ind.*, **5**, 599 (1988).
4. L.A. King, *Bio Chem. Soc. Trans*, **17**, 841 (1989).
5. Yutaka Kawamatsu and Mitsunu Sakai, *Kagaku to Kogyo*, **42**, 718 (1989).
6. Vlastimil Ruzicka and Milan Zabransky, *Chem. Prum.*, **38**, 543 (1988).
7. R. Dannenfelser and S.H. Yalkowsky, *Comput. Appl. Bio Sci.*, **5**, 253 (1989).
8. R. Brand and G. Jaroma-Weiland, *CFI, Ceram, Fourum Int.*, **66**, 122 (1989).
9. Mastomo Togasi and Kimio Ohno, *Kagaku, Zokam*, **106**, 219 (1985).
10. Peiming. Wang and Jisuani Leshan, *Yu Ying Yong Heaxue*, **3**, 25 (1986).
11. G. Bergerhoft, *Crystallog. Comput.; Data collect; Struct. Determ.; Proteins; Data bases; pap. Int. Summer Sch. 9th 1984 (pub. 1985). pp 85-95.*
12. Edward. C. Long, *Am. Lab.*, **16**, 112, 114 (1984).
13. Wieczyslaw F. Pazdur and Marian Porwol, *Zesz. Nank. Politech. Slask., Mat-Fiz* **56**, 19 (1987).
14. J. Zupan, M. Nonc, S. Bohance, M. Razinger, L. La, M. Tusar and I. Kosin, *Vestn. Slov. Kem. Drug*, **34**, 83 (1987).
15. H. Noeth, E. Striedl and Mitteisungsbl, *Gos. Dtsch. Chem; Fachgruppe Chem-Inf.*, **15**, pp. 3-10, (1989).
16. John C. Marshall *J. Chem. Inf. Comput. Sci.*, **27**, 47 (1987).
17. I. Paeva, Z. Lateva and G. Dimitrov, *Int. J. Bio-Med. Comput.*, **18**, 7 (1986).