1992RMxAA..23..119V

Rev. Mexicana Astron. Astrof., **23**, 119-125 (1992)

MOGADOC — A PERSONAL COMPUTER DATABASE FOR ATMOSPHERIC AND INTERSTELLAR MOLECULES IN MICROWAVE SPECTROSCOPY AND RADIO ASTRONOMY

J. Vogt

University of Ulm

RESUMEN. MOGADOC es una amplia base de datos de moléculas en la fase gaseosa que se investigan en la espectroscopía de microondas, en radio-astronomía y en la difracción de electrones. Contiene información sobre las propiedades eléctricas, magnéticas, dinámicas y espectroscópicas de compuestos inorgánicos, orgánicos y organometálicos en la fase gaseosa. Como característica especial, esta base de datos interna que corre en computadores personales por medio del bien conocido lenguaje de acceso *Messenger*, contiene conjuntos de datos sobre parámetros estructurales tales como las distancias internucleares y los ángulos de enlace.

ABSTRACT. MOGADOC is a comprehensive database for gas-phase molecules, investigated by microwave spectroscopy, radio astronomy and electron diffraction. It contains data on electrical, magnetic, dynamical and spectroscopic properties of inorganic, organic and organometallic compounds in the gas phase. As a special feature the in-house database, which can be run on a personal computer by means of the well known Messenger retrieval language, contains numerical data sets for structural parameters such as internuclear distances and bond angles.

Key words: INTERSTELLAR MOLECULES – MOLECULAR PROCESSES – RADIO-SOURCES: SPECTRA

I. INTRODUCTION

Molecular physical properties and structural parameters of gas-phase compounds are often published in sources that are not well known, causing these data to be easily overlooked. Thus in order to facilitate the access to structural and related properties of free molecules, the Section for Spectra and Structure Documentation at the University of Ulm has compiled and critically evaluated for more than two decades literature in the field of high-resolution spectroscopy (especially in the long-wavelength region) and gas-phase electron diffraction (Lohr et al. 1983, Buck et al. 1981, Herde et al. 1985, Starck et al. 1977, Callomon et al. 1976, Callomon et al. 1987, Starck 1967, Demaison et al. 1974, Demaison et al. 1982, Brown et al. 1983).

This complete and worldwide unique compilation has been the basis for the machine-readable database MOGADOC. This acronym stands for MOlecular GAs-phase DOCumentation. In connection with other in-house database projects of Fachinformationszentrum Energie Physik Mathematik in Karlsruhe, the European base of STN International (The Scientific and Technical Information Network), a special database management system has been developed for IBM and IBM-compatible personal computers under MS-DOS disk operating system. This in-house database, which can be run locally by the Messenger retrieval language (widely used for STN International online databases), enables the user to trace back literature by means of bibliographic, physical and chemical search terms. As for the STN International online databases many interactive help facilities are accessible.

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V

II. SCOPE OF THE DATABASE

The MOGADOC database is a comprehensive machine-readable data compilation for gas-phase molecules, which have been investigated by microwave spectroscopy, radio astronomy as well as electron diffraction. It contains data on spectroscopic, structural, dynamical, electric and magnetic properties of inorganic, organometallic and organic compounds. All information has been compiled mainly from scientific journals. Moreover the grey literature such as dissertation theses, reports and conference proceedings is also included. Hereby relevant abstracts from the Austin Symposia on Molecular Structure, the Ohio Symposia on Molecular Spectroscopy as well as from the European High Resolution Molecular Spectroscopy Conferences are taken into account. The literature is recorded back to

- 1930 for gas-phase electron diffraction
- 1945 for microwave spectroscopy
- 1960 for molecular radio astronomy

In total about 16,500 references (until 1991) are presently implemented for about 4500 compounds.

III. STRUCTURE OF THE DATABASE

Different types of information are stored in separate fields of the database. These fields form two greater entities called files. Hereby the fields compound number (CN) and sum formula in the Hill system (MF) belong to the bibliographic file MGDLIT and the compound file MGDCOM as well.

The bibliographic file MGDLIT contains information on authors, document titles, sources and publication years. Moreover errata are given and taken into account. The content of the quoted publications is characterized by means of keywords, which form a hierarchical controlled vocabulary of about 1900 items. The topics of the keywords are summarized in Table 1. They describe the content of documents in much more detail than titles can do. Experts, who specially review the documents, select appropriate keywords. The presence of keywords devoted to theoretical calculations (such as CNDO, MNDO, *ab initio*, etc.) and force field calculations implies that these papers also report results from microwave spectroscopy, molecular radio astronomy or electron diffraction. Because of the special scope of MOGADOC pure theoretical papers on CNDO, MNDO, *ab initio* and force constants are not implemented. However, publications that only deal with theoretical aspects in microwave spectroscopy, rotation-vibration spectroscopy, molecular radio astronomy and gas-phase electron diffraction are accessible in MOGADOC.

The compound file MGDCOM deals with the corresponding compounds. It contains compound names, synonyms, Chemical Abstracts registry numbers, Hill sum formulas, structural formulas, stereo drawings as well as conformational descriptions. Various entries also list numerical data with experimental uncertainties for structural parameters such as internuclear distances and bond angles. Presently about 500 numerical data sets are available in MOGADOC.

It is emphasized that no other parameters are given numerically. However, the keywords give hints to the appropriate references for rotational and centrifugal distortion constants, rotation vibration interaction parameters, hyperfine structure constants, magnetic susceptibilities, dipole moments, intensity coefficients, pressure broadening parameters, potential barrier parameters, etc.

IV. FEATURES OF THE DATABASE

Because of the detailed structure of MOGADOC rather specific retrievals can be performed. Hereby truncation and nested Boolean expressions are allowed. The features of the database may be illustrated by means of typical retrieval examples. For example the command

```
search C3 H2/MF and cycl?                    (1)
```

applied in the compound file MGDCOM, looks for compounds with the sum formula (MF) C_3H_2. The second part of the command requires that the compound name or synonym contains the fragment terms "cyclo" or

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V

TABLE 1: Topics of keywords that characterize the content of the quoted documents and form a hierarchical controlled vocabulary.

1. classification of molecular type
2. spectral ranges, spectroscopic methods and methods of assignment
3. rotational and rotation vibration interaction parameters
4. structure of free molecules
5. nuclear quadrupole hyperfine structure
6. Zeeman effect of diamagnetic molecules
7. technical details of microwave setups
8. dipole moments
9. Stark effect
10. collisional effects (pressure broadening, line shifts, etc.)
11. intensity measurement and line shapes
12. large amplitude motions: internal rotation, ring puckering, inversion, pseudo-rotation, quasilinear and quasisymmetric top molecules, loosely bound complexes
13. technical details of electron diffraction setups
14. theoretical calculations: CNDO, MNDO, *ab initio*, etc.
15. specific isotopic labels
16. spectra of linear paramagnetic molecules
17. spectra of non-linear paramagnetic molecules
18. Zeeman effect of paramagnetic molecules
19. preparation of unstable molecules
20. astrophysical and environmental observations
21. force constants
22. theory on rotation vibration spectroscopy

"cyclic". The fragment terms are stored in the basic index (BI) of the compound file. Hereby the question mark ? serves as truncation symbol.

As for the STN online databases, run by Messenger retrieval language, MOGADOC gives detailed protocols on the computer monitors by reporting intermediate results, which enables the user to interpret the size of the final answer sets. In the case of Command (1) the following display may result:

```
      2    "C3 H2"/MF
    551    CYCL?/BI
L15   1    C3 H2/MF AND CYCL?/BI
```

According to this protocol there are two entries for the Hill formula C_3H_2 and 551 entries for compounds, whose compounds names or synonyms contain the fragment terms "cyclo" or "cyclic" in the basic index of the compound file. However, only one entry fulfills the logical requirement in Command (1) and forms the answer set L15. In this example the resulting final answer set deals with 2-cyclopropen-1-ylidene. Due to the second part of Command (1) the entry for the non-cyclic isomer vinylidenecarbene (propadienylidene), $H_2C{=}C{=}C{:}$, is excluded. In Figure 1 the resulting printout is given.

In order to get the bibliographic entry, which deals with this cyclic compound, a file crossover should be first performed by the command

```
file MGDLIT                                                    (2)
```

It switches retrieval from the compound file to the bibliographic file MGDLIT. In the case that not all bibliographic entries for this compound are desired, the retrieval command should be specified in more detail. For example the command

```
search L15 and (laboratory and (astro?  or RAS/CT)) and
     PY>=1986 not C/DT                                         (3)
```

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

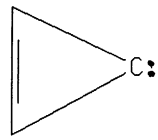
1992RMxAA..23..119V

MGDCOM FILE SEARCH RESULTS 09.08.1991 12:39:53

L1 ANSWER 1 OF 1

CN **102408**
MF C3 H2
RN 16165-40-5
NA 2-Cyclopropen-1-ylidene

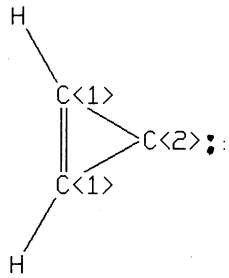
SF



Method of study : Microwave Spectroscopy

Temperature : Not specified

SD



© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V


```
MGDCOM FILE SEARCH RESULTS        09.08.1991  12:39:53

Refined geometrical parameters

Errors of values are parenthesized in units of the last digit.

Structure type for distances: r(s)

Parameter(s)      | Value(s)      | Note(s)
                  | (Angstroem)   |
------------------|---------------|---------
C1-C2             | 1.4195        |    a
C1-C1             | 1.3242        |
C1-H              | 1.0754        |

Note(s):
a) error limits not given
   however, error limits given for Cartesian coordinates

Structure type for angles: r(s)

Parameter(s)      | Value(s)      | Note(s)
                  | (degrees)     |
------------------|---------------|---------
C1-C2-C1          |  55.607       |    a
C1-C1-H           | 149.829       |

Note(s):
a) error limits not given

Other relevant information:

C(2v) symmetry assumed.
```

FIGURE 1 (including page before): Printout of the answer set obtained by the compound search specified in Command (1).

looks for all bibliographic entries dealing with this cyclic compound. The resulting answer set is then narrowed down by requiring the term "laboratory" as well as terms such as "astronomical", "astronomy" or "astrophysical" within the document titles. Instead of the terms "astro?" within the titles, the desired documents should be alternatively devoted to radio astronomy. Those documents are annotated by the keyword (CT) RAS, an acronym for radio astronomy. Furthermore the command looks for recent publication since 1986 inclusively. Finally it excludes conference proceedings. The following protocol may result:

```
      116  ASTRO?
     2007  RAS/CT
      106  LABORATORY
       28  102408
     3725  PY>=1986
     2746  C/DT
L16     1  L15 AND (LABORATORY AND (ASTRO? OR RAS/CT)) AND
             PY>=1986 NOT C/DT
```

From 16,500 bibliographic entries 116 of them contain the term "astro?" (such as "astronomical", "astronomy" or "astrophysical") within the document titles whereas 106 documents have the term "laboratory" in their titles. On the other hand 2007 bibliographic entries are attributed to radio astronomy, independent of whether this

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V


```
MGDLIT FILE SEARCH RESULTS        09.08.1991  12:54:07


L2  ANSWER 1 OF 1

AN  54462
AU  Vrtilek, J. M.; Gottlieb, C. A.; Thaddeus, P.
TI  Laboratory and astronomical spectroscopy of C3H2, the first interstellar
    organic ring.
SO  Astrophys.J., 314, 716
CO  9
PY  1987
DT  J
IN  1   CN  102408
        MF  C3 H2
        KW  ASYMTOP         asymmetric top
            MW              microwave or millimeter wave spectroscopy
            RAS             radioastronomy
            MWRMW           microwave region: 2 - 100 GHz
            RCEGRST         rovibrational constants: ground state
            RCOCFD          rotational constants (nonrigid rotor)
            RCOALL          all rotational constants determined
            INDGRST         ground state inertial defect
            CDQCEXSP        quartic centrifugal distortion constants from
                            experimental spectrum
            SEG             structure in ground state
            DISTO           r(0) distance(s)
            ANGLO           r(0) angle(s)
            APOIM           interstellar molecule
            APOCD           astrophysical column density estimated
```

FIGURE 2: Printout of the corresponding bibliographic entry for 2-cyclopropen-1-ylidene. The answer set was obtained by the file crossover and further search demonstrated in Commands (2) and (3). The keywords are given as acronyms with full-text explanations.

method is mentioned within the titles or not. 28 documents report investigations on 2-cyclopropen-1-ylidene, whose MOGADOC specific compound number is 102408. Among 16,500 documents there are 3725 entries since 1986 (inclusively) and 2746 conference contributions. It is emphasized that the six partial results are related to the total amount of documents. They do not necessarily deal with identical references. However, only one document fulfills the logical requirement in Command (3). The resulting bibliographic answer set is given in Figure 2.

V. OUTLOOK

The release of the MOGADOC database is set for the end of 1991 and will be distributed by means of floppy diskettes. It is planned to update the database every year. Moreover the implementation of more recent and even older numerical data will be continued. Furthermore high-resolution infrared papers will be added to MOGADOC.

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V

Acknowledgments.

The author would like to express his sincere thanks to Prof. H. K. Bodenseh, Prof. I. Botskor, Dr. D. Boucher, Dr. B. N. Cyvin, Prof. S. J. Cyvin, Dr. J. Demaison, Prof. H. Dreizler, Prof. J. L. Duncan, Prof. I. Hargittai, Dr. M. Hargittai, Prof. W. Hüttner, Prof. H. Jones, Dr. Z. Kisiel, Dipl.-Chem. G. R. Kramer, Dipl.-Math. A. Lohr, Dr. A. H. Lowrey, Dr. B. Mez-Starck, Dipl.-Chem. R. Mutter, Prof. K. P. R. Nair, Prof. H. Oberhammer, Prof. H. D. Rudolph, Prof. A. Ruoff, Dr. K. H. Schmidt, Dr. K. Volka, Dr. D. G. Watson, Prof. M. Winnewisser, Dr. G. Wlodarczak, and K. M. T. Yamada for their stimulating discussions and valuable contributions. The project has been supported by Fachinformationszentrum Energie Physik Mathematik in Karlsruhe.

REFERENCES

Brown, J. M., Demaison, J., Dubrulle, A., Hüttner, W., and Tiemann, E.: Molecular Constants Mostly from Microwave, Molecular Beam, and Electron Resonance Spectroscopy. Landolt Börnstein New Series II, Vol. 14b, Springer, Berlin, 1983.

Buck, I., Maier, E., Mutter, R., Seiter, U., Spreter, C., Starck, B., Hargittai, I., Kennard, O., Watson, D. G., Lohr, A., Pirzadeh, T., Schirdewahn, H. G., and Majer, Z.: Bibliography of Gas-Phase Electron Diffraction 1930 - 1979. Physics Data No. 21-1, Fachinformationszentrum Energie Physik Mathematik, Karlsruhe, 1981.

Callomon, J. H., Hirota, E., Kuchitsu, K., Lafferty, W. J., Maki, A. G., Pote, C. S., Buck, I., and Starck, B.: Structure Data of Free Polyatomic Molecules. Landolt Börnstein New Series I I, Vol. 7, Springer, Berlin,

Callomon, J. H., Hirota, E., Iijima, T., Kuchitsu, K., Lafferty, W. J.: Structure Data of Free Molecules. Landolt Börnstein New Series II, Vol. 15, Springer, Berlin, 1987.

Demaison, J., Hüttner, W., Starck, B., Buck, I., Tischer, R., and Winnewisser, M.: Molecular Constants from Microwave, Molecular Beam, and Electron Spin Resonance Spectroscopy. Landolt Börnstein New Series II, Vol. 6, Springer, Berlin, 1974.

Demaison, J., Dubrulle, A., Hüttner, W., and Tiemann, E.: Molecular Constants Mostly from Microwave, Molecular Beam, and Electron Resonance Spectroscopy. Landolt Börnstein New Series II, Vol. 14a, Springer, Berlin, 1982.

Herde, E., Maier, E., Mez-Starck, B., Mutter, R., Seiter, U., Spreter, C., Hargittai, I., Watson, D. G., Lohr, A., and Selz, G. J.: Bibliography of Gas-Phase Electron Diffraction. Supplement 1980 - 1982. Physics Data No. 21-2, Fachinformationszentrum Energie Physik Mathematik, Karlsruhe, 1985.

Lohr, A., Mez-Starck, B., Schirdewahn, H. G., and Watson, D. G.: J. Mol. Struct., 97 (1983) 57.

Starck, B.: Molecular Constants from Microwave Spectroscopy. Landolt Börnstein New Series II, Vol. 4, Springer, Berlin, 1967.

Starck, B., Mutter, R., Spreter, C., Kettemann, K., Boggs, A., Botskor, M., and Jones, M.: Bibliography of Microwave Spectroscopy 1945 - 1975. Physics Data No. 9-1. Fachinformationszentrum Energie Physik Mathematik, Karlsruhe, 1977.

J. Vogt: Sektion für Spektren- und Strukturdokumentation, Universität Ulm, P.O. Box 4066, D-7900 Ulm, Federal Republic of Germany.

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System

1992RMxAA..23..119V

© Universidad Nacional Autónoma de México • Provided by the NASA Astrophysics Data System