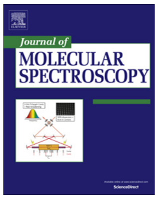
Contents lists available at ScienceDirect

Journal of Molecular Spectroscopy

journal homepage: www.elsevier.com/locate/jms



The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC

Christian P. Endres [1], Stephan Schlemmer, Peter Schilke, Jürgen Stutzki, Holger S.P. Müller *

I. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany

ARTICLE INFO

Article history:
Received 20 January 2016
In revised form 2 March 2016
Accepted 9 March 2016
Available online xxxx

Keywords:
Database
Line identification
Rotational spectroscopy
Hamiltonian
Radio astronomy
Astrochemistry

ABSTRACT

The Cologne Database for Molecular Spectroscopy, CDMS, was founded 1998 to provide in its catalog section line lists of mostly molecular species which are or may be observed in various astronomical sources (usually) by radio astronomical means. The line lists contain transition frequencies with qualified accuracies, intensities, quantum numbers, as well as further auxiliary information. They have been generated from critically evaluated experimental line lists, mostly from laboratory experiments, employing established Hamiltonian models. Separate entries exist for different isotopic species and usually also for different vibrational states. As of December 2015, the number of entries is 792. They are available online as ascii tables with additional files documenting information on the entries.

The Virtual Atomic and Molecular Data Centre, VAMDC, was founded more than 5 years ago as a common platform for atomic and molecular data. This platform facilitates exchange not only between spectroscopic databases related to astrophysics or astrochemistry, but also with collisional and kinetic databases. A dedicated infrastructure was developed to provide a common data format in the various databases enabling queries to a large variety of databases on atomic and molecular data at once.

For CDMS, the incorporation in VAMDC was combined with several modifications on the generation of CDMS catalog entries. Here we introduce related changes to the data structure and the data content in the CDMS. The new data scheme allows us to incorporate all previous data entries but in addition allows us also to include entries based on new theoretical descriptions. Moreover, the CDMS entries have been transferred into a mySQL database format. These developments within the VAMDC framework have in part been driven by the needs of the astronomical community to be able to deal efficiently with large data sets obtained with the Herschel Space Telescope or, more recently, with the Atacama Large Millimeter Array.

1. Introduction

The unambiguous assignment of a spectral line feature observed in space to a specific molecular or atomic species and its quantitative interpretation, e.g. in terms of column densities or relative abundances, depends decisively on the availability of rest frequencies and intensity information with sufficient accuracy. Such information is usually generated from laboratory data, sometimes supplemented by astronomical observations for rest frequencies and quite often by quantum chemical calculations for intensity information. The reliability of the derived frequencies and intensities depends on the quality of the Hamiltonian model, the accuracies of the rest frequencies and of the intensity information, but also very much on the proper judgement of the line frequency accuracies. We will discuss important aspects on this topic in Section 2.

The CDMS[2] [1,2] was founded in the year 1998 to provide such information in its catalog section following the example of and complementing the JPL catalog[3] [3] because of limited activities in the JPL catalog in recent years. The emphasis in the CDMS catalog is put on molecular species of interest for astrophysics or astrochemistry which have been or may be observed by radio astronomical means. The catalog also contains entries for fine structure transitions of atoms, such as C, ^{13}C, C^+, $^{13}C^+$, N^+, O, and Al. In selected cases, rovibrational transitions are available not only in the submillimeter (sub-mm) or far-infrared (FIR) region, such as for C_3 or C_3O_2, but

---

* Corresponding author.
E-mail address: hspm@ph1.uni-koeln.de (H.S.P. Müller).

[1] Present address: Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany.

[2] http://www.astro.uni-koeln.de/cdms/.

[3] http://spec.jpl.nasa.gov/.

http://dx.doi.org/10.1016/j.jms.2016.03.005
0022-2852/© 2016 Elsevier Inc. All rights reserved.

sometimes also in the mid-IR, such as for CH^+, CCH, HCO^+, and N_2H^+. Some of the species are also important for studies of planetary atmospheres, including that of Earth, for example HCN, H_2CO, PH_3, CH_3CN, and SO_2. Individual entries are created for different isotopologues as far as they may be relevant for astronomers. In addition, transitions of different vibrational states are usually provided in different entries as well; they may be in one single entry for simple diatomics, such as NaCl and KCl, or strongly coupled vibrational states. The number of entries approaches 800 at the end of 2015; an up-to-date list is available online.[4] These entries, together with those in the JPL catalog, cover a large fraction of the almost 200 molecules detected in the interstellar medium (ISM) or in circumstellar envelopes (CSEs)[5] of late type stars with minor isotopologues or excited vibrational states. In fact, the CDMS and JPL catalogs were instrumental for the analyses of spectral recordings obtained, e.g., with the Herschel Space Telescope. In addition, data collections, such as the Leiden Atomic and Molecular DAtabase (LAMDA) [4], CASSIS [5], and splatalogue [6,7], draw heavily on the entries of the CDMS and JPL catalogs. Other databases, such as GEISA [8,9] and HITRAN [10], use many data from the CDMS and JPL databases, in particular in their pure rotational parts.

The line list entries are generated from spectroscopic parameters, their uncertainties, and their correlation coefficients. These values are obtained from experimental data employing established Hamiltonian models. The experimental data usually stem from laboratory spectroscopic investigations, but data from astronomical observations can also be used as far as appropriate. In most cases, the SPFIT/SPCAT program suite [11] is used for fitting and prediction of molecular spectra. Other programs have been used on occasion, details are provided in Section 2.2.

In the early years of the CDMS, considerable effort had been put on terahertz data of light hydrides, such as entries on H_2D^+, HD_2^+, CH, CH^+ and its isotopologues, CH_2, NH, ND, NH_2, $^{15}NH_3$, OH^+, H_2O^+, SH^+, and H_2S, H_2Cl^+, and ArH^+ along with their isotopologues. The aim was to support missions such as the Herschel Space Telescope [12], especially with respect to its high resolution Heterodyne Instrument for the Far-Infrared, HIFI [13], and the Stratospheric Observatory For Infrared Astronomy, SOFIA [14], with its German REceiver for Astronomy at Terahertz frequencies, GREAT [15]. *Herschel* in particular has expanded our knowledge in the field of terahertz astronomy tremendously [16], in particular with respect to hydrides. One of the main goals were low-energy rotational transitions of H_2O and its isotopologues [17]. Large scale molecular line surveys were another important goal. They led to the detection of several hydrides in Galactic or extragalactic sources, the latest detection being ArH^+ [18,19], a tracer of the almost completely atomic ISM, that was also detected in an extragalactic source recently with ALMA [20]. Important discoveries with SOFIA/GREAT include the detection of SH [21] and of the ground state rotational transition of *para*-H_2D^+ [22]; the latter observation benefitted from earlier laboratory work [23], see also Section 2.1. Terahertz observations are possible to some extent even from the ground, as demonstrated, e.g., by the early observations of OH^+ [24] and SH^+ [25] with the Atacama Pathfinder EXperiment, APEX [26].

The focus of new or updated entries in the CDMS has shifted somewhat with the advent of the Atacama Large Millimeter Array, ALMA. Due to the very high sensitivity and high spatial resolution, which reduces confusion caused by multiple source components in the telescope beam, interferometers are superior to single dish telescopes in detecting complex organic molecules as exemplified by the detection of *iso*-propyl cyanide (i-C_3H_7CN) as the first branched alkyl compound [27] whose detection was only possible because of fairly recent laboratory work [28]. Transitions belonging to vibrationally excited states are easier to detect, e.g. highly excited HNC toward the carbon-rich late-type star CW Leonis [29]; such transitions provide deeper insight into excitation mechanisms in general and insight into the dust formation zone in the particular case of late-type stars. Interferometric observations of late-type stars provide considerably more detail than single dish observations and can lead to the detection of new species, such as the recent detection of TiO_2 and the radio astronomical detection of TiO in the circumstellar envelope of the O-rich AGB star VY Canis Majoris [30]. TiO_2 was also observed with ALMA during science verification observations [31]. Interferometric observations will also be very beneficial for the investigations of extragalactic sources, such as the unnamed foreground galaxy in the direction of the quasar PKS 1830-211 [32], in which also ArH^+ was detected [20].

Radio astronomical spectra are often analyzed under the assumption of local thermodynamic equilibrium (LTE). This is usually a good assumption in warm and dense parts of molecular clouds, but deviations from LTE may be considerable at lower temperatures, e.g. in dark clouds, or in less dense regions, especially in the diffuse ISM. In these cases it is often necessary to take into account collisional processes with H_2 and He in the denser ISM, and/or with H and electrons in the more diffuse ISM. A review on collisional processes in the ISM has been published fairly recently [33]. Accurate data are needed for a plethora of molecules to avoid misinterpretations. An interesting and important aspect was the detection of seemingly more HNC than the more stable HCN. These results were obtained by assuming the somewhat similar molecular structure would lead to similar collisional rates with He and H_2. In the absence of collisional data for HNC, those of HCN were used instead. Initial calculations of HNC with He and later with H_2 [34] revealed that the collisional rates are in fact very different and the column densities of HNC had to be revised to values lower than those of HCN in all instances.

Collisional rates are available from laboratory experiments for some systems. In many cases, however, such data are not available, and often it may be difficult to obtain such data. Resorting to quantum chemical calculations is a common alternative. But the calculation of collisional processes by quantum chemical means with high accuracy is demanding even nowadays. Therefore, calculations are often restricted to He as the collider and rates of collisions with H_2 are often estimated from these. Also, these calculations consider mostly collisions with small molecules, which play a more important role in cold or less dense environments than complex organic molecules. Nevertheless, the need for collisional data involving more complex molecules has become increasingly apparent in recent years. The complex absorption and emission spectrum of methyl formate (CH_3OCHO) in the GBT PRIMOS survey of Sagittarius B2(N) between ~0.3 and ~50 GHz is an interesting case because the observed low energy transitions sample the less dense envelope of this source, and their intensity distribution, including weak maser activity in some lines, could not be explained by an LTE model [35].

The need to access an increasing number of accurate molecular spectroscopy data together with an increasing number of accurate collisional data has become apparent in recent years. An early effort to link molecular spectroscopy data (here from the CDMS and JPL catalogs) with collisional data in a virtual observatory compliant environment was undertaken within the BASECOL collisional data base [36,37]. Eventually, this effort led to the Virtual Atomic and Molecular Data Centre (VAMDC)[6] [38], an

---

[4] http://www.astro.uni-koeln.de/cdms/entries.

[5] See, e.g., http://www.astro.uni-koeln.de/cdms/molecules.

[6] http://www.vamdc.eu/; http://www.vamdc.org/.