What Constitutes Spectroscopic Proof for the Detection of Large "Hot Core" Molecules?

Lucy M. Ziurys and Aldo J. Apponi

Departments of Chemistry and Astronomy, University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721 USA

Abstract. The presence of large organic species in interstellar gas has important implications for the origin of life and pre-biotic chemistry. Accurate identifications of such molecules, however, are problematic in molecular clouds. There are many reasons for such difficulties. One is that the spectral density is very high in objects where the chemistry is sufficiently complex to produce such species—at least 10 lines per 100 MHz in Sgr B2(N), for example, at 3 mm, at a sensitivity of 10 mK peak-to-peak. Hence, the possibility of chance coincidences is large. Another reason is the presence of many large organic molecules at relatively high temperatures; these large asymmetric tops have vast numbers of favorable transitions under these conditions, including those originating from low-lying vibrational states. Confusion and blending of transitions of one large molecule with those of another add to the risks of an inaccurate identification.

A case in point is that of glycolaldehyde, CH_2OHCHO. In order to confirm the identification of this molecule in Sgr B2(N), we searched for its most favorable transitions in the 2 and 3 mm windows, spanning the energy range of $\sim$10–100 K—a total of 43 individual transitions. Of all these lines, only seven were not heavily blended or contaminated by other molecules, i.e. so-called "clean" features. Emission, however, was visibly detected at 34 of the other transitions, but one transition was clearly absent. The "missing" line corresponds to a weak transition originating in the $K_a = 3$ ladder, and its absence is consistent with the other detected features. Based on the clean features only, glycolaldehyde has a $V_{LSR} = 61.7 \pm 1.5$ km s^{-1} and $\Delta V_{1/2} = 7.8 \pm 1.8$ km s^{-1} with intensities consistently in the range $T_R^* \sim 20-70$ mK. Given these data, the identification of glycolaldehyde is 99.9% secure. A rotational diagram from this data set yields a column density of $N_{tot} \sim 6 \times 10^{13}$ cm^{-2} for CH_2OHCHO—roughly a factor of 27 less than that of H_2CO. These data illustrate the problems and subtleties in identifying large organic molecules in space.

Keywords. astrobiology — astrochemistry — ISM: abundances — ISM: molecules — molecular processes — radio lines: ISM

---

1. Introduction

The large molecules found in dense interstellar clouds are typically simple organic species containing common functional groups. Ketones, aldehydes, acids, esters, and amides, for example, are well represented in interstellar molecules, as demonstrated by the presence of compounds such as dimethyl ether (CH_3OCH_3), formic acid (HCOOH), methyl formate ($HCOOCH_3$), acetaldehyde (CH_3CHO), and formamide (NH_2CHO). In fact, interstellar space is the one common place that organic molecules are routinely found other than in living organisms on Earth.

The presence of organic molecules in these two vastly different environments suggests that there may be a connection. In fact, it is possible that interstellar chemistry actually supplied the initial compounds that led to terrestrial life. It is now thought that the evolution of life on Earth did not always result in the best biochemical solutions (e.g. Benner *et al.* 2004). There are many other chemical possibilities for living systems. For
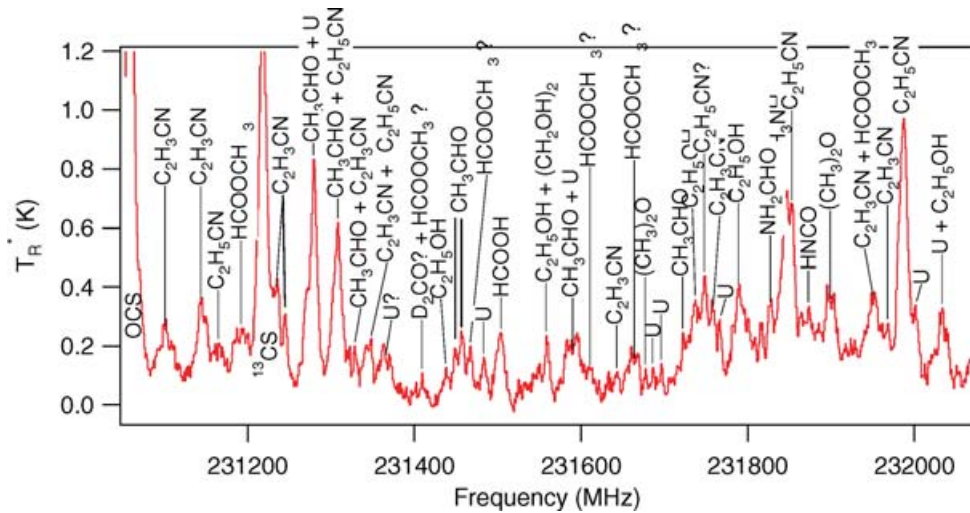
Figure 1. Single-sideband spectrum of Sgr B2(N) taken with the ARO SMT 10-meter telescope near 231.6 GHz. With an integration time of 6 hours, the spectrum is near the confusion limit with no baseline regions at the theoretical $3\,\sigma$ noise level ($\sim$10 mK).

example, it is not clear why ribose and deoxyribose are used in DNA and RNA. Why not glycol, a tetrose or even a hexose? Work by Albert Eschenmoser (1999) has suggested that other sugars are viable substitutes for ribose in the DNA structure. Why are only twenty standard amino acids used in living proteins? Studies have shown that there are no biochemical reasons to exclude other unnatural amino acids in proteins (e.g. Hohsaka & Masahiko 2002). However, life may have chosen one set of chemical possibilities over others not because they were the best solution to the problem, but because of synthetic contingency, namely, that what evolved is a result of what was available as starting materials. Could interstellar chemistry have provided these contingencies via molecule formation in dense clouds? Such compounds may have been transported to planet surfaces via comets and meteorites, and hence became the substances that led to living systems.

To examine such a hypothesis, it is clearly necessary to establish what chemical compounds are present in dense molecular clouds and in what quantities. Only then can the pieces be put together to examine the possible contribution of interstellar molecules to early biochemistry. However, accurate identification of such large organic species is not trivial. In this paper, we discuss such difficulties and formulate criteria for secure detections of large molecules in dense interstellar clouds.

2. The Millimeter-Wave Spectrum of Sgr B2(N)

The spectrum of Sgr B2(N) shown in Figure 1 is a good illustration of a confusion limited spectrum. This 6-hour integration on Sgr B2(N) using the ARO SMT 10-meter telescope has a theoretical $3\,\sigma$ noise level of about 10 mK. No significant baseline regions are found in this spectrum at this level, indicating that the entire bandpass is made up of partially overlapping emission features. Only the strongest of these features are easily identifiable; i.e., those belonging to vinyl cyanide, carbon monosulfide, ethyl cyanide, and carbonyl sulfide. The intensity of these lines typically are greater than 500 mK. Lines below that level start to become contaminated with other features and show up as blends of two or more transitions. These data illustrate the difficulty in unambiguously assigning the lines near the confusion limit.
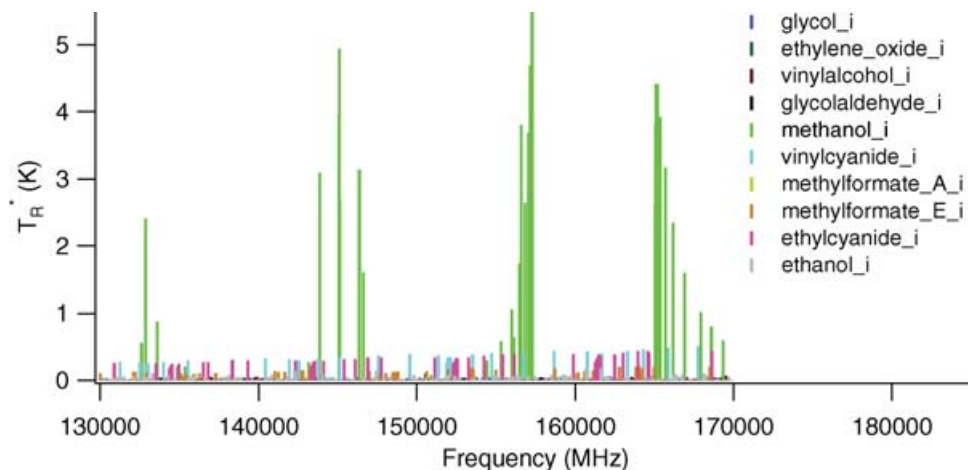
Figure 2. Simulated 2 mm spectrum of Sgr B2(N) using the column density and rotational temperatures given in the literature for 11 interstellar molecules of varying abundance. The combination of only these few molecules quite effectively shows that source confusion starts to become very significant in Sgr B2(N) at about the 500 mK level. See the on-line presentation for the color version of this figure.

The lines that make up the confusion-limited spectra of molecular clouds in general can be attributed to relatively large asymmetric top organic molecules. These types of species possess intrinsically complex spectra owing to their three unique moments of inertia and their system of complex energy levels that contribute to large rotational partition functions. The result is a dense population of rotational transitions that peak in the millimeter region of the spectrum.

Figures 2 and 3 illustrate the three levels of complexity that can be expected while observing astronomical sources containing large organic molecules. Figure 2 is a model of the 2 mm spectrum of Sgr B2(N), only considering the 11 molecules listed in the legend. At first glance, the spectrum is dominated by strong emission lines of the most abundant species present (in this case methanol). These lines are very strong (several Kelvin) and are sparsely distributed throughout the spectrum. It is noteworthy that the density of lines this strong is small (only a few per 5000 $km\,s^{-1}$); hence, any contamination of weak features on these lines will be insignificant. These lines are easily identifiable.

The next level of complexity found in Sgr B2(N) is at about the 300 to 500 mK level. Here a pattern of lines of increasing density is found, resulting from molecules that are of lesser abundance, as well as weaker features from molecules that are of high abundance. The emission features of methyl formate, ethyl cyanide and vinyl cyanide were selected for this illustration. These lines are about the same size as many of the weak methanol lines, which adds a level of complexity. The transitions are more densely spaced because these molecules are generally heavier and more asymmetric. The line density at this level of sensitivity is about an order of magnitude higher.

The last level of complexity is at about the 40 mK level, where there is nearly a continuum of closely space lines (see Figure 3). The weak features of the molecules of high and intermediate abundance plus the strongest features of yet lesser abundant species all contribute at this level. The molecules selected to illustrate this case are ethanol, ethylene glycol, ethylene oxide, vinyl alcohol and glycolaldehyde. Even with these few additional species, the problem in assigning spectra becomes very apparent. The line density increases yet again by about an order of magnitude to about 5 lines per
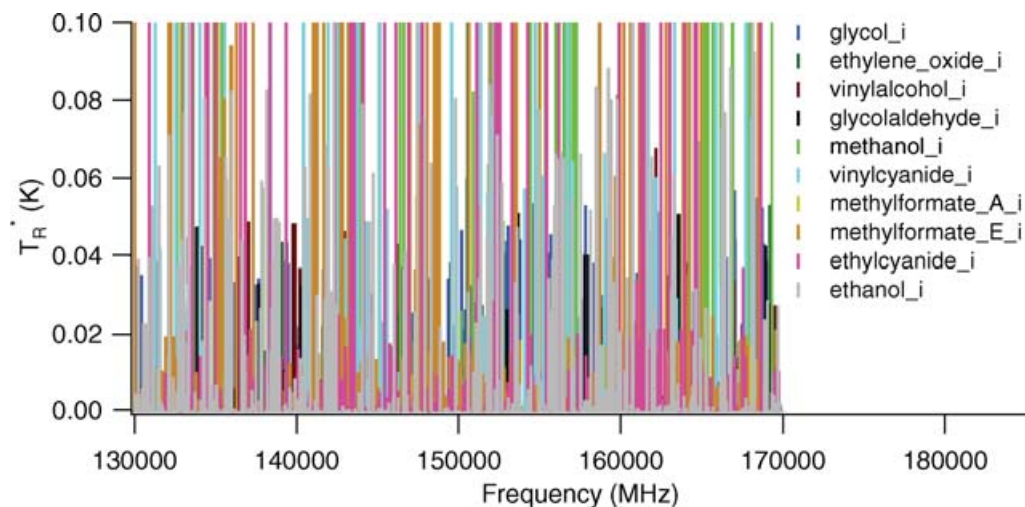
Figure 3. Same as figure 2, but the intensity scale is magnified to illustrate the increased complexity that occurs at about the 40 mK level in Sgr B2(N).

100 km s^{-1} in Sgr B2(N). Contamination at this level is inevitable and clean transitions will be few in number.

3. Approaching This Complex Problem

The statistical probability of finding a particular spectral feature of a molecule at a particular frequency is a function of: (i) the physics associated with the transition under investigation (N_{tot}, T_{rot}, S_{ij}, E_u, etc.) and (ii) the number of abundant contaminating species with favorable transitions that significantly overlap the same spectral region. Ideally, we would like the first of these two factors to dominate, and we have all become accustomed to assuming that this is true. However, as the sensitivity of an observation is increased, the second factor becomes significant. Hence, only through the observation of a sufficient number of clean lines (lines free from significant contamination) can a molecule be unambiguously identified near the confusion limit.

To date, spectral surveys at millimeter wavelengths have not been sensitive enough to identify new molecules or even some weak known molecules. For instance, Nummelin *et al.* (2000) surveyed the 1.3 mm atmospheric window down to about 100 mK peak to peak and Friedel et al. (2004) sparsely covered the 3 mm window at modest sensitivity compared to that required for identifying new molecules. Neither of these surveys of Sgr B2(N) was able to detect glycolaldehyde or ethylene glycol, nor were they able to help confirm or dismiss other recently claimed molecular identifications. To identify new asymmetric organic molecules with low abundance, it is clear that higher sensitivity is needed. Moreover, owing to high contamination levels, a vast number of transitions need to be covered to ensure that clean lines are observed over the most probable energy ranges.

The approach that we have adopted to help solve this problem is to conduct deep, confusion-limited, spectral line surveys that cover a large frequency range (65–270 GHz). This large dataset will provide the means to locate clean transitions of new molecules over a large energy range. The overall method can be summarized as follows: (i) obtain an accurate set of rest frequencies, preferably by direct laboratory measurement, (ii) locate clean transitions that match in line shape and width, (iii) analyze the intensity

of those lines to ensure that they represent a physically connected system, (iv) locate a sufficient number of lines that sample an appropriate energy range, (v) identify and model intensities of all known species, particularly weak lines that may overlap to make strong lines, (vi) attempt to separate out contaminating features to strengthen the case, and (vii) account for beam-dilution where appropriate (Snyder *et al.* 2005; Halfen *et al.* 2005).

4. Case Studies of Glycolaldhyde and Dihydroxyacetone

Two case studies illustrate the power of this technique for identification of new interstellar molecules. These studies involve the molecules glycolaldehyde and dihydroxyacetone. Glycolaldehyde was first reported by Hollis *et al.* (2000) to be present in Sgr B2(N) using the Kitt Peak 12-meter telescope at millimeter wavelengths. These observations were followed by measurements at the NRAO GBT telescope in the K-band region (Hollis *et al.* 2004). Dihydroxyacetone was first reported by Widicus Weaver & Blake (2005) to be present in Sgr B2(N) using the CSO 10-meter telescope.

4.1. *Confirmation of Glycolaldehyde*

In Hollis *et al.* (2000), several lines were reported in the 3 mm window; however, most of these are apparently contaminated with other lines because they exhibit unusually broad line widths of 20 to 30 km s^{-1}. Recently, Halfen *et al.* (2005) have searched for an additional 43 transitions of this molecule. In this work, we discovered that only about 20% of the transitions were free from contamination, but we were able to observed 7 clean features that agree in line width and line center (V_{LSR}), and their intensities are consistent for a gas with a rotational temperature of 34 K.

Figure 4 is an example of some of the 40 spectra that we have measured for glycolaldehyde. These spectra represent the complete dataset. The degree of contamination ranges from total obscuration (102.5 GHz) to partial blending (103.4 GHz). The spectrum at 103.7 GHz shows an example of a clean feature, which exhibits the canonical line width and line center of glycolaldehyde in Sgr B2(N). Still, some ambiguities remain. For example, in the spectrum at 95.1 GHz, vibrationally excited lines of vinyl cyanide are seen throughout the spectrum (Nummelin & Bergman 1999; Halfen *et al.* 2005). The relative intensities of these lines closely match the observed data. However, there does seem to be some excess emission near the center frequency, which can be attributed to glycolaldehyde although it is difficult to quantify how much. Contamination such as this is seen throughout the dataset and highlights the importance of a complete accounting of all the emission from known constituents.

To help prove that the lines observed for glycolaldehyde are all physically connected, a rotational diagram was used (see Figure 5). As shown in the figure, both the clean lines and the partially blended lines all fit quite closely to the same physical model.

4.2. *A Rigorous Attempt to Confirm Interstellar Dihydroxyacetone*

Widicus Weaver & Blake (2005) reported the detection of 9 transitions of dihydroxyacetone (DHA) in Sgr B2(N) at 1.3 mm that closely matched rest frequencies measured by the same group (Widicus *et al.* 2004). From their observations, they derived a column density of 5×10^{15} cm^{-2} and a rotational temperature of 220 K. Figure 6 shows a single sideband spectrum of DHA taken at the ARO SMT 10-meter telescope that reproduces Figure 1 shown in Widicus Weaver & Blake (2005). The 220 K rotational temperature does not appear to reproduce the relative intensities of the observed lines as well as had been reported. The figure also shows that if the temperature is lowered, the predicted
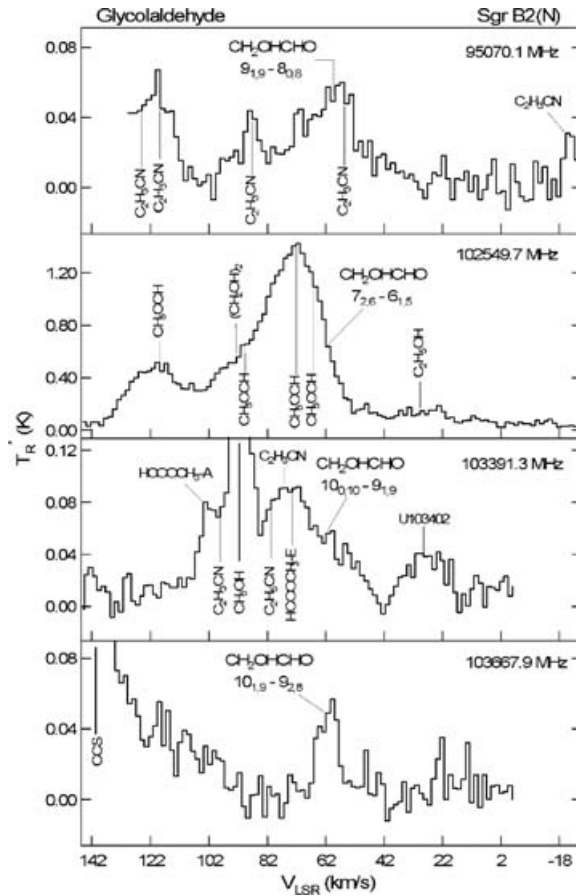
Figure 4. Single-sideband spectra of Sgr B2(N) taken with the ARO KP 12-meter telescope (Halfen *et al.* 2005).

relative intensities change dramatically. Also, the reported column density is about a factor of 6 higher than that of the common interstellar molecule H_2CO towards the same region (Apponi *et al.*, in preparation), clearly raising some questions about the validity of their detection. The intense lines observed at 1.3 mm (400 mK) and the high rotational temperature proposed indicate that DHA should possess a very intense spectrum from 100 GHz through 300 GHz. Figure 7 shows the simulated spectrum of DHA from 100 GHz to 245 GHz. The blue trace (see on-line presentation for color version) indicates the positions of the transitions, while the red trace is the modeled spectrum assuming 10 km s^{-1} line width and 10 mK rms noise. It is easy to see that there should be no problem detecting this molecule at lower frequencies.

In addition, a closer examination of the rotational diagram of the original 9 lines recorded at 1.3 mm further shows that their intensities are not reproduced well (see Figure 8). On the left is the rotational diagram reproduced from the results of Widicus Weaver & Blake (2005). At first glance, the data all seem reasonable, except that not all of the data points fall on the fitted line. The right panel of the figure shows the optical depth for each transition that would be required to bring the respective point onto the fitted line. These data show that the optical depths vary from –1 to 1.5 and suggest that these lines are not all physically connected.
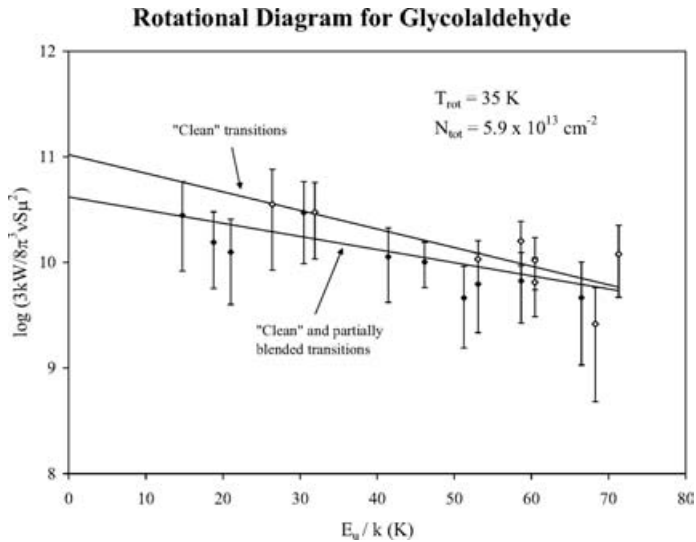
Rotational Diagram for Glycolaldehyde

Figure 5. Rotational energy diagram for clean and partially blended lines of glycolaldehyde in Sgr B2(N).
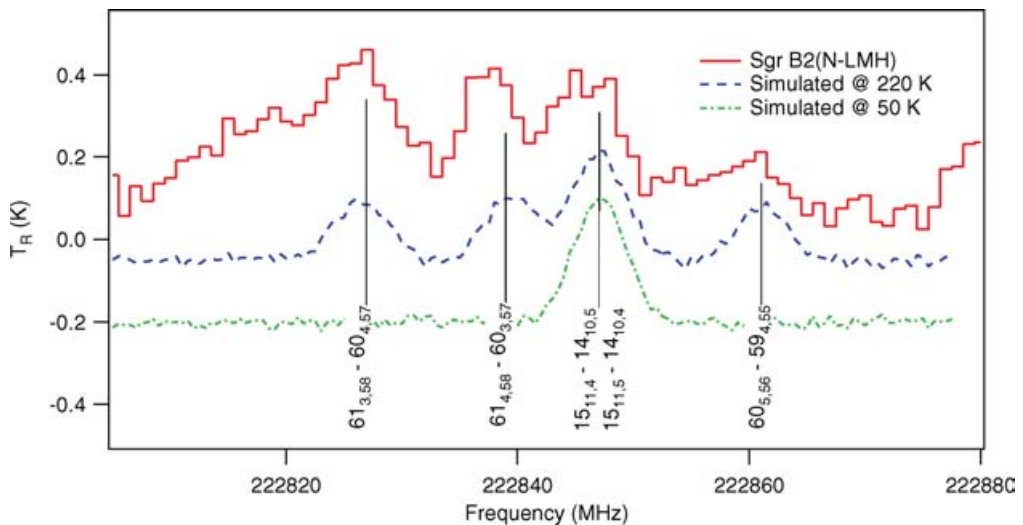

Figure 6. Single sideband spectrum taken at the ARO SMT 10-meter telescope covering the same spectral features as shown in Figure 1 of Widicus Weaver & Blake (2005). Overlaid on the spectrum are simulated spectra at 220 K and 50 K showing that these data are not well reproduced at either temperature even though they do show a close match in frequency.

The Ziurys group, in collaboration with Hollis and coworkers, also conducted searches for DHA in the same source in the K-band at the NRAO GBT and at 2 and 3 mm at the ARO 12-meter telescope at nearly the same time as the CSO measurements. Of 40 separate frequencies covering 61 favorable rotational transitions of DHA, only 2 of them show reasonable intensity correlation with the original work reported by Widicus Weaver & Blake (2005). Fourteen of the regions were contaminated with other molecules and 20
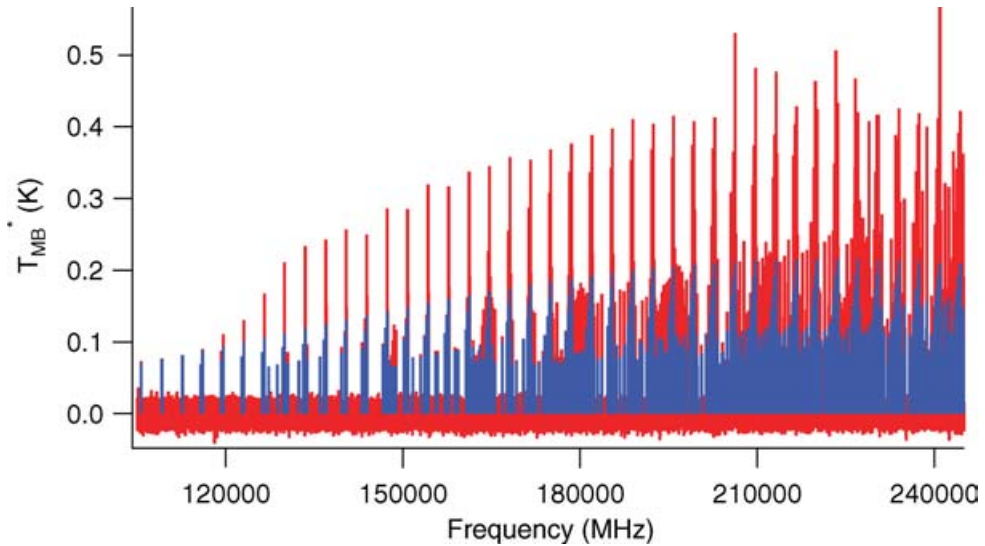
Figure 7. Simulated spectrum of DHA in Sgr B2(N) using the physical parameters listed in Widicus Weaver & Blake (2005). For a color version of this figure see the on-line presentation.
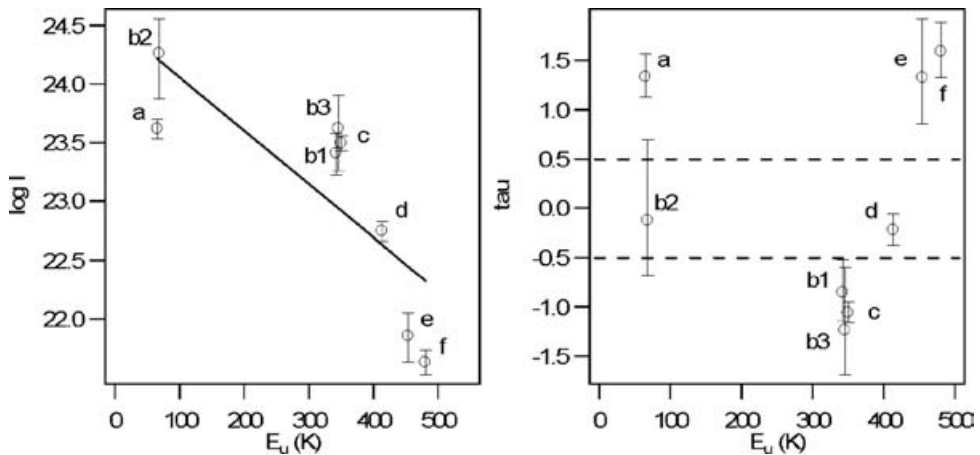

Figure 8. Rotational diagram analysis of the 9 lines taken from Widicus Weaver & Blake (2005) (*left*) and the respective opacities that would be required to correct each data point back onto the fitted line (*right*).

of them showed no significant emission. Figure 9 shows a few of these spectra. With these new results, we have constrained the temperature of DHA to below 50 K and have reduced the upper limit of the column density to 5×10^{13} cm^{-2}. It is doubtful from these new data that DHA is abundant enough in Sgr B2(N) for a definitive detection.

5. Conclusion

To obtain reliable spectroscopic evidence for the presence of large molecules, a careful and consistent dataset needs to be assembled. Although such data collection is time consuming, it is clearly necessary for a reliable molecular identification. Previous
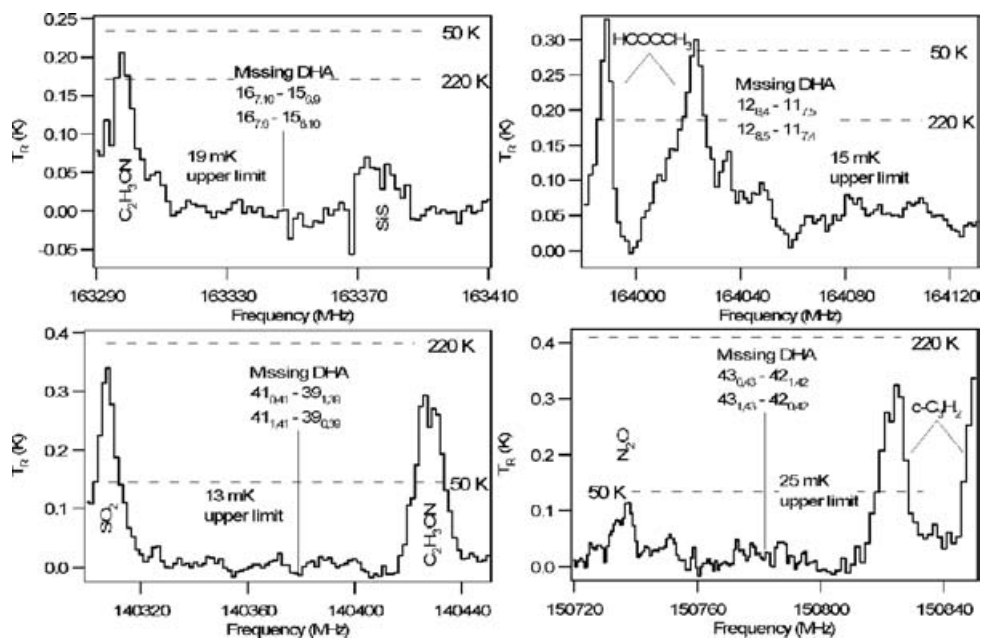
Figure 9. Spectra covering 4 favorable regions that should have significant DHA emission. The spectra were taken with the ARO 12-meter telescope in single sideband mode (Apponi *et al.* in preparation). The dashed lines indicate the predicted intensity of DHA emission based on T_{rot} = 220 K and T_{rot} = 50 K as labeled.

identifications of large species based on a limited number of transitions need to be closely re-examined. An accurate picture of interstellar organic chemistry will never be obtained otherwise.

Acknowledgements

References

Apponi, A. J., Halfen, D. T., Ziurys, L. M., Hollis, J. M., Lovas, F. J, & Remijan, A., in preparation.
Benner, S. A., Ricardo, A., & Carriga, M. A., 2004, *Curr. Opin. Chem. Biol.* 8, 672
Eschenmoser, A., 1999, *Science* 284, 2118
Friedel, D. N., Snyder, L. E., Turner, B. E., & Remijan, A., 2004, *Ap. J.* 600, 234
Halfen, D. T., Apponi, A. J., Woolf, N., Polt, R., & Ziurys, L. M. 2005, *Ap. J.*, accepted
Hohsaka, T. & Masahiko, S. M., 2002, *Curr. Opin. Chem. Biol.* 6, 809
Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, *Ap. J.* 540, L107
Hollis, J. M., Jewell, P. R., Lovas, F. J., & Remijan, A. 2004, *Ap. J.* 613, L45
Nummelin, A. & Bergman, P., 1999, *A&A* 341, L59
Nummelin, A. et al., 2000, *Ap. J. Suppl* 128, 213
Snyder, L. E. et al. 2005, *Ap. J.* 619, 914
Widicus, S. L., Braakman, R., Kent, D. R., & Blake, G. A., 2004, *J. Mol. Spectrosc.* 224, 101
Widicus Weaver, S. L., & Blake, G. A., 2005, *Ap. J.* 624, 33

Photo: E. Herbst