The Molecular Universe
Proceedings IAU Symposium No. 280, 2011
José Cernicharo & Rafael Bachiller, eds.

doi:10.1017/S1743921311025178

How Can We Use Complete Experimental Catalogs in the Complex Spectra Limit?

Frank C. De Lucia[1], **Sarah M. Fortman**[1], **Ivan R. Medvedev**[2], and **Christopher F. Neese**[1]

[1]Department of Physics, Ohio State University
[2]Department of Physics, Wright State University

Abstract. There is a broad consensus that many, if not most, of the unidentified spectral lines in astrophysical spectra are due to transitions in excited vibrational states of a relatively small number of molecules, the astrophysical weeds. For these unidentified lines, it is somewhat less well understood that the spectroscopic effort required to characterize them in the traditional quantum mechanical catalog approach is substantially larger because of significant perturbations. We have previously discussed a new experimental approach that addresses this challenge. This approach is based on the analysis of many complete, intensity-calibrated spectra taken over a range of temperatures. However, the spectroscopic completeness of this approach results in a much larger database. These data can be transfer to the astrophysical community in a variety of ways, but because an order of magnitude larger number of lines is included, consideration must be given to implementation strategies.

1. Introduction

We have previously proposed and demonstrated a method to record complete, intensity calibrated spectra as a function of temperature and to use these results to provide complete astrophysical catalogs or complete point-by-point simulations as a function of temperature (Medvedev & De Lucia (2007), Fortman *et al.* (2010a, 2010b, 2010c, 2011). Table 1 shows an example of a small portion of a resultant catalog and Figure 1 shows an example of a small region in the point-by-point simulation.

Table 1. A small portion of a catalog derived from an experimental catalog.

Frequency (MHz)	$S_{ij}\mu^2$ (D^2)	Energy cm^{-1}	Figures of Merit
227820.916	197	1056	W, G, T
227821.777	600	1196	W
227826.268	335	1106	W
227827.710	315	1093	W, T
227834.675	630	1257	W, G, T
227838.149	9	395	W, G, T
227850.297	703	1244	W
227857.065	429	1146	W
227858.221	310	637	
227866.026	532	825	
227870.968	304	654	
227872.144	415	1145	W
227892.603	485	1229	W, G, T
227897.612	635	157	
227906.700	694	150	

Figure 1. A point-by-point simulation for a small spectral region. The green is a single experimental trace at 300 K. The black is a simulation of this spectrum based on an analysis of ∼400 different spectral scans. The red sticks are lines included in the QM catalog.

In Figure 1 the black line, which represents the spectrum calculated from a fit to all ∼400 temperature scans, is of much higher sensitivity than the single green experimental scan because of statistical redundancy and contains many lines which are unobservable on the scale of the figure. The red sticks represent the predictions of the quantum mechanical (QM) model. The figure illustrates that many relatively weak lines (e. g. the line near 227.84 GHz) are in the catalogs, whereas much stronger lines (e. g. the line near 227.87 GHz) are not. Additionally, this spectrum shows that the weak line predicted near 227.86 GHz is overlapped by a much stronger line not included in the QM catalog.

One of the challenges for the experimental catalog approach is blended lines because their shapes vary, often significantly, with temperature. To meet this challenge we developed and tested what we refer to as a point-by-point analysis that uses a method similar to (but augmented by knowledge of the Gaussian profile associated with laboratory Doppler limited spectra) our line-by-line catalog approach, but which treats each ($\sim 10^6$) frequency point independently. Figure 2 illustrates the application of this approach to blends.

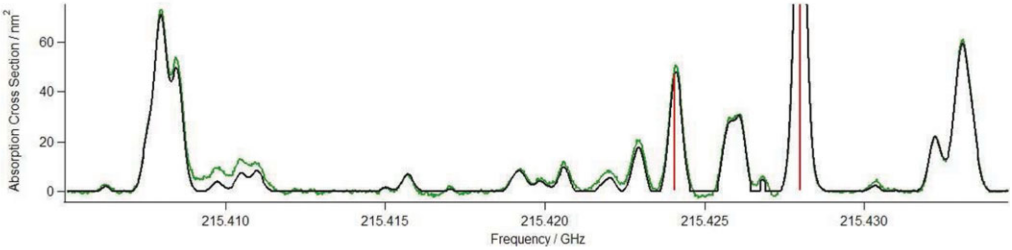

Figure 2. A blow up of a small portion of blended spectra. The black is a simulation of this spectrum based on an analysis of ∼400 different spectral scans. The red sticks are lines included in the QM catalog.

In this figure the green trace is a single experimental scan near 300 K, and the black trace is a simulation from the point-by-point analysis. Inspection of the feature near 215.410 GHz shows that the simulation faithfully reproduces the complex lineshape. Additionally, because of the statistical redundancy of the fit to ∼400 spectral scans, the noise and baseline issues are both significantly reduced.

In this paper we will consider the use of these complete experimental catalogs in an astrophysical context. Figure 3 shows on a very compressed scale both the astrophysical data and five laboratory spectra.

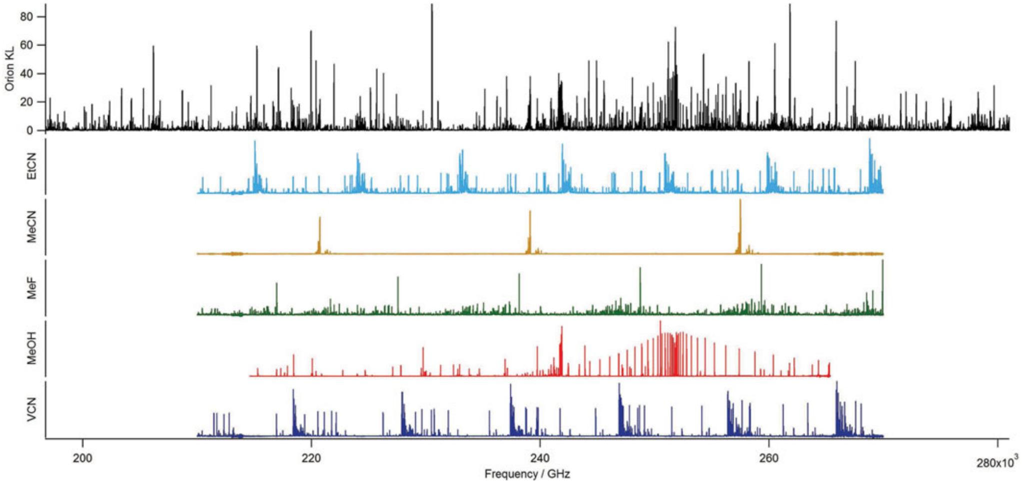

Figure 3. Astrophysical data (upper panel) and intensity calibrated laboratory data (lower five panels, top to bottom: ethyl cyanide, methyl cyanide, methyl formate, methanol, and vinyl cyanide) on highly compressed scales. Astrophysical data from IRAM, courtesy of Jose Cernicharo and Belen Tercero.

2. Quantum mechanical catalogs and catalogs/spectra derived from experimental spectra

Quantum mechanical catalogs and catalogs/spectra derived from experimental data provide complimentary information. Table 2 summarizes this.

Table 2. Comparison of QM catalogs and intensity calibrated experimental catalogs.

	Quantum Mechanical Catalogs	**Experimental Spectra**
Line Frequencies	Model redundancy Model extrapolation error	Directly measured
Lower State energy	Very accurate from model	Accurate enough to give ∼1% intensity error
Completness	Only for analyzed vibrational states and for rotational states within cutoffs	Currently down to ∼1/1000 of strongest line for species
Quantum numbers	Known	Unknown

2.1. *Line frequencies*

In the QM catalogs the frequencies are calculated via a Hamiltonian model from a large number of assigned lines, whereas in the experimental catalogs each measured frequency stands on its own. The QM frequencies have the advantage that the statistical redundancies of the fit can provide increased accuracy. A comparison for vinyl cyanide in the 210-270 GHz band is shown in Fig. 4, which compares the experimental and QM catalog frequencies. This figure shows that for lines for which there are not QM model errors, the agreement is on the order of 5 kHz. This is better than we expected for the experimental frequencies and results from the fact that the varying temperature randomizes baseline effects that can be traced to standing waves. These standing wave effects can be one of the limitations on frequency measurement accuracy. The larger systematic differences are due to model errors in the QM analyses.

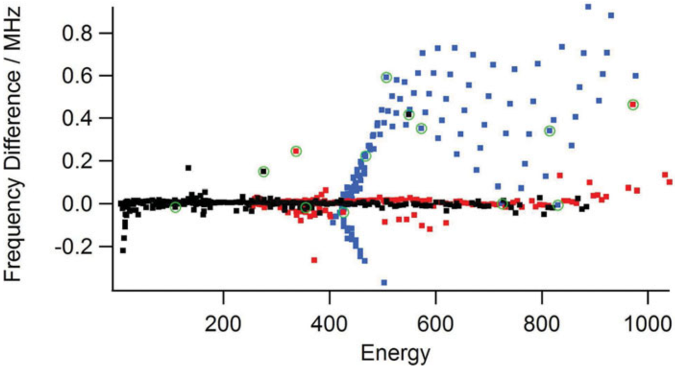

Figure 4. A comparison between QM catalog frequencies and experimentally measured frequencies for vinyl cyanide. The black points are ground vibrational state lines, the red points are first excited state lines, and the blue points are lines in the second excited vibrational state. The systematic differences are presumably due to model errors caused by perturbations in the QM analysis.

The shifts in higher frequency bands are considerably larger and it is not possible to make a figure similar to that of Fig. 4 because it can be hard to match experimental lines with QM model predictions. Figure 5 shows an example of shifts in a small frequency region. In Fig. 5 the calibrated intensities greatly aid the spectral assignments shown.

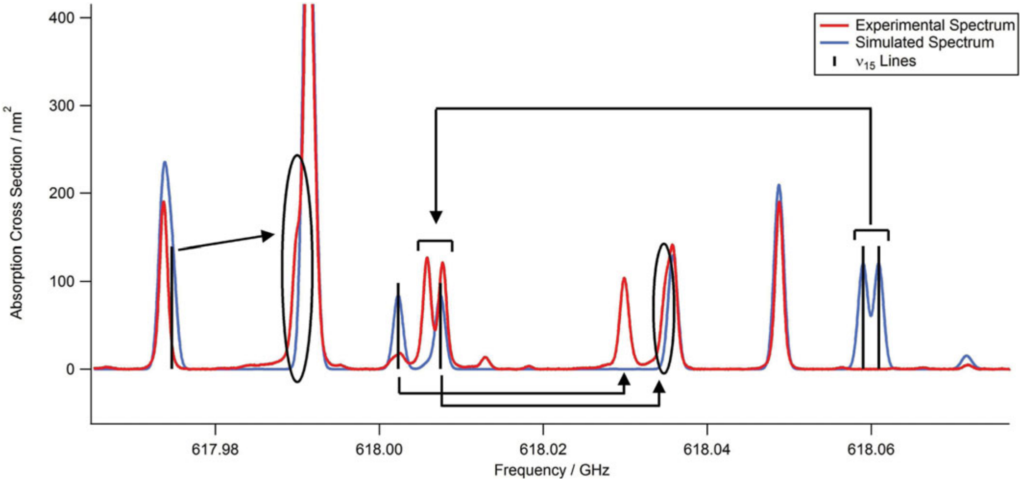

Figure 5. Frequency shifts in the calculated spectrum of vinyl cyanide.

2.2. *Lower state energies/calculated intensities*

QM analyses provide extremely accurate lower state energies, very small fractions of a wave number, whereas the uncertainties associated with the lower state energies of the experimental fits are much greater, ∼10 cm^{-1} for unblended lines. For blended lines, the calculated lower state energies have no meaning. In fact, the calculated lower state energies for blended lines are best considered as fitting parameters for calculation of the spectral line intensities as a function of temperature. The accuracy of these calculated intensities is typically 1%, with greater fractional uncertainties for weaker lines.

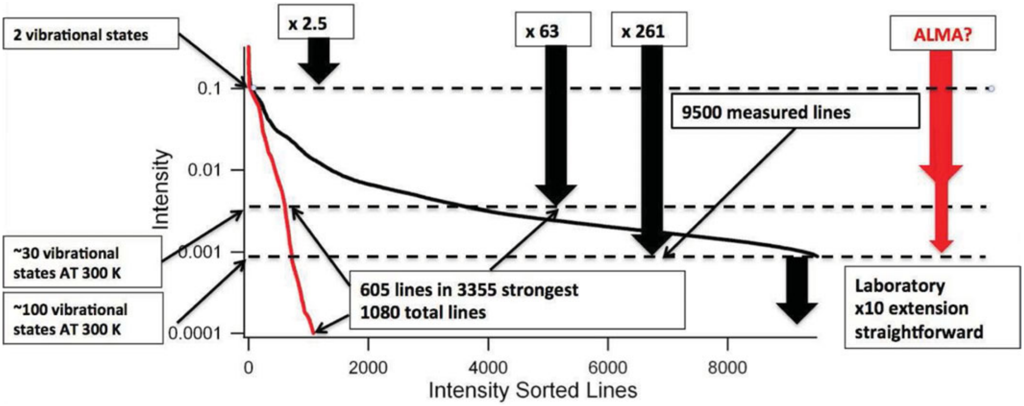

Figure 6. Completeness of the ethyl cyanide QM catalog. The red points (which are so numerous as to appear as lines) are the QM catalog intensities, sorted by strength and the black points are the experimental intensities sorted by strength. These diverge at the strength of the first experimental line that is not included in the catalog.

2.3. *Completeness*

The most important feature of the experimental method is that it is complete. We have discussed this issue in quantitative detail (Fortman *et al.* 2010c). Figure 6 shows this for ethyl cyanide and Figure 7 shows an overview for eight astrophysical weeds.

QM catalogs are complete in the sense that for every vibrational state that has been analyzed as many rotational transitions as desired can be calculated, but with varying accuracy according to model extrapolation and the significance of perturbations. The experimental catalogs only include those lines that are strong enough to observe and are thus intensity incomplete. However, this is not a practical limitation because the experimental simulations include lines that are weaker by $\sim 10^3$–10^4 of the strongest lines in their spectrum. In contrast, the QM catalogs are incomplete because they only include lines for which there has been an analysis. It is this incompleteness that is believed to be the source of many, if not most, of the unidentified astrophysical lines.

2.4. *Quantum numbers*

The experimental approach provides a means for the calculation of spectra as a function of temperature. While its completeness, intensity calibration, and known lower state energies can greatly aid a QM assignment (Fortman *et al.* 2010d), it does not provide assigned quantum numbers. There are many astrophysical circumstances for which the knowledge of the quantum numbers would be advantageous. These include vibrational temperatures that are not in equilibrium with rotational temperatures and state populations that cannot be described by a temperature.

3. Astrophysical Impact

In this section we will use a comparison between our experimental spectra and astrophysical spectra to show their impact as well as challenges in making use of them. We are grateful to Jose Cernicharo and Belen Tercero for making prepublication IRAM data for Orion KL available to us for these preliminary comparisons.

Because of the large amount of laboratory experimental data involved, it was our goal to do these comparisons and analyses in as automated a fashion as possible. This led us to fitting procedures similar to those that we have developed for laboratory analytical

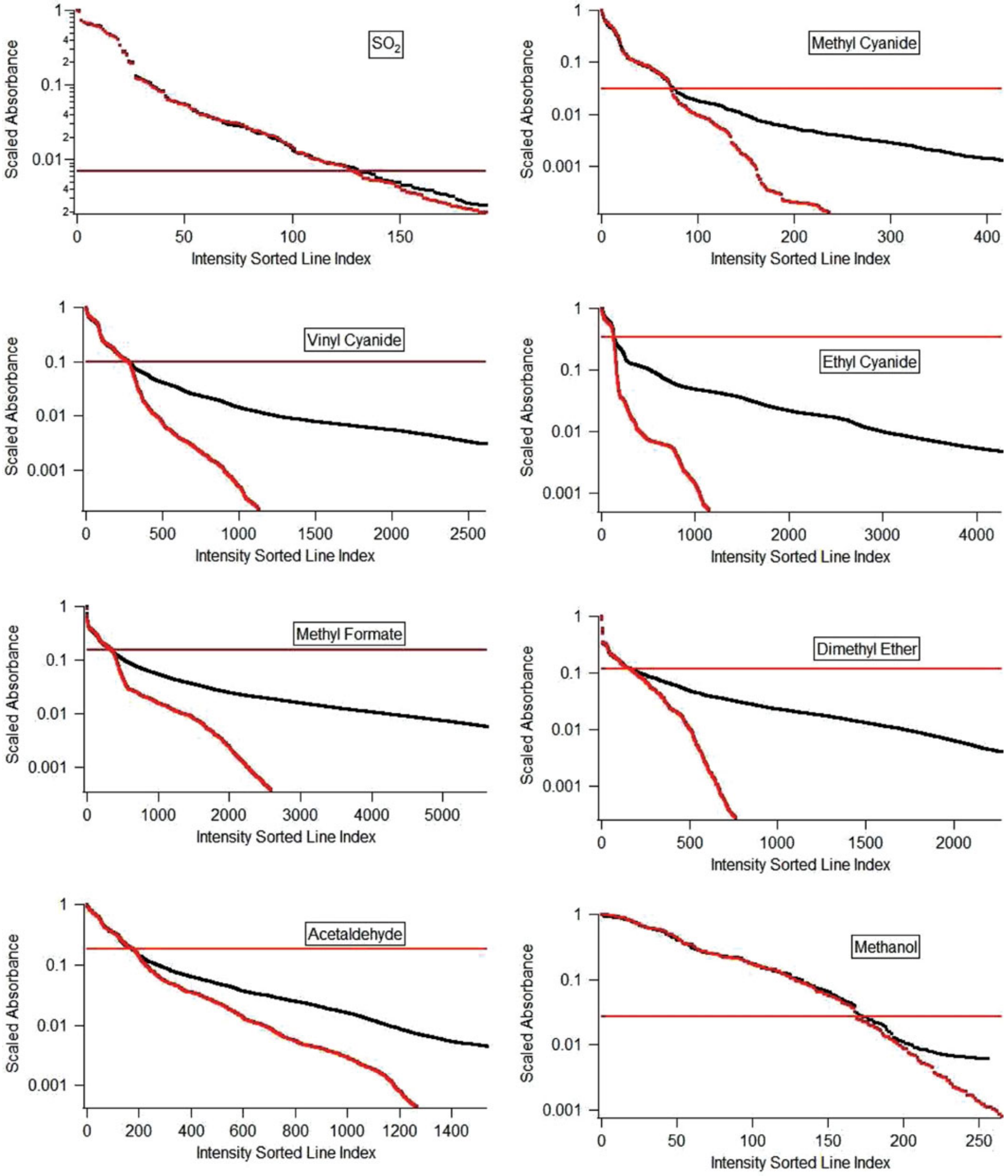

Figure 7. Overview of the completeness for eight astrophysical weeds. The red points (which are so numerous as to appear as lines) are the QM catalog intensities, sorted by strength and the black points are the experimental intensities sorted by strength. These diverge at the strength of the first experimental line that is not included in the catalog. The expected location of this divergence is determined by the Boltzmann factor of the lowest excited vibrational state not included in the analysis. In each figure this is indicated by the horizontal red line.

chemistry, but with the recognition that the interstellar medium is a much more complex environment.

To begin, we needed to convolve our laboratory data with appropriate astrophysical lineshapes, one for each species. Indeed, a fundamental question is whether or not this is possible. To do this, we chose a number of assigned lines for each of the species and

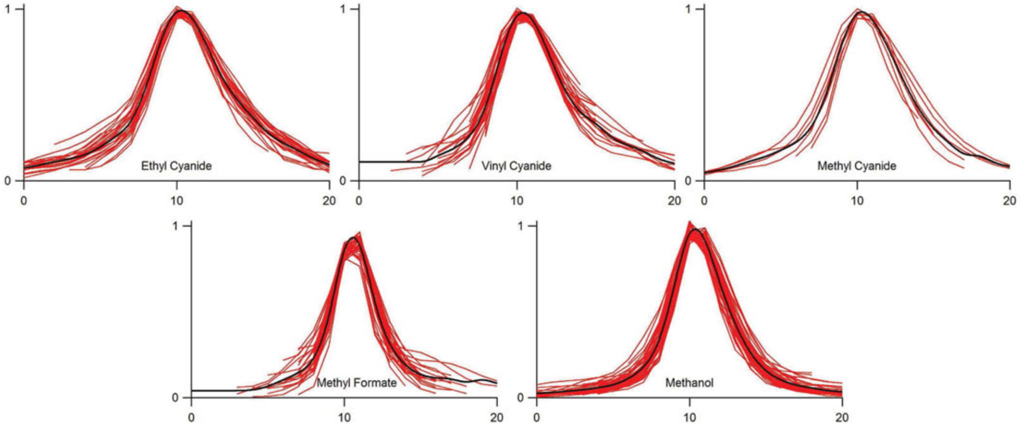

Figure 8. Lineshapes for five of the observed species in the astrophysical data. The red lines are from fits to relatively isolated astrophysical lines. The black lines are the averages of the astrophysical lines that are used for convolution with the astrophysical data.

overlaid the lineshapes derived from the astrophysical data as shown in Fig. 8. This figure shows that to within the bin resolution of the astrophysical data and some overlap on the wings of the spectra, that it was possible to obtain a single lineshape for each species that was representative of the astrophysical data. These lineshapes are considerably broader than those observed by interferometers, and it is inferred that beam averaging plays a significant role.

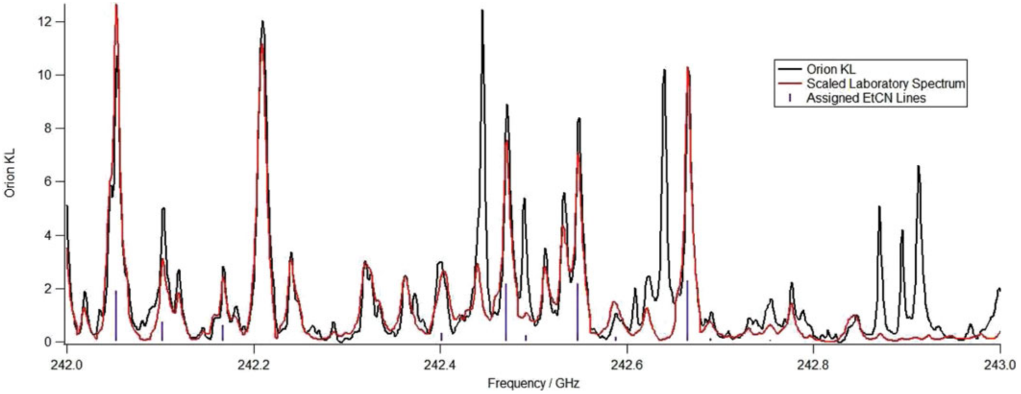

Figure 9. Fit of an experimental simulation of 190 K ethyl cyanide to the Orion KL spectrum.

Above in Figs. 6 and 7, we showed catalog completeness in statistical terms for eight astrophysical weeds. Figure 9 shows this in spectral detail. This figure shows the results of a fit of an experimental simulation of a 190 K spectrum of ethyl cyanide to the astrophysical data, with the abundance as the only free variable. This fit was performed iteratively with any residual that resulted from the astrophysical data being stronger removed from the fit under the assumption that the extra intensity is from another species. There are many similar examples in the astrophysical data.

These results are very encouraging. In this spectral region there are 10 features that can be ascribed to QM cataloged line, but more than 15 additional lines are predicted

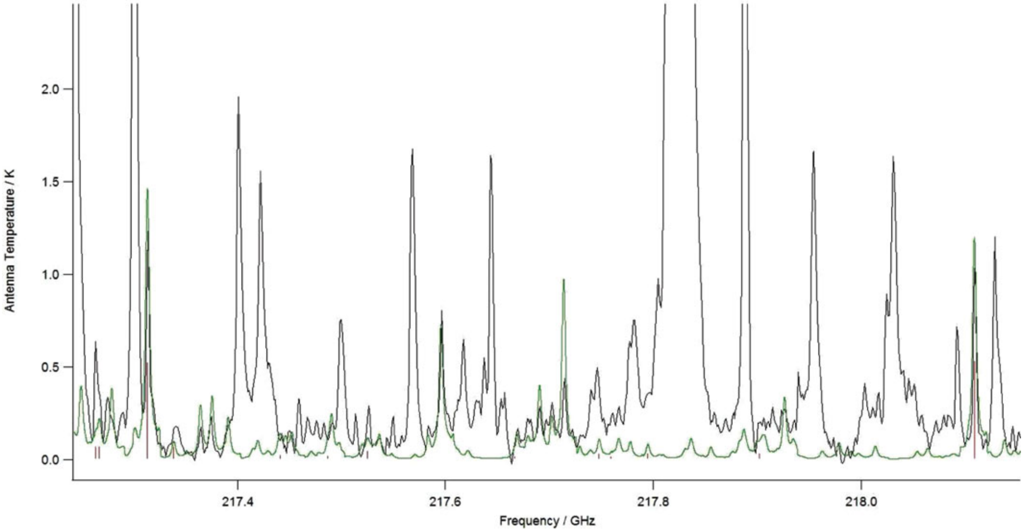

Figure 10. A comparison of the astrophysical spectrum of Orion KL (black trace) and the 190 K spectrum of methyl formate (green trace). The predictions of the QM catalog are shown as red sticks.

by the experimental spectrum. For a number of these additional lines, the experimental spectrum provides additional detail for the structure of the blends that make up these features.

Figure 10 shows a similar comparison for methyl formate in a relatively crowed region of the astrophysical spectrum. Again, there are many additional methyl formate lines that can be assigned from the methyl formate spectrum, many due to lines that are considerably stronger than other lines that are include in the QM catalog. However, in this preliminary comparison the intensity matches in some cases are not as good as those shown in Fig. 9.

4. Observations and conclusions

From these comparisons we make several preliminary observations:

(*a*) The experimental simulation accounts for many features that are not in the QM catalog, including intensities and details of blends. Thus, these astrophysical features are not available for assignment to “new” species.

(*b*) The experimental simulation gets the intensity right in regions where the QM calculation is a clear underestimate-presumably due to a missing line in the catalog. See the example near 242.4 GHz in Fig. 9.

(*c*) The experimental simulation does not predict any lines that are not in the astrophysical spectrum.

(*d*) These one temperature fits vary in the accuracy with which they reproduce intensities. It will be interesting to see if multiple temperature fits or interferometer data from arrays such as ALMA that reduce beam averaging improve the intensity fit.

(*e*) The experimental fits are based on terrestrial isotopic abundances. Either experimental fits on samples with different abundances or the use of ground state QM analyses will be needed to accurately account for astronomical isotopic abundances.

References

Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C., 2010a, *ApJ*, 714, 476

Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C., 2010b, *ApJ*, 725, 1682

Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C., 2010c, *ApJL*, 725, L11

Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C., 2010d, *Chem. Phys. Lett.*, 493, 212

Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C., 2011, *ApJL*, in press

Medvedev, I. R. & De Lucia, F. C., 2007, *ApJ*, 656, 621

Discussion

R. Mauersberger: In star-forming regions, visible mm lines can contribute 30-50% to the total flux (line + continuum). Is there any guess of the contribution of blended weed lines to the “continuum”?

F. De Lucia: We do not have an estimate at this time, but since the experimental spectra are complete down to a relatively low level, we should be able to make such an estimate. It would be interesting to compare this estimate with an estimate that one would get from a calculated vibrational partition function.

K. Menten: 1) What do you mean when you sy you toolo 400 spectra? 2) How many species did you do? did you do methanol?

F. De Lucia: 1) We scan 50 to 100 GHz in about 30 seconds (120 scans/hr). Each scan is a complete spectrum. While we scan for about 4 hours, we ramp the temperature by about 200 K. Thus, the temperature varies by about 0.5 K over each scan. 2) In the 210-270 GHz region we have done 8 molecules, including methanol. In the 575-650 GHz region, we have done 2. We expect do the remaining 6 in the next few months.