ROTATIONAL SPECTROSCOPY AND PRECISE EQUILIBRIUM STRUCTURE DETERMINATION OF NITROGENOUS MOLECULES

by

Zachary N. Heim

A senior honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

(Chemistry)

at the

University of Wisconsin – Madison

Spring 2017

Advisors:

Professor Robert J. McMahon

Professor R. Claude Woods

TABLE OF CONTENTS

ABSTRACT

Rotational spectroscopy is a powerful spectroscopic method for both the identification of molecules in the gas-phase and for determination of molecular structure. This is due to the inherent relationship between molecular rotation and the distribution of mass in any given molecule. For molecules with permanent dipole moments, this relationship can be exploited to confidently identify various molecules in the interstellar medium (ISM) as well as determine some of the most precise experimental molecular structures. For identification of these molecules in the ISM, detailed laboratory studies of potential targets are required. Towards this end, the experimental rotational constants of pyrimidine have been determined in the millimeter-wave range. These constants were determined from the least-squares fitting of several thousand rotational transitions to a centrifugal distortion Hamiltonian. From these constants, extremely precise moments of inertia can be determined for the main isotopologue as well as various heavy atom- and deuterium-substituted isotopologues which can then be used to determine the structure of pyrimidine. Additionally, guanidine was identified as a similar target for study by rotational spectroscopy. Work is currently underway to isolate a pure sample of guanidine for study by rotational spectroscopy and structure determination.

CHAPTER 1

INTRODUCTION

Spectroscopic Methods

Rotational spectroscopy is uniquely suited for the determination of molecular structure as well as the unambiguous identification of molecules because the spectrum obtained from any given molecule depends almost solely on the distribution of mass within the molecule. The principle rotational constants obtained are generally determined to eight or nine significant figures. The precise determination of these constants, and by extension the moments of inertia, permits the most precise experimental determination of molecular structure to date. These structures are often an order of magnitude better than the best crystal structures produced for the same molecule. Hence, the structures determined using rotational spectroscopy are often the best option to use to benchmark theoretical, computationally optimized structures. Additionally, the relationship between an individual molecule's spectrum and the molecular structure produces a spectral "fingerprint" that can be used towards unambiguous identification. Combined with the fact that the atmosphere has a convenient spectral window with few absorptions to inhibit ground-based radio telescope observations, just a few rotational transitions can lead to identification of a molecule in the ISM. These identifications do, however, rely on the precisely determined rotational constants resulting from fairly extensive laboratory studies to support these ground-based surveys.

Pyrimidine

Pyrimidine has, for some time, been a target of astrochemical interest due to its potential to serve as polar trace for other aromatic species. IR spectroscopy has been used to detect benzene in regions such as the protoplanetary nebula CRL 618.[1] In addition, it could be used as a trace for the presence of polycyclic aromatic hydrocarbons (PANH's). These are thought to be responsible for unidentified infrared bands (UIB's) and/or diffuse interstellar bands.[2-4] Observation of a nitrogen-containing heterocycle would be an important step in understanding the chemistry of aromatic molecules in the ISM. Pyrimidine is an ideal candidate for detection in the ISM by rotational spectroscopy for two reasons: it has an appreciable dipole moment (μ = 2.334 D)[5] and it is the most thermodynamically stable of the *c*-$C_4H_4N_2$ family. By B3LYP/aug-cc-pVTZ calculations, pyrimidine is 4 kcal/mol more stable than pyrazine and 22 kcal/mol than pyridazine. Even though pyridazine has a significantly larger dipole moment (μ = 4.22 D)[6], it is possible that the relative stability of pyrimidine will result in higher abundance in the ISM and a higher likelihood of detection.

Previous studies of pyrimidine[5, 7-9] investigated the rotational spectra of the main isotopologue as well as several heavy-atom and deuterium substituted isotopologues. The previous work included a substitution structure including only the position of the heavy atoms,[9] providing an ideal point to expand and improve upon. Recent work determining the structure of pyridazine[6] gave a simple synthetic route towards obtaining a multitude of deuterated isotopologues, giving even more motivation towards determining the total molecular structure of pyrimidine by rotational spectroscopy.

Guanidine

Guanidine was the ideal follow-up to the study of pyrimidine due to the highly-nitrogenous nature and lack of microwave studies. Recent work on the structure of guanidine has largely been theoretical. Notable exceptions to this include the recent x-ray[10] and neutron diffraction[11] experiments yielding the solid-state structure of guanidine. Using rotational spectroscopy, these structures could be immensely improved upon by determining the position of the hydrogens, in particular. Guanidine is similar to urea in that it has conjugated amines with significant pyramidal character, presenting an interesting problem that rotational spectroscopy is uniquely suited to solve. The relative position of the hydrogens is highly susceptible to the effects of hydrogen-bonding, such as that expected to be present in a crystal of guanidine. Using rotational spectroscopy, these interactions can be removed by studying a gas-phase sample at low pressure. For the determination of the true equilibrium structure of guanidine, rotational spectroscopy is arguably the best method that can be employed.

CHAPTER 2

ROTATIONAL SPECTROSCOPY OF PYRIMIDINE

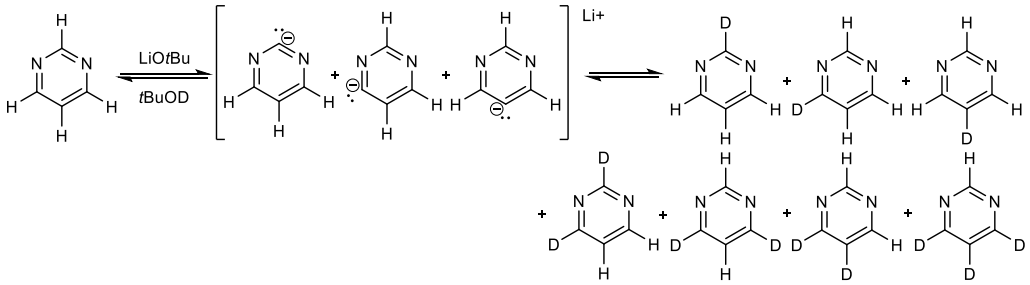
Experimental Setup and Data Analysis

A millimeter-wave spectrometer described previously was used to collect broadband spectra from 235 GHz to 360 GHz.[6, 12] All line measurements reported are estimated to have an uncertainty of ± 50 kHz. A sample of pyrimidine, purchased from Sigma-Aldrich, was used without further purification (≥ 98.5% purity) and spectra were collected at a pressure of approximately 30 mTorr using either a static sample or a flow system and at room temperature. More detailed scans of smaller regions, typically around the bandhead of the ^{15}N- and ^{13}C-isotopologues studied in natural abundance, were obtained at 18 mTorr. The data were analyzed using the Assignment and Analysis of Broadband Spectra (AABS) program suite.[13-14] All transitions observed were least-squares fit to either a quartic or sextic centrifugal distortion/rigid rotor Hamiltonian. The A_I^r, S_I^r, and S_III^r reductions were all used to be certain all transitions observed fit well regardless of reduction. Rotational constants and other data obtained from the spectra can be found in Appendix A.

Synthesis of Deuterated Isotopologues

Deuterium enriched samples of pyrimidine were synthesized using the method previously employed to generate deuterated pyridazines using *t*-BuOD/*t*-BuOLi, as shown in Scheme 1.

Pyrimidine samples with varying degrees of deuterium incorporation were obtained using reaction times of 1 hour, approximately 2.5 days, and approximately 25 days.



Scheme 1. The synthetic method for deuteration of pyrimidine.[6]

In comparison to pyridazine[6], deuteration of pyrimidine proved to be more difficult, requiring greater quantities of *t*-BuOD and longer reaction times. This method is advantageous because it generates a single sample containing varying amounts of mono-, di-, tri-, and tetradeuterio species, allowing for study of all 11 deuterated species in a single broadband spectrum. These spectra were used to identify and characterize the 11 deuterated isotopologues, of which only 2,4,6-trideuteriopyrimidine had been studied previously.[7]

Ground State Isotopologues

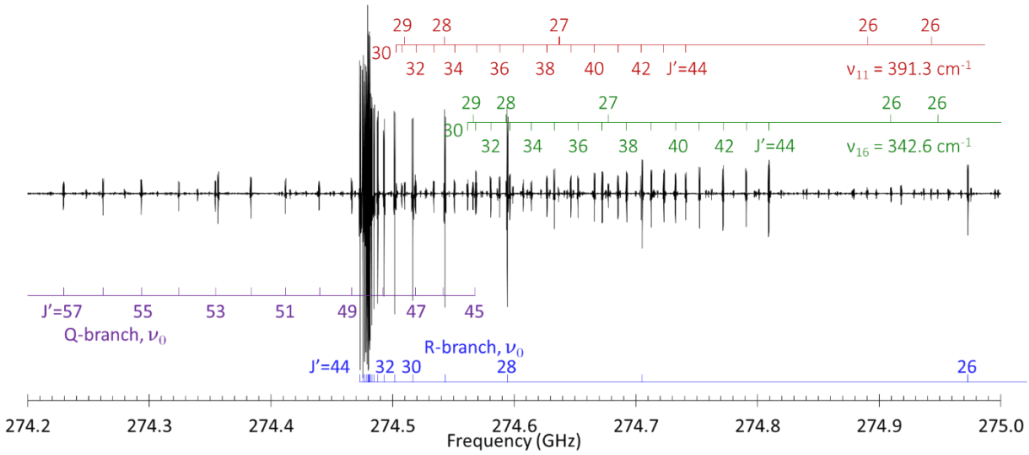


Figure 1. A representative bandhead structure of the main isotopologue of pyrimidine.

The rotational spectrum of the main isotopologue and the ^{15}N- and ^{13}C-substituted isotopologues was relatively straightforward due to the existence of well determined rotational constants from the most recent previous study of pyrimidine by Kisiel *et al.*[9] As a nearly-oblate asymmetric rotor ($\kappa = 0.869$), a majority of the most intense transitions are grouped into compact band structures with small spacings between transitions as shown in Figure 1. For the main isotopologue, the spectrum consists of only *b*-type transitions of either *R*- transitions ($^{b}R_{1,1}$, $^{b}R_{-1,1}$, $^{b}R_{1,-1}$, $^{b}R_{-1,3}$, and $^{b}R_{3,-1}$) or *Q*-branch ($^{b}Q_{1,-1}$) transitions.

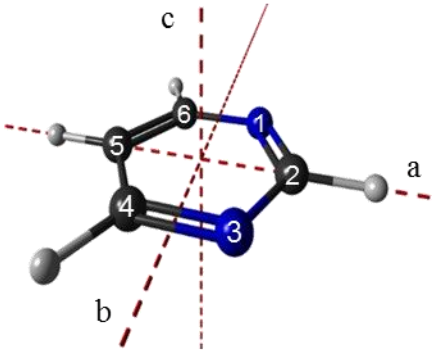


Figure 2. The structure of pyrimidine with principle inertial axes and atom numbering displayed.

The ^{13}C(2)- and ^{13}C(5)-substituted isotopologues and several deuterated isotopologues, with isotopic substitution symmetric about the same axis as the dipole moment, also exhibit the same types of transitions as the main isotopologue. The ^{15}N- and ^{13}C(4)-substituted isotopologues as well as several deuterated isotopologues, on the other hand, are not substituted symmetrically about the same axis as the dipole component. The change in center of mass and the location of the atoms relative to the principle inertial axes leads to these isotopologues exhibiting both *a*- and *b*-type transitions or just *a*-type transitions. In all cases, the transitions near the bandhead either occur as degenerate signals. As transitions of that band are observed at higher frequency, the degeneracy of the transitions breaks down and the signals become doublets or quartets, depending on whether the spectrum is made up of *b*-type or *a*-type transitions only or a combination of *a*- and

b-type transitions. For the ^{15}N- and ^{13}C(4)-substituted isotopologues, this quartet structure allowed for unambiguous identification of off-axis substitution. The measurable transitions for each isotopologue have been least-squares fit as mentioned before.

To precisely determine the rotational constants of any given isotopologue, a wide range of transitions were needed. For example, the least-squares fit of the transitions from the main isotopologue contained R-branch transitions ranging from $J' = 16$ to $J' = 54$ and Q-branch transitions ranging from $J' = 40$ to $J' = 100$. The final least-squares fit of the main isotopologue contained over 1700 individual transitions, including transitions from previous works.[8-9] This allowed for the determination of principle rotational constants with error in the eighth or ninth significant figures. Similarly, the quartic and sextic centrifugal distortion constants are all determined with error in the fifth and sixth decimal places.

As a consequence of the synthetic method utilized, all spectra collected for the deuterated isotopologues were quite congested with transitions from many different isotopologues and their excited vibrational states. This method was beneficial for the convenient production and detection of many isotopologues and was significantly less challenging than synthetic routes used previously. This also limited the number of usable transitions to R-type transitions and only the most intense Q-type transitions. The quantity of measurable lines for all but the most abundant deuterated isotopologues was approximately 200-400 transitions, most of which were degenerate sets of R-branch transitions. Due to the large effects of centrifugal distortion on these high-*J* transitions, it was necessary to fit approximately 200 transitions to accurately determine the rotational constants for a quartic centrifugal distortion Hamiltonian. Several hundred more transitions were required to obtain a satisfactory sextic fit. Therefore, most deuterated isotopologues were only fit to a full quartic Hamiltonian.

Vibrational Satellites

Nine vibrational satellites of pyrimidine were studied, in part, to assess the accuracy of the vibration-rotation interaction corrections calculated using both DFT and *ab initio* methods. Of these nine satellites, six are fundamentals and the remaining three are overtones or combinations. Transition frequencies for five of these satellites were initially predicted using rotational and distortion constants from the previous pure rotational work of Kisiel *et al*.[9] for the fundamentals ν_{16}, ν_{11}, and ν_{24} and rovibrational studies of Albert *et al.*[15] for the fundamentals ν_{24}, ν_{15}, and ν_{14}. Transitions of all of these states were well fit to single-state Hamiltonian models. Rotational and distortion constants were used to determine the experimental α_i values. The assignment of these satellites, which match previous assignments,[9 15] is unambiguous based upon the agreement of their CCSD(T)/ANO1 calculated and experimental α_i values and the relative transition intensities. The vibration-rotation interactions for the two lowest-energy, ring-deformation modes (ν_{16} and ν_{11}) are predicted very precisely by CCSD(T)/ANO1 with absolute errors < 0.08 MHz. The DFT equivalent α_i values, on the other hand, have errors that are as large as 12 MHz. Most importantly, at the DFT level, it is very difficult to distinguish between these two vibrational satellites, as they have qualitatively very similar predicted α_i values at that level. This same general trend holds for the other vibrational satellites investigated with the worst CCSD(T) prediction being that of α_A for ν_{15}, which is under-predicted by 3.2 MHz (23 %). The analysis of these vibration-rotation constants, as well as values for each α_i, is extensively discussed in the text provided in Appendix A.

CHAPTER 3

EQUILIBRIUM STRUCTURE DETERMINATION OF PYRIMIDINE

Methods

Determination of the molecular equilibrium structure was made possible due to the substantial number of isotopologues studied. To attain the highest degree of precision possible, several corrections to the rotational constant must be made. The first step taken is to remove the effects of using a reduced centrifugal distortion Hamiltonian to least-squares fit the transition energies to rotational constants. Converting from the A- and S-reduced constants to the determinable constants (A_0'', B_0'', C_0'') is achieved through the use of equations 1-6.[16]

$$A_0'' = A^{(A)} + 2\Delta_J \quad (1)$$

$$B_0'' = B^{(A)} + 2\Delta_J + \Delta_{JK} - 2\delta_J - 2\delta_K \quad (2)$$

$$C_0'' = C^{(A)} + 2\Delta_J + \Delta_{JK} + 2\delta_J + 2\delta_K \quad (3)$$

$$A_0'' = A^{(S)} + 2D_J + 6d_2 \quad (4)$$

$$B_0'' = B^{(S)} + 2D_J + D_{JK} + 2d_1 + 4d_2 \quad (5)$$

$$C_0'' = C^{(S)} + 2D_J + D_{JK} - 2d_1 + 4d_2 \quad (6)$$

Determinable constants from both reduction were averaged and used in the following analysis. Further corrections were needed to determine the equilibrium structure. The determinable constants, although corrected for the effects of centrifugal distortion, are still vibrationally

averaged. To correct for this, the effect of vibration-rotation interaction corrections ($\frac{1}{2}\Sigma\alpha_i$) must be removed. As the vibration-rotation interaction cannot be directly measured by rotational spectroscopy, the correction must be calculated by an anharmonic frequency calculation. Additional correction must be made for the effect of electron mass. For systems with conjugated π systems, like pyrimidine and pyridazine, have been shown to require this correction due to their large out-of-plane electron mass distribution.[6, 17-19] This is achieved through calculation of the electronic component of the inertial *g*-tensor, g^{bb}, where η is the electron-proton mass ratio and B_{calc} is the rotational constant determined through the calculation. Together, these contributions result in the equilibrium rotational constants as given in equation 7.

$$B_e = B_0 + \frac{1}{2}\sum \alpha_i - \eta g^{bb} B_{calc} \qquad (7)$$

These corrected rotational constants can then be used to obtain a least-squares fit equilibrium geometry for pyrimidine. To determine whether the equilibrium structure determined is consistent with what we expect, a useful metric to use is the inertial defect. For a planar molecule such as pyrimidine, the moments of inertia must, by definition, cancel out exactly in a true equilibrium structure.

Equilibrium Structure and Comparison to Other High-Level Structures

The equilibrium geometry, with and without electron mass correction, as well as a vibrationally average geometry have been compared to previous experimental structures and the highest level theoretical structures feasible for a molecule this size. In expected fashion, the CCSD(T)/cc-pCVQZ and CCSD(T)/cc-pwCVQZ calculated structures have the closest agreement with our experimental equilibrium structure. The quadruple-ζ calculations are expected to be the best purely theoretical equilibrium structures to compare to as the effects of increasing the

treatment of electron correlation from CCSD(T) to FCI should be offset well by the effects of increasing the basis set size from quadruple-ζ to the infinite basis set limit.[20] From Figure 3 it is apparent that the computational parameters, even using the highest levels of theory, are frequently well outside the statistical error bars.

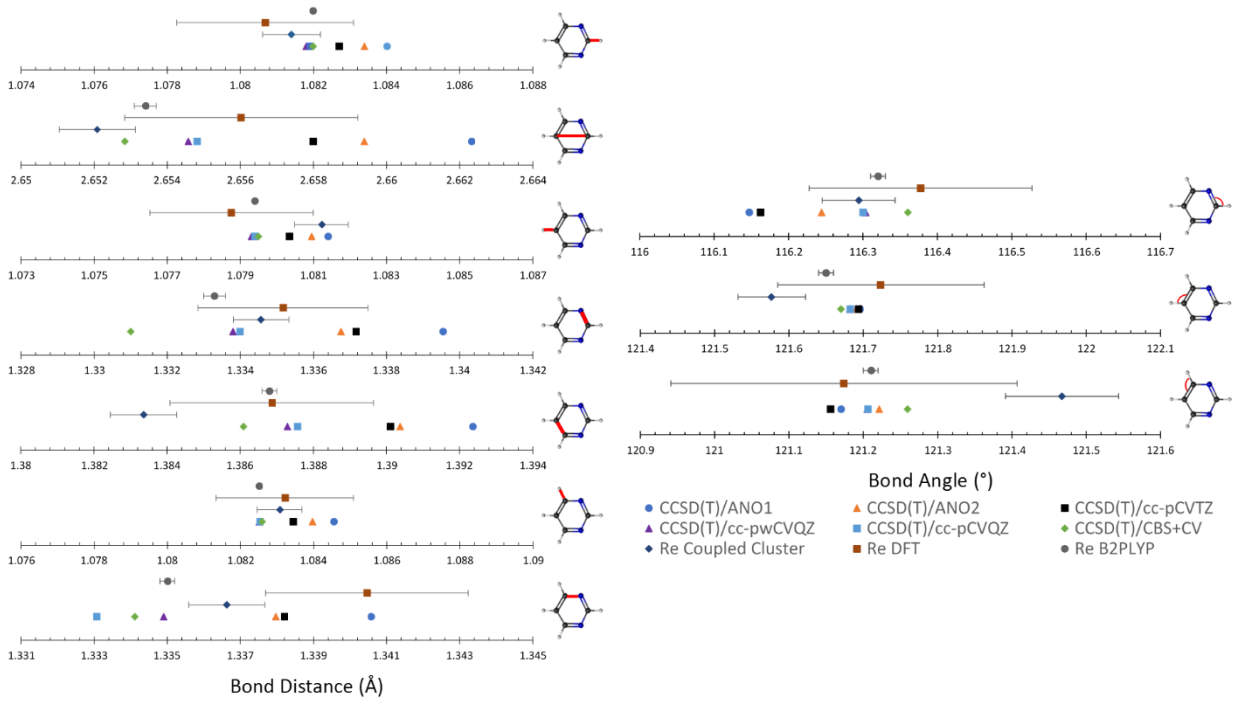


Figure 3. A graphical comparison of the interatomic distances and angles determined in this work to high-level computational structures. Re B2PLYP denotes the R_e^{SE} structure from Penocchio *et al.*,[21] CCSD(T)/CBS+CV denotes the calculated structure from Biczysko *et al.*,[22] and CCSD(T)/cc-pwCVQZ denotes the calculated structure from Csaszar *et al.*[17]

Upon comparing the equilibrium structures determined using *ab initio* and DFT methods, there is an apparent difference in both the errors and values determined for the structural parameters. As these structures are derived from the same experimental data, the most likely cause for these differences is the theoretical corrections applied to the experimental constants. As such, we believe that the existing differences between the purely theoretical equilibrium structures and the experimental equilibrium structures could be resolved by theoretical corrections calculated at

a higher level of theory and purely theoretical structures calculated with a basis set larger than cc-pCVQZ and treatment of electron correlation greater than CCSD(T), but such calculations are prohibitive.

Inertial Defect Trend Analysis and Treatment

In the analysis of the experimental equilibrium structures obtained by the previously described method, it was apparent upon inspection that there was a clear trend in the inertial defect of the isotopologues after correcting for electronic mass and vibration-rotation interaction. As shown in Figure 4, this trend is both present and remarkably similar to the inertial defect of the isotopologues of pyridazine.[6]

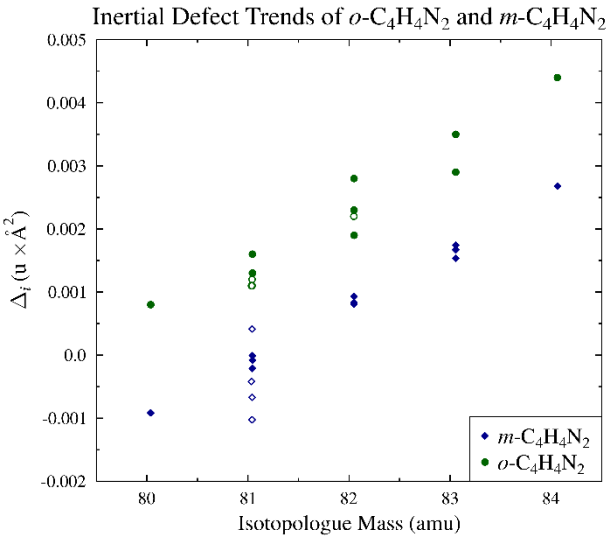


Figure 4. A trend in the inertial defect when compared to isotopologue mass.

This trend is particularly linear for the deuterium substituted isotopologues, which is reveals a clear systematic error in the methods used. This trend also suggests that the error lies almost entirely in the methodology employed rather than the experimental constants. If the error were to be entirely in the experimental constants, the inertial defect of all isotopologues would be expected

to be randomly distributed about 0 amu Å^2. As such, close inspection of the methodology employed was warranted. Initial investigation surrounded the vibrational-rotation interaction correction to the vibrationally-averaged rotational constants due to the known effects of out-of-plane vibrations on the inertial defect of planar molecules.[23] The propagation of the vibration-rotation interaction constant into the equilibrium inertial defect was determined starting from equation (7), absent the electronic mass correction. This was determined to take the form of equation (8).

$$I_e^C - I_e^B - I_e^A = (I_0^C - I_0^B - I_0^A) - \left(\frac{I_0^C}{2C_0}\sum \alpha_i^C - \frac{I_0^B}{2B_0}\sum \alpha_i^B - \frac{I_0^A}{2A_0}\sum \alpha_i^A\right) \quad (8)$$

It is suspected that at least part of the inaccuracy in the method somehow stems from inexact force constants, particularly for the five out-of-plane modes for the hydrogens. Anharmonic frequency calculations at the HF/-, B3LYP/-, and MP2/cc-pVTZ level have been performed but comparison of the force constants and correlation to the inertial defect trend has not yet been performed due to the length of the calculations required. In the process of investigating the effect of the vibration-rotation interaction correction on the inertial defect, the propagation of the electronic mass correction was determined, also starting from equation (7). The final form of the electronic mass correction takes the form of equation (9), where $B_e' = B_0 + \frac{1}{2}\sum \alpha_i^B$

$$I_e^C - I_e^B - I_e^A = (I_e^{C\prime} - I_e^{B\prime} - I_e^{A\prime}) + \left(\frac{I_0^C \eta g^{cc} C_e}{C_e'} - \frac{I_0^B \eta g^{bb} B_e}{B_e'} - \frac{I_0^A \eta g^{aa} A_e}{A_e'}\right) \quad (9)$$

This equation yielded significant insight into a way to improve this methodology. Previously, the electronic contribution to the rotational constant had been assumed to be constant for all isotopologues, and as such is only calculated once for the main isotopologue and applied to all isotopologues. This is correct to a certain degree, the distribution of electronic mass does not change upon isotopic substitution. This distribution of electronic mass is what is directly

calculated, g^{bb}, but this is only a single component of the electronic contribution to the equilibrium rotational constants. The factor of $\frac{I_0^B B_e}{B_e'}$ in the electronic contribution demonstrates the importance to use the rotational constants specific to the individual isotopologues for each of the individual corrections.

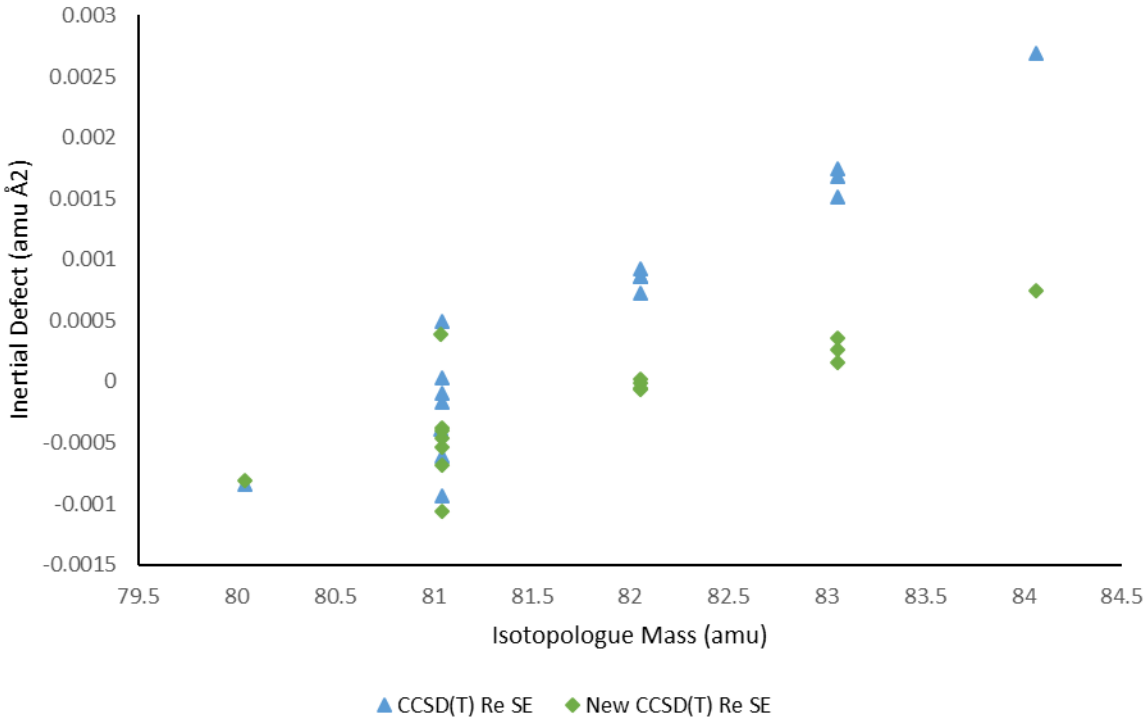


Figure 5. The inertial defect of isotopologues before and after using the new method for determining the electronic contribution to the rotational constants.

As seen in Figure 5, there is an appreciable improvement towards zero inertial defect upon use of the new methodology for determining the electronic contribution to the rotational constants. However, there is still a clear trend in the inertial defect that mirrors that of the original trend, with only a modestly shallower slope. Additional adaption to the methodology includes calculating the inertial *g*-tensor using the experimental equilibrium geometry while additionally using a larger basis set and a higher level of theory to better model the distribution of electronic mass. Though this correction to the methodology could potentially fix the remaining error in the slope, the error

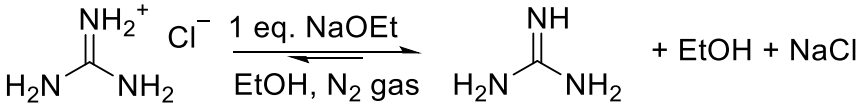
in intercept will likely remain. Even though more investigation into this remains, this advance shows a significant development in our understanding of the methodology behind the determination of equilibrium structures by rotational spectroscopy.

CHAPTER 4

SYNTHETIC EFFORTS TOWARDS ISOLATING NEUTRAL GUANIDINE FROM GUANIDINIUM CHLORIDE

Methods Employed

Guanidine has been both synthesized and studied previously by several groups.[10-11] These studies, however, were only in the solid phase and did not have concerns over the presence of volatile compounds in trace amounts. For a precise rotational spectroscopy study of guanidine, however, any trace contaminants more volatile than guanidine will be present in much higher quantity in the gas phase than in the solid phase due to the low volatility of guanidine. This in mind, the synthetic route employed in the was attempted, albeit under a nitrogen rather than argon gas as shown in Scheme 2.



Scheme 2. An adaption of the synthetic method for obtaining crystals of guanidine.

The workup was also slightly different in that the solvent was removed by rotary evaporation to precipitate guanidine solid rather than by cooling the solution. Upon rotary evaporation, an orange solid formed rather than the small white crystals expected based upon previous work.

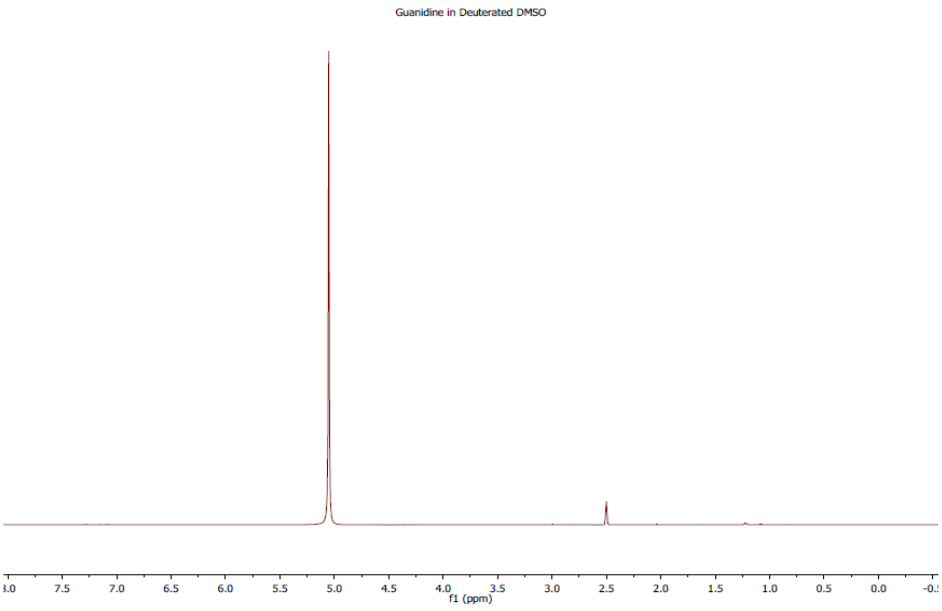
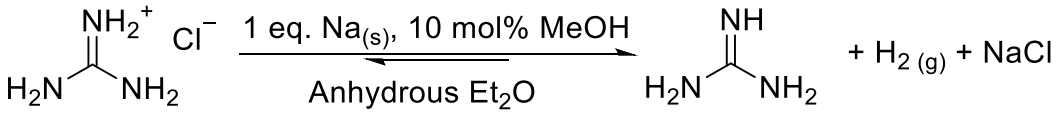


Figure 6. The NMR spectrum obtained from the reaction shown in Scheme 2.

This orange solid is thought to be a polymeric product of neutral guanidine reacting with itself under the basic conditions. An ^{1}H-NMR in deuterated DMSO was taken, yielding the spectrum shown in Figure 6. Based on this spectrum, it seemed apparent that any reaction materials Additionally, the ethanol proved to be very difficult to remove efficiently, so the same reaction pictured above, only with methanol rather than ethanol as the solvent and sodium methoxide rather than sodium ethoxide, was performed in an attempt to find a solvent system that would be more easily removed. Likewise, this reaction produced the same orange sludge as the ethoxide experiments. Based on these results, using diethyl ether as a new solvent system that would be even easier to remove quickly was explored. The reaction was carried out as shown in Scheme 3. In diethyl ether, neither the guanidinium chloride or sodium are very soluble, so it was discovered that a small amount of an alcohol was needed to mediate the reaction and form an ethoxide or methoxide *in situ*.



Scheme 3. The neutralization of guanidinium chloride in diethyl ether.

Again, upon workup an orange solid was formed even though the solvent was readily removed by rotary evaporation. As such, new methods for isolating the pure solid from the reaction mixture were explored. The first new method involved using a closed vacuum system to evaporate the ether solvent from the guanidine solution after filtering off the solid product. This was achieved by using a liquid nitrogen trap adjacent to the sample to remove vapor from the closed system, causing the solvent to boil off. Again, an orange solid was obtained after removing the solvent. Once it was clear that the orange impurity was not volatile and could not be removed by vacuum, the guanidine purification was attempted by sublimation of the guanidine onto a cold finger. This was evidently achieved by heating the sample to about 200 ^{O}C under an aspirator vacuum. However, mass spectrometry analysis showed the tan product obtained was likely just stopcock grease and not the desired product. The cumulative failure of each of the preceding experiments indicated that the desired product may have significant air, water, and temperature sensitivity issues, so the diethyl ether neutralization was attempted again only with a rapid filtration and removal of the ether solvent, and finally storage at low temperature. This yielded a collection of small white crystals, the same as was seen by previous experiments. However, once this sample was obtained the spectrometer was found to be in need of maintenance and analysis by rotational spectroscopy could be performed. The sample had since turned a slight orange and was thought to have formed the polymeric contaminant, but a successful method for obtaining a sample of the purified, free guanidine suitable for study by rotational spectroscopy is thought to have been found.

Computational Predictions

In anticipation of collecting rotational spectra of the various isomers of guanidine, the principle rotational constants of the *anti*- and *syn*-isomers of guanidine were calculated at the B2PLYP/cc-pVTZ level. In addition, anharmonic frequency calculations were also performed at the same level of theory and basis to determine the quartic centrifugal distortion constants as well as the vibration-rotation interaction constants. Given in Appendix B, the predicted rotational constants for both isomers are surprisingly similar. This would be anticipated to significantly complicate interpretation of the spectrum collected as the *syn*-isomer, due to the energy difference and the ease of conversion between *syn*- and *anti*-isomers, is expected to be an order of magnitude less abundant than the *anti*-isomer. Detection of the *syn*-isomer will therefore depend on the larger predicted b- and c-components of the dipole moment. As seen in Table 1, μ_C of the *syn*-isomer is nearly two Debye, which indicates the spectrum will have intense *c*-type R-branch transitions.

Table 1. The predicted dipole components of the *anti*- and *syn*-isomers of guanidine.

	Anti Conf (D)	Syn Conf (D)
μ_A	1.74	1.51
μ_B	1.88	2.46
μ_C	0.08	1.88

Expecting to also calculate an experimental equilibrium structure for one or both isomers of guanidine, high level theoretical computations of the equilibrium structure have been performed. These can be found in Appendix B.

CHAPTER 5

FUTURE DIRECTION

Both the pyrimidine and guanidine projects could be expanded on significantly in the future. Of particular interest is the elucidation of the cause for the trend in inertial defect. In particular, parsing through the force constants obtained at various levels of theory and comparing the change in force constants for individual vibrational modes to the change in trend would seem to be a worthwhile endeavor. Additionally, using an improved electronic correction by performing a calculation using the existing equilibrium geometry at both a higher level of theory and with an expanded basis set might prove useful. Finally, if none of these avenues of research remove the trend in inertial defect, a new anharmonic frequency calculation at the CCSD(T)/cc-pCVQZ level could yield a better calculated value of the vibration-rotation interaction correction to the rotational constants. As this last option is likely the most computationally prohibitive, it should only be explored once the other directions have been exhausted.

If the work on guanidine were to be continued, some directions of research may include looking into alternative methods of forcing the molecule into the gas phase other than just opening a sample of the purified solid to vacuum. In particular, investigation into outfitting the spectrometer with a laser oblation apparatus would prove useful for looking at molecules such as guanidine where volatility is a concern. This apparatus could also be used for the study of gas-phase ions that cannot currently be obtained by discharge chemistry. Improving in this particular area could vastly increase the range of molecules able to be studied with our instrument.

Another potential direction would involve the development of a synthetic route towards isotopically-labeled guanidine from urea and ammonium chloride, which are substantially cheaper isotopically-labeled reagents than labeled guanidine. Due to the low-lying isomers of guanidine as well as the low frequency vibrational modes, the spectrum of any given guanidine isotopologue could easily be masked by that of an isotopologue which is present in greater abundance. This would dictate that any synthesis of isotopically labeled guanidine be clean and result in few isotopologues as products under any given set of conditions, quite opposite the goals of the pyrimidine synthetic work.

Bibliography

1. Cernicharo, J.; Heras, A. M.; Tielens, A. G. G. M.; Pardo, J. R.; Herpin, F.; Guelin, M.; Waters, L. B. F. M., Infrared space observatory's discovery of C_4H_2, C_6H_2, and benzene in CRL 618. *Astrophys. J.* **2001,** *546* (2), L123-L126.
2. Sarre, P. J., The diffuse interstellar bands: A major problem in astronomical spectroscopy. *J. Mol. Spectrosc.* **2006,** *238* (1), 1-10.
3. Duley, W. W., Polycyclic aromatic hydrocarbons, carbon nanoparticles and the diffuse interstellar bands. *Faraday Discuss.* **2006,** *133* (0), 415-425.
4. Douglas, M. H.; Charles W. Bauschlicher, Jr.; Allamandola, L. J., Variations in the Peak Position of the 6.2 µm Interstellar Emission Feature: A Tracer of N in the Interstellar Polycyclic Aromatic Hydrocarbon Population. *Astrophys. J.* **2005,** *632* (1), 316.
5. Blackman, G. L.; Brown, R. D.; Burden, F. R., Microwave spectrum, dipole moment, and nuclear quadrupole coupling constants of pyrimidine. *J. Mol. Spectrosc.* **1970,** *35*, 444-54.
6. Esselman, B. J.; Amberger, B. K.; Shutter, J. D.; Daane, M. A.; Stanton, J. F.; Woods, R. C.; McMahon, R. J., Rotational spectroscopy of pyridazine and its isotopologs from 235-360 GHz: Equilibrium structure and vibrational satellites. *J. Chem. Phys.* **2013,** *139*, 224304-1 - 224304-13.
7. Schneider, R. F. The Microwave Spectrum of Pyrimidine. Ph.D. Dissertation, Columbia University, New York, 1959.
8. Caminati, W.; Damiani, D., Easy assignment of rotational spectra of slightly abundant isotopic species in natural abundance: carbon-13 and nitrogen-15 isotopic species of pyrimidine. *Chem. Phys. Lett.* **1991,** *179*, 460-2.
9. Kisiel, Z.; Pszczolkowski, L.; Lopez, J. C.; Alonso, J. L.; Maris, A.; Caminati, W., Investigation of the Rotational Spectrum of Pyrimidine from 3 to 337 GHz: Molecular Structure, Nuclear Quadrupole Coupling, and Vibrational Satellites. *J. Mol. Spectrosc.* **1999,** *195*, 332-339.
10. Yamada, T.; Liu, X.; Englert, U.; Yamane, H.; Dronskowski, R., Solid-State Structure of Free Base Guanidine Achieved at Last. *Chemistry – A European Journal* **2009,** *15* (23), 5651-5655.
11. Sawinski, P. K.; Meven, M.; Englert, U.; Dronskowski, R., Single-Crystal Neutron Diffraction Study on Guanidine, CN3H5. *Crystal Growth & Design* **2013,** *13* (4), 1730-1735.
12. Amberger, B. K.; Esselman, B. J.; Woods, R. C.; McMahon, R. J., Millimeter-wave Spectroscopy of Carbonyl Diazide, $OC(N_3)_2$. *J. Mol. Spectrosc.* **2014,** *295*, 15-20.
13. Kisiel, Z., Assignment and Analysis of Complex Rotational Spectra. In *Spectroscopy from Space*, Demaison, J.; Sarka, K.; Cohen, E. A., Eds. Springer Netherlands: Dordrecht, 2001; pp 91-106.
14. Kisiel, Z. PROSPE - Programs for ROtational SPEctroscopy. http://info.ifpan.edu.pl/~kisiel/prospe.htm (accessed 06-19-2016).
15. Albert, S.; Quack, M., High resolution rovibrational spectroscopy of pyrimidine: Analysis of the B_1 modes ν_{10b} and ν_4 and B_2 mode ν_{6b}. *J. Mol. Spectrosc.* **2007,** *243*, 280-291.
16. Gordy, W.; Cook, R., *Microwave Molecular Spectra*. 3rd. ed.; Wiley Interscience: New York, 1984; Vol. XVIII.
17. Csaszar, A. G.; Demaison, J.; Rudolph, H. D., Equilibrium Structures of Three-, Four-, Five-, Six-, and Seven-Membered Unsaturated N-Containing Heterocycles. *J. Phys. Chem. A* **2015,** *119*, 1731-1746.
18. Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V., Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules. *J. Phys. Chem. A* **2015,** *119*, 2058-2082.
19. Stanton, J. F.; Gauss, J.; Christiansen, O., Equilibrium geometries of cyclic SiC3 isomers. *J. Chem. Phys.* **2001,** *114* (7), 2993-2995.

20. Bartlett, R. J.; Stanton, J. F., Applications of Post-Hartree—Fock Methods: A Tutorial. In *Rev. Comput. Chem.*, John Wiley & Sons, Inc.: 1994; Vol. 5, pp 65-169.
21. Penocchio, E.; Piccardo, M.; Barone, V., Semiexperimental Equilibrium Structures for Building Blocks of Organic and Biological Molecules: The B2PLYP Route. *Journal of Chemical Theory and Computation* **2015,** *11* (10), 4689-4707.
22. Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G.; Puzzarini, C.; Santoro, F.; Trani, F.; Villani, G., Integrated computational approaches for spectroscopic studies of molecular systems in the gas phase and in solution: pyrimidine as a test case. *Theor. Chem. Acc.* **2012,** *131*, 1-19.
23. Oka, T., On negative inertial defect. *J. Mol. Struct.* **1995,** *352*, 225-233.

Appendix A

Pyrimidine Rotational Constants and Structure Data

Table A1. CCSD(T)/ANO1 Calculated and Experimental Isotopologue Rotational Constants

	Calculated[a]			Experimental		
Pyrimidine Isotopologue	A_0 **(MHz)**	B_0 **(MHz)**	C_0 **(MHz)**	A_0 **(MHz)**	B_0 **(MHz)**	C_0 **(MHz)**
$C_4H_4N_2$	6276.818198	6067.159924	3084.431113	6276.82775	6067.16576	3084.44922
[2-^{13}C]	6152.624515	6067.534924	3054.227736	6152.67866	6067.53928	3054.25147
[4-^{13}C]	6256.759348	5956.689819	3050.849265	6256.09741	5957.21851	3050.84565
[5-^{13}C]	6132.413141	6068.119924	3049.408255	6132.81727	6067.35963	3049.30507
[1-^{15}N]	6253.930391	5953.836988	3049.435897	6253.95822	5954.15235	3049.53046
[2-^{2}H]	6066.955924	5871.677661	2983.304305	6066.85370	5871.75220	2983.30009
[4-^{2}H]	6242.436949	5694.446029	2977.371313	6242.39729	5694.50625	2977.37753
[5-^{2}H]	6067.032924	5840.068036	2975.117329	6066.96171	5840.14871	2975.12051
[2,4-^{2}H]	5934.927054	5604.023233	2881.909095	5934.80590	5604.13160	2881.89348
[2,5-^{2}H]	6066.779924	5479.168460	2878.547947	6066.66450	5479.25420	2878.52975
[4,5-^{2}H]	5914.004794	5602.322107	2876.497317	5913.95177	5602.39335	2876.48887
[4,6-^{2}H]	6082.429608	5457.710748	2876.121966	6082.41737	5457.75988	2876.11603
[2,4,5-^{2}H]	5809.776500	5350.987305	2785.101163	5809.65990	5351.08530	2785.07418
[2,4,6-^{2}H]	5692.633442	5457.520748	2785.932967	5692.64799	5457.51365	2785.90867
[4,5,6-^{2}H]	5680.003899	5457.613748	2782.910390	5680.01344	5457.63849	2782.89077
[2,4,5,6-^{2}H]	5457.464748	5330.617899	2696.353015	5457.39248	5330.69699	2696.31571

[a]Derived from calculated equilibrium constants with corrections for rotation-vibration interactions and electron mass calculated at the same level.

Table A2. CCSD(T)/ANO1 Calculated Isotopologue Quartic Distortion Constants in the S-reduction, I^r and III^r representations

	S-reduction, III^r representation					S-reduction, I^r representation				
Pyrimidine Isotopologue	D_J **(kHz)**	D_{JK} **(kHz)**	D_K **(kHz)**	d_1 **(kHz)**	d_2 **(kHz)**	D_J **(kHz)**	D_{JK} **(kHz)**	D_K **(kHz)**	d_1 **(kHz)**	d_2 **(kHz)**
$C_4H_4N_2$	1.44226	-2.39873	0.98818	0.00696	0.02131	0.48895	1.42497	-0.52819	-0.32299	-0.13932
[2-^{13}C]	1.41469	-2.35482	1.01737	0.01720	0.01777	0.48535	1.38238	-0.52299	-0.32367	-0.14043
[4-^{13}C]	1.42606	-2.42508	0.99038	0.01741	0.02366	0.48361	1.39209	-0.53180	-0.32367	-0.14141
[5-^{13}C]	1.40207	-2.28644	0.97155	0.00102	0.01717	0.47595	1.40049	-0.51076	-0.31265	-0.13426
[1-^{15}N]	1.41352	-2.35443	0.93817	0.00105	0.02353	0.47478	1.41056	-0.52098	-0.31230	-0.13460
[2-^{2}H]	1.30995	-2.17255	0.91082	-0.05934	0.00754	0.40358	1.31145	-0.30148	-0.26636	-0.11995
[4-^{2}H]	1.30803	-2.18832	0.79845	-0.02175	0.02383	0.42896	1.29717	-0.42224	-0.27641	-0.11755
[5-^{2}H]	1.30762	-2.16643	0.89646	-0.06042	0.00745	0.40498	1.29546	-0.28688	-0.26523	-0.11821
[2,4-^{2}H]	1.19784	-2.03942	0.82595	0.02627	0.01675	0.42729	1.06853	-0.38402	-0.27882	-0.11597
[2,5-^{2}H]	1.19850	-1.97119	0.71911	-0.10153	0.00600	0.37510	1.00142	0.03704	-0.22644	-0.08973
[4,5-^{2}H]	1.19455	-2.03554	0.83150	0.02825	0.01695	0.42583	1.07284	-0.39452	-0.27891	-0.11675
[4,6-^{2}H]	1.18368	-1.95171	0.67679	-0.04314	0.02006	0.37980	1.17984	-0.32982	-0.23881	-0.09992
[2,4,5-^{2}H]	1.08762	-1.80881	0.68906	-0.06599	0.00123	0.34467	0.96312	-0.09065	-0.21498	-0.08928
[2,4,6-^{2}H]	1.07833	-1.78338	0.72094	0.00102	0.01155	0.37733	0.98287	-0.30702	-0.24097	-0.09900
[4,5,6-^{2}H]	1.07340	-1.77727	0.72339	0.00373	0.01180	0.37632	0.98141	-0.31539	-0.24103	-0.09944
[2,4,5,6-^{2}H]	0.98553	-1.62578	0.69317	-0.03847	0.00260	0.31676	0.92959	-0.18908	-0.20441	-0.08892

[a]Derived from the calculated distortion constants in the S-reduction, I^r representation and the calculated rotational constants.

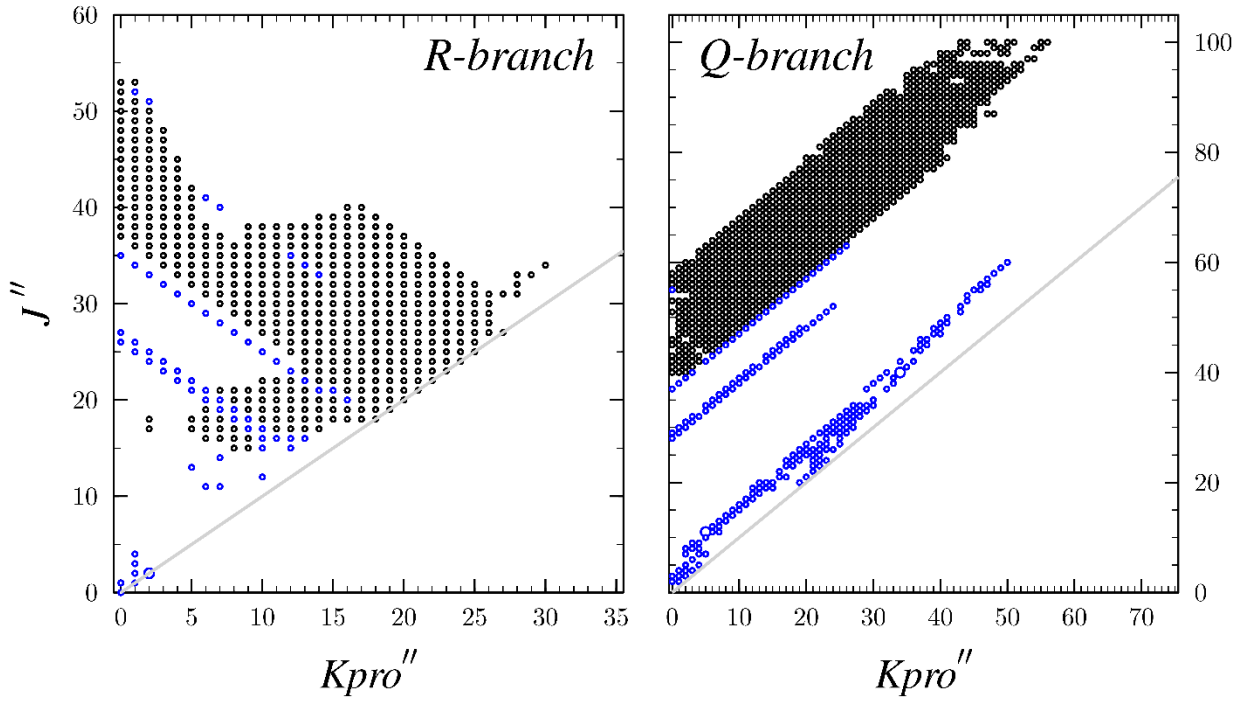
Table A3. Conversion from S- and A-reduced constants, representation-I^r to the average determinable constants

	$C_4H_4N_2$	[2-^{13}C]	[4-^{13}C]	[5-^{13}C]	[1-^{15}N]	[2-^{2}H]	[4-^{2}H]	[5-^{2}H]
Molecular Mass (u)	80.03744814	81.04080296	81.04080296	81.04080296	81.03448303	81.04372488	81.04372488	81.04372488
A^S	6276.828245	6152.67980	6256.09809	6132.81856	6253.95922	6066.874	6242.39794	6066.96219
A^A	6276.826815	6152.67835	6256.09679	6132.81683	6253.95788	6066.8731	6242.39693	6066.96101
A'' (A)	6276.828391	6152.67992	6256.09833	6132.81840	6253.95942	6066.87439	6242.39830	6066.96232
A'' (S)	6276.828385	6152.67992	6256.09831	6132.81868	6253.95944	6066.87409	6242.39829	6066.96229
A'' avg	6276.828388	6152.67992	6256.09832	6132.81854	6253.95943	6066.8742	6242.39830	6066.96230
A'' diff	0.000007	0.00000	0.00002	-0.00028	-0.00002	0.000	0.00001	0.00003
B^S	6067.165081	6067.53869	5957.21716	6067.35969	5954.15114	5871.7314	5694.50540	5840.14762
B^A	6067.166963	6067.54051	5957.21897	6067.36125	5954.15295	5871.7325	5694.50696	5840.1492
B'' (A)	6067.166296	6067.53987	5957.21829	6067.36062	5954.15226	5871.73205	5694.50630	5840.14873
B'' (S)	6067.166299	6067.53986	5957.21828	6067.36086	5954.15225	5871.73252	5694.50630	5840.14875
B'' avg	6067.166297	6067.53986	5957.21828	6067.36074	5954.15226	5871.7323	5694.50630	5840.14874
B'' diff	-0.000003	0.00002	0.00001	-0.00024	0.00001	-0.0005	0.00001	0.0000
C^S	3084.44974	3054.25179	3050.84571	3049.30578	3049.52992	2983.30057	2977.37812	2975.12069
C^A	3084.44901	3054.25114	3050.84497	3049.30503	3049.52944	2983.30009	2977.37737	2975.12005
C'' (A)	3084.45229	3054.25428	3050.84813	3049.30816	3049.53260	2983.30287	2977.38015	2975.12290
C'' (S)	3084.45229	3054.25428	3050.84813	3049.30828	3049.53233	2983.30275	2977.38016	2975.12290
C'' avg	3084.45229	3054.25428	3050.84813	3049.30822	3049.53246	2983.30281	2977.38016	2975.12290
C'' diff	0.00001	-0.00001	0.00000	-0.00012	0.00027	0.00011	-0.00001	0.00000
D_J	0.000500826	0.00049599	0.000505401	0.00049552	0.00050484	0.00040642	0.000482	0.000410701
D_{JK}	0.00145454	0.00141487	0.00128292	0.00142232	0.00128046	0.00131632	0.00091409	0.00133416
D_K	-0.00055031	-0.00054355	-0.00037453	-0.00055551	-0.0003708	-0.0002364	0.00004394	-0.00029249
d_1	-0.000331803	-0.00033169	-0.000323461	-0.000332143	-0.000324353	-0.000266359	-0.000285982	-0.000269444
d_2	-0.000143655	-0.000144596	-0.000131863	-0.000145492	-0.000132246	-0.000119946	-0.000102092	-0.000121029
Δ_J	0.000788135	0.00078519	0.00076913	0.00078641	0.00076942	0.00064635	0.00068619	0.00065274
Δ_{JK}	-0.00026929	-0.00032029	-0.00029934	-0.00032364	-0.00030612	-0.0001311	-0.00031084	-0.0001182
Δ_K	0.0008862	0.0009024	0.00094396	0.00089938	0.00095111	0.0009765	0.00106456	0.00091793
δ_J	0.000331802	0.000331696	0.000323461	0.000332102	0.000324319	0.000266359	0.000285986	0.000269434
δ_K	0.000655419	0.00061109	0.000635985	0.00060715	0.00063809	0.000540375	0.000573158	0.000560737

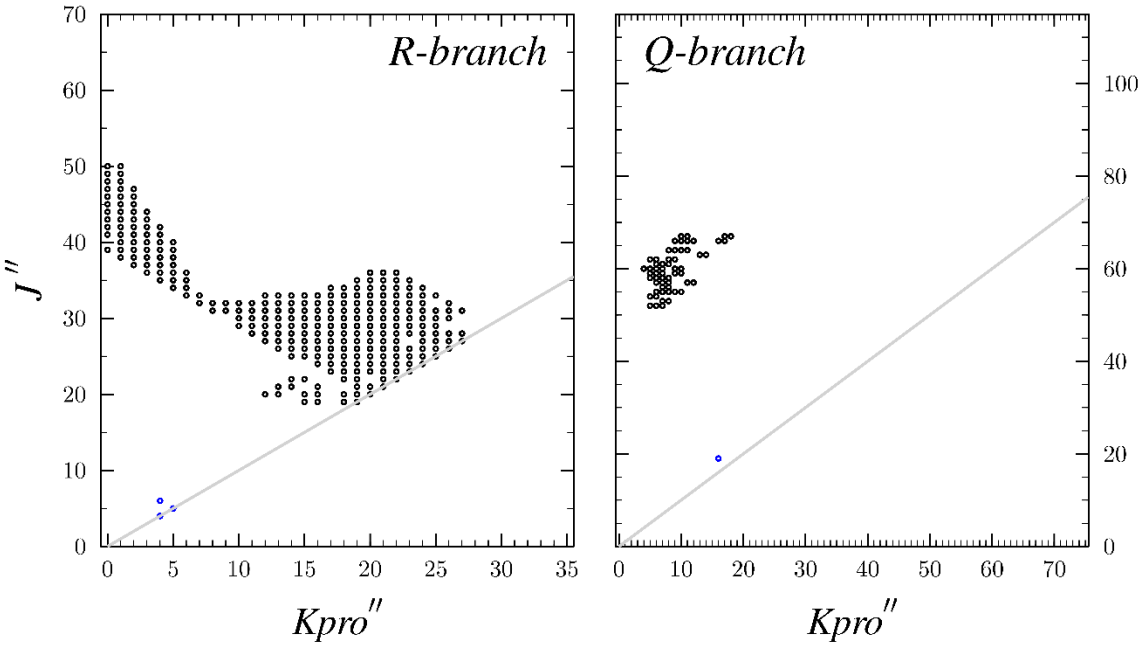
	[2,4-2H]	[2,5-2H]	[4,5-2H]	[4,6-2H]	[2,4,5-2H]	[2,4,6-2H]	[4,5,6-2H]	[2,4,5,6-2H]
Molecular Mass (u)	82.05000163	82.05000163	82.05000163	82.05000163	83.05627838	83.05627838	83.05627838	84.06255512
A^S	5934.8062	6066.6726	5913.95218	6082.41781	5809.6601	5692.64824	5680.01420	5457.39321
A^A	5934.8187	6066.6734	5913.95123	6082.41675	5809.6592	5692.64724	5680.01301	5457.39215
A'' (A)	5934.8200	6066.6745	5913.95249	6082.41794	5809.6602	5692.64842	5680.01419	5457.39316
A'' (S)	5934.8064	6066.6728	5913.95250	6082.41797	5809.6602	5692.64840	5680.01436	5457.39331
A'' ave.	5934.8132	6066.6736	5913.95249	6082.41796	5809.6602	5692.64841	5680.01427	5457.39323
A'' diff	0.0137	0.0017	0.00000	-0.00003	0.0000	0.00002	-0.00017	-0.00015
B^S	5604.1314	5479.2463	5602.39261	5457.75992	5351.0851	5457.51337	5457.63853	5330.69703
B^A	5604.1185	5479.2462	5602.39392	5457.76138	5351.0864	5457.51467	5457.63972	5330.69808
B'' (A)	5604.1179	5479.2458	5602.39339	5457.76083	5351.0860	5457.51418	5457.63923	5330.69774
B'' (S)	5604.1324	5479.2473	5602.39339	5457.76100	5351.0861	5457.51424	5457.63941	5330.69785
B'' ave.	5604.1251	5479.2465	5602.39339	5457.76091	5351.0860	5457.51421	5457.63932	5330.69779
B'' diff	-0.0145	-0.0015	0.00000	-0.00017	-0.0001	-0.00006	-0.00018	-0.00011
C^S	2881.89391	2878.53030	2876.489452	2876.11639	2785.07458	2785.90915	2782.89115	2696.31562
C^A	2881.89358	2878.52965	2876.488899	2876.11553	2785.07397	2785.90853	2782.89053	2696.31535
C'' (A)	2881.89600	2878.53217	2876.491283	2876.11840	2785.07641	2785.91094	2782.89294	2696.31743
C'' (S)	2881.89603	2878.53216	2876.491283	2876.11845	2785.07639	2785.91101	2782.89302	2696.31727
C'' ave.	2881.89602	2878.53217	2876.491283	2876.11842	2785.07640	2785.91097	2782.89298	2696.31735
C'' diff	-0.00003	0.00002	0.000000	-0.00005	0.00002	-0.00007	-0.00007	0.00016
D_J	0.000429771	0.0003679	0.000442341	0.00038949	0.00033863	0.00038668	0.00038485	0.00032167
D_{JK}	0.00117047	0.001037	0.00079866	0.00120078	0.0010727	0.0009954	0.0010137	0.0009531
D_K	-0.0005169	0.0001066	0.00007668	-0.00033979	-0.0001687	-0.0003155	-0.00032929	-0.0001915
d_1	-0.000278824	-0.0002217	-0.000263652	-0.000245132	-0.00021318	-0.00024683	-0.00024686	-0.00020771
d_2	-0.000115965	-0.00008973	-0.000094787	-0.000102708	-0.00009174	-0.000101515	-0.000102441	-0.000090812
Δ_J	0.000661735	0.0005461	0.000631922	0.00059483	0.00052205	0.00058967	0.000589677	0.00050328
Δ_{JK}	-0.00041579	-0.000043	-0.00033874	-0.00003162	-0.0000283	-0.0002227	-0.00021562	-0.00013655
Δ_K	0.0009825	0.0010208	0.00102448	0.00068727	0.0007493	0.0006996	0.00069513	0.00071638
δ_J	0.000278824	0.000221	0.000263656	0.000245096	0.00021315	0.00024682	0.000246833	0.00020769
δ_K	0.00047949	0.000516452	0.000465849	0.0006097	0.0004978	0.00047764	0.00047788	0.00039805

Pyrimidine Data Set Distribution Plots

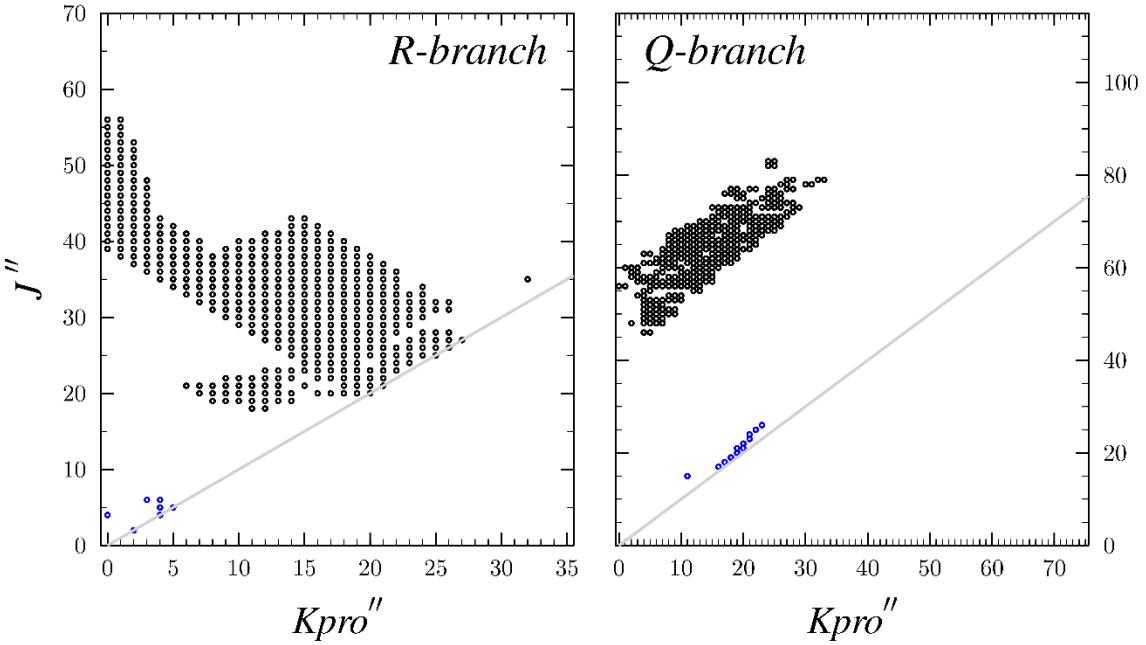
$C_4H_4N_2$



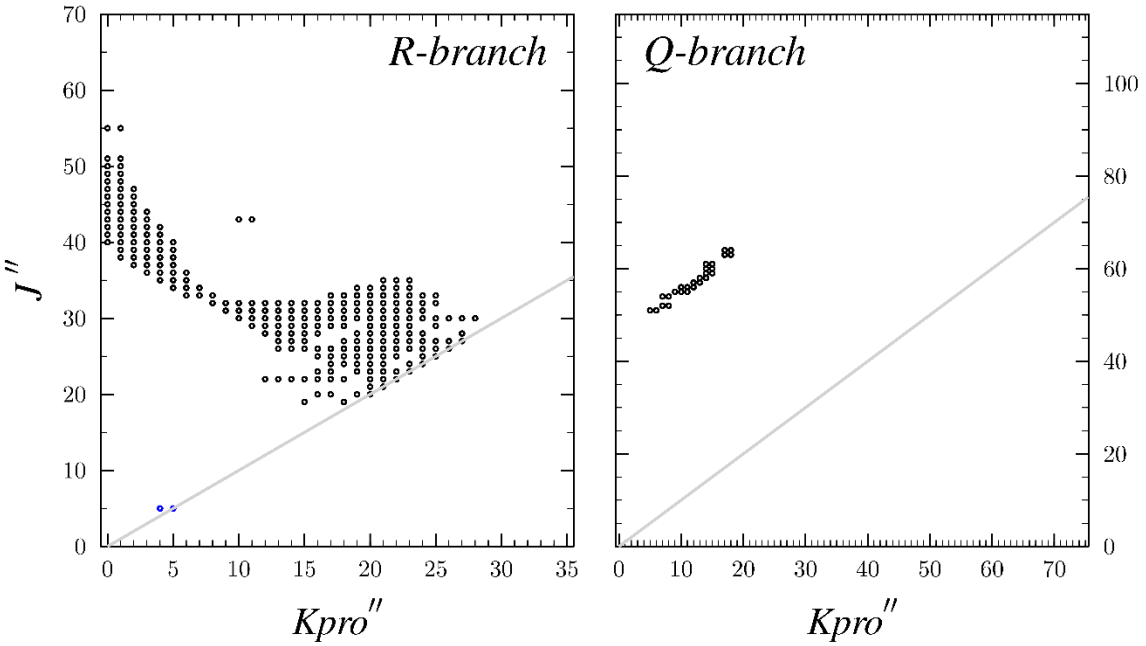
[2-^{13}C]



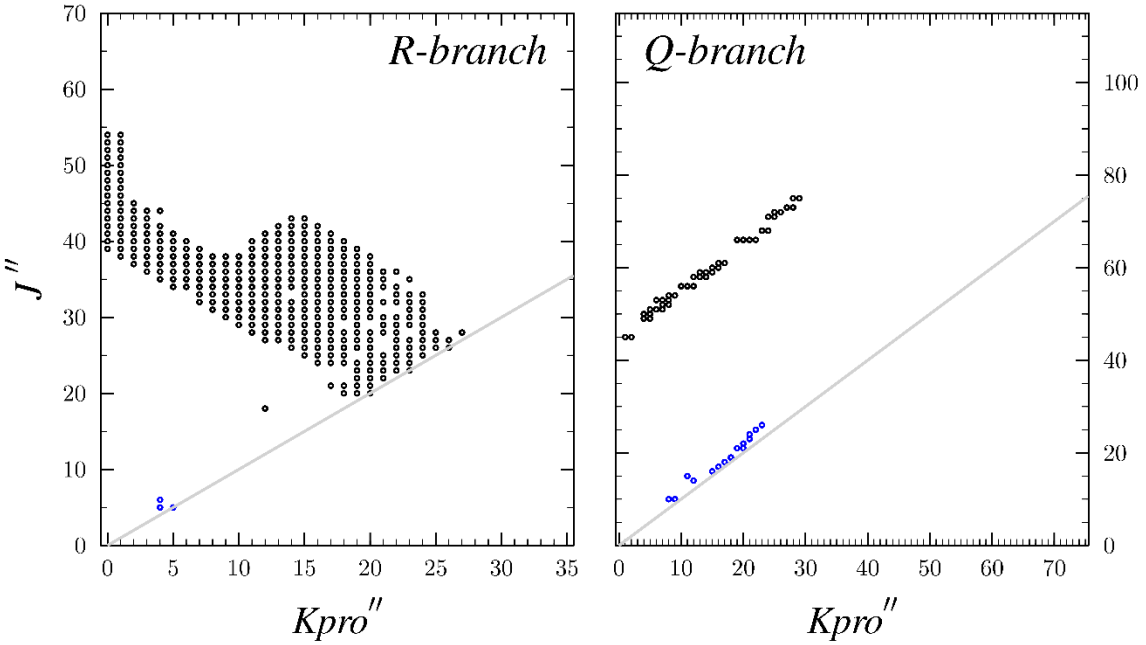
[4-^{13}C]



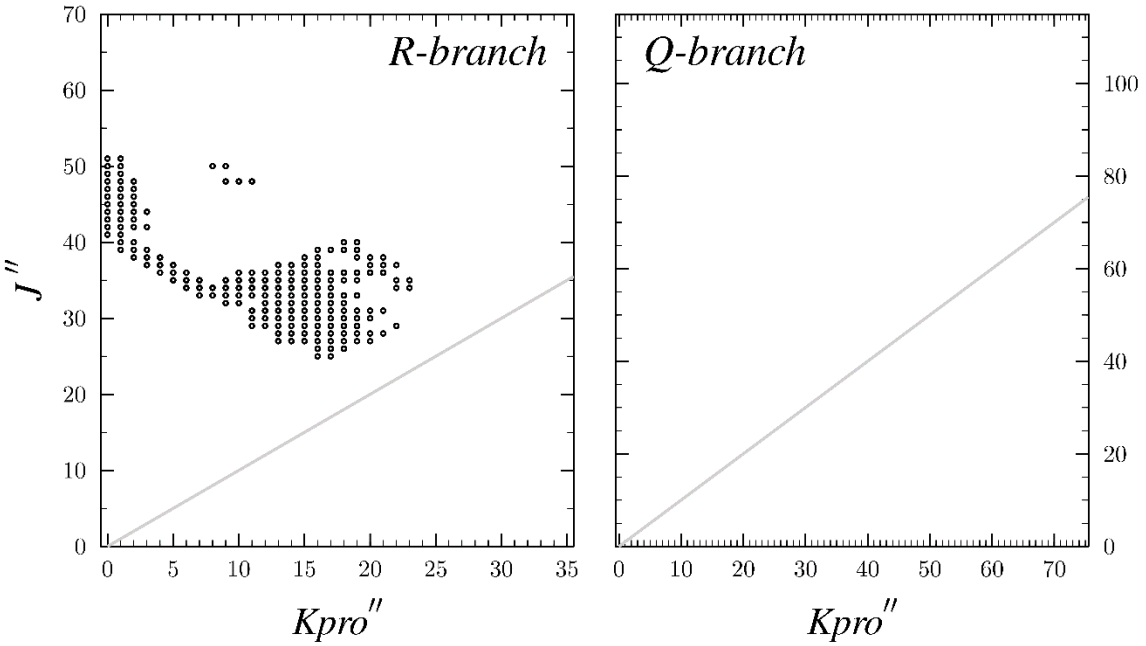
[5-^{13}C]



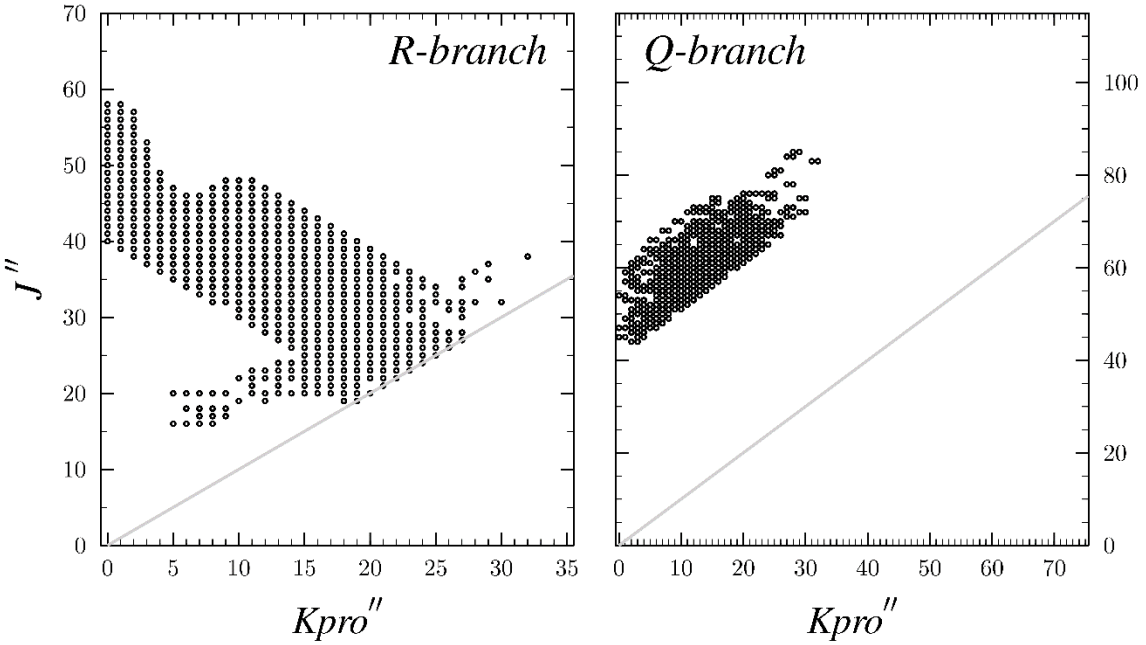
[1-^{15}N]



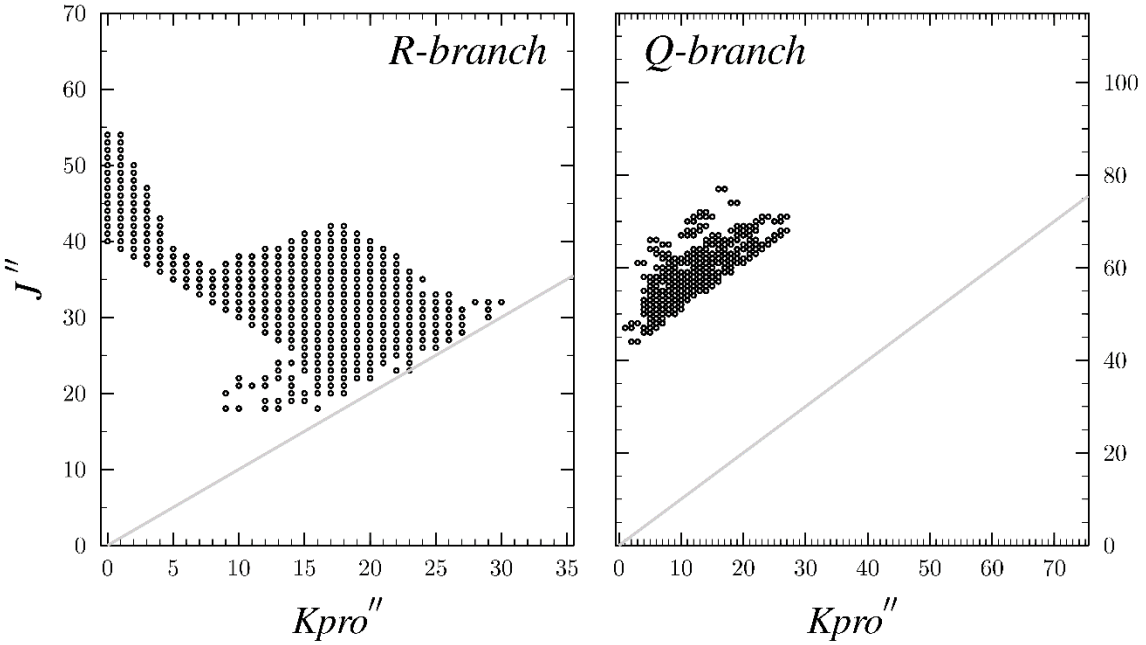
[2-^{2}H]



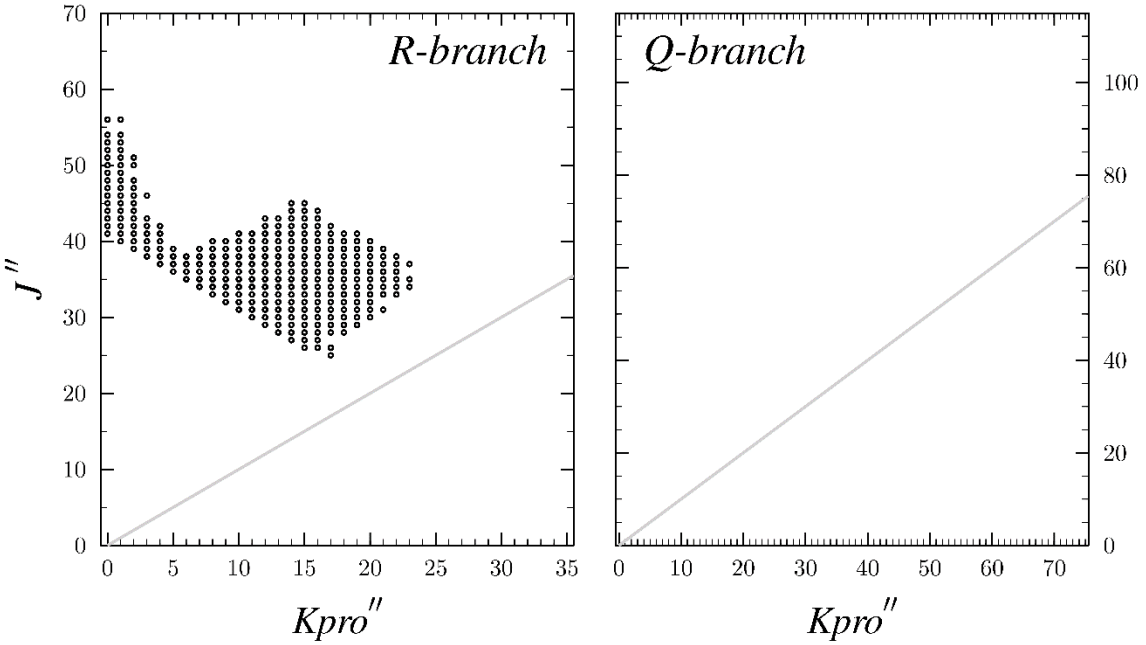
[4-^{2}H]



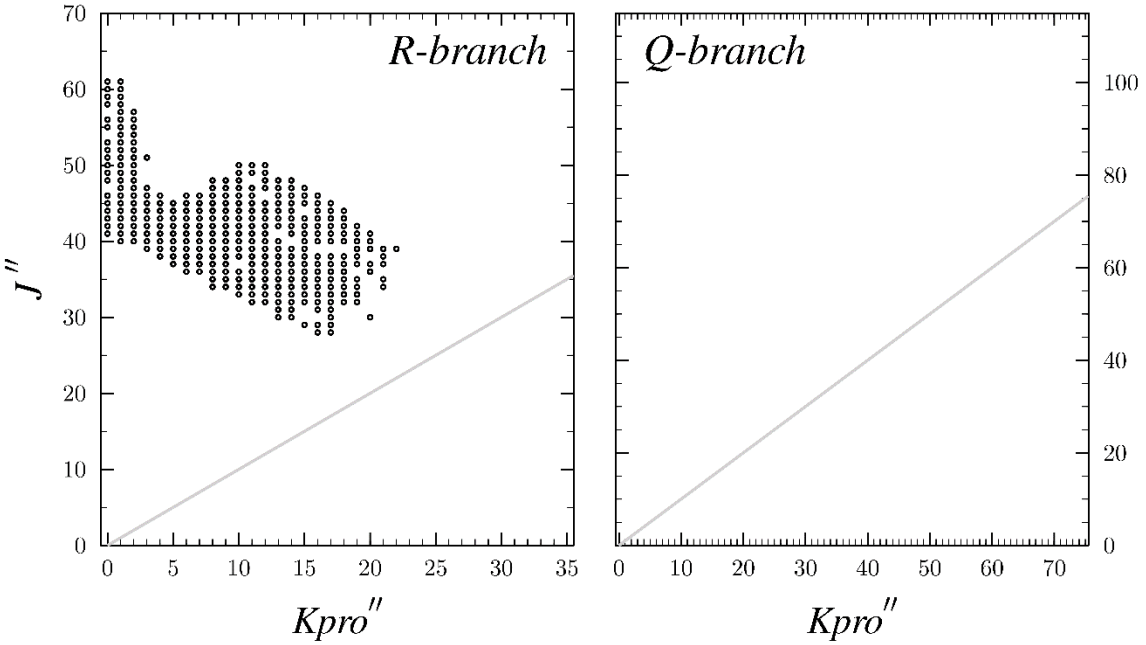
[5-^{2}H]



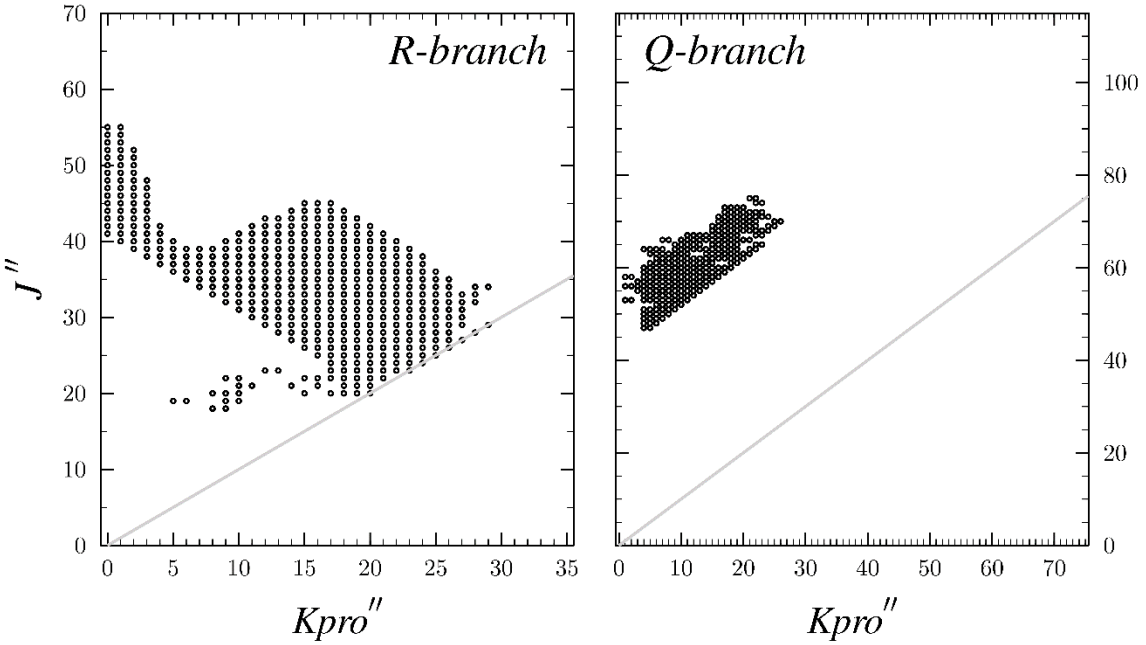
[2,4-^{2}H]



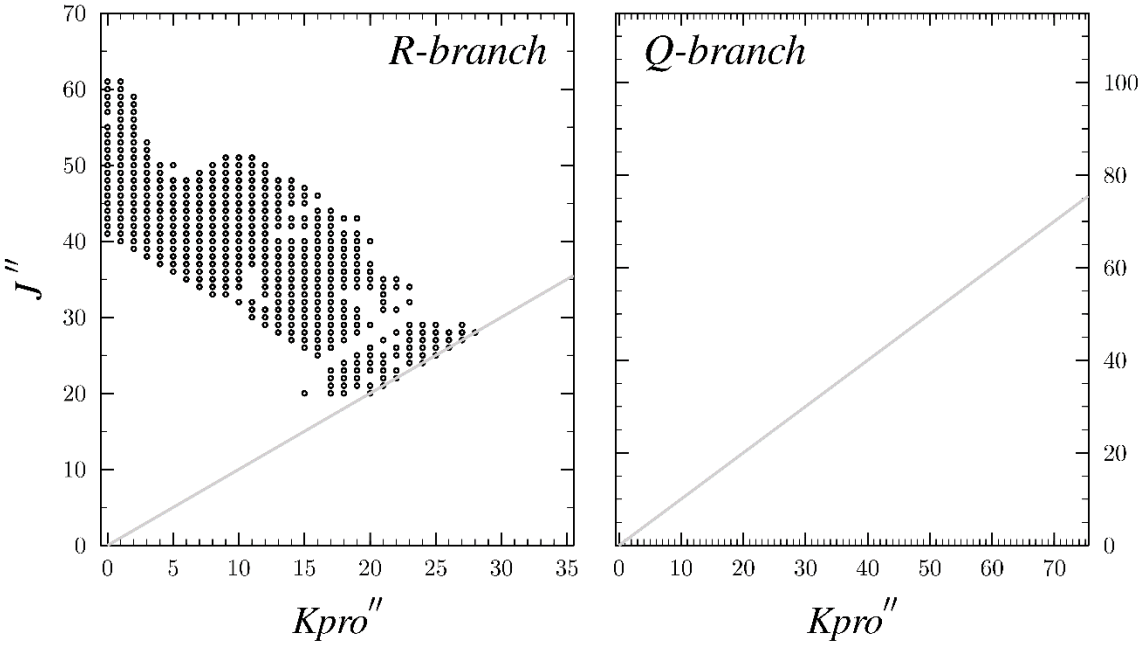
[2,5-^{2}H]



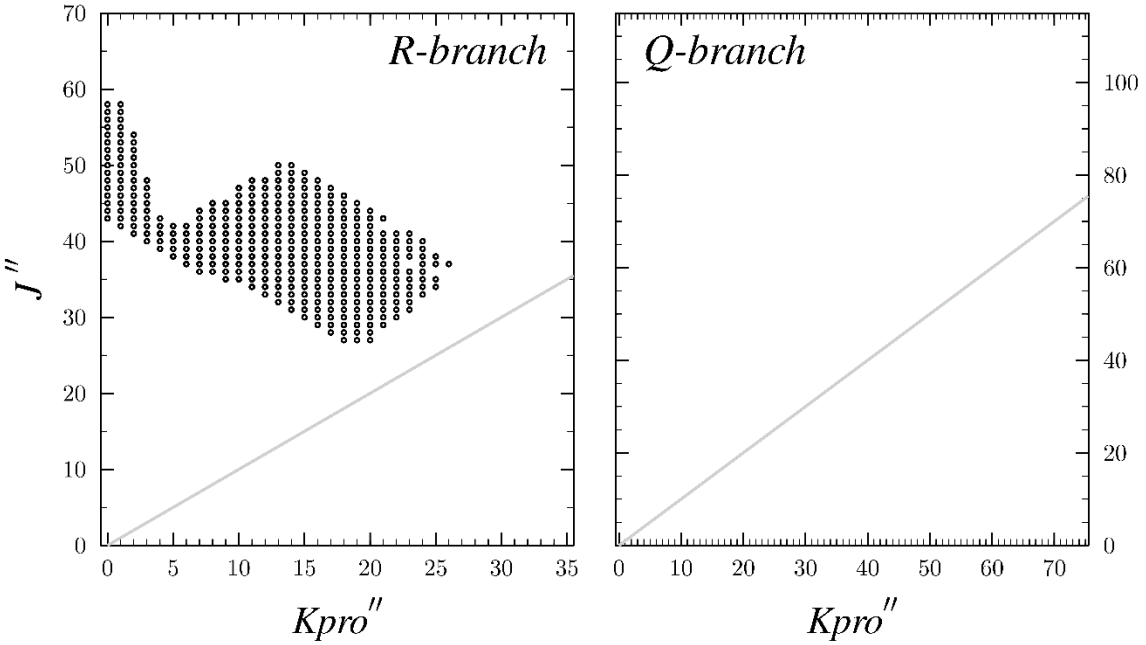
[4,5-^{2}H]



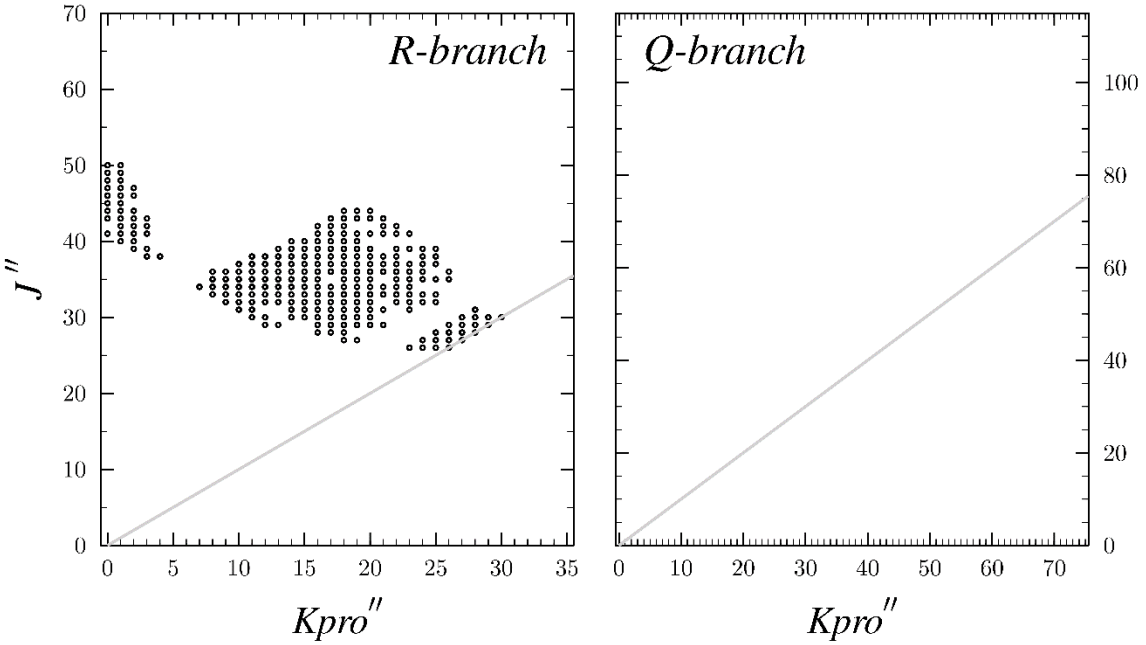
[4,6-^{2}H]



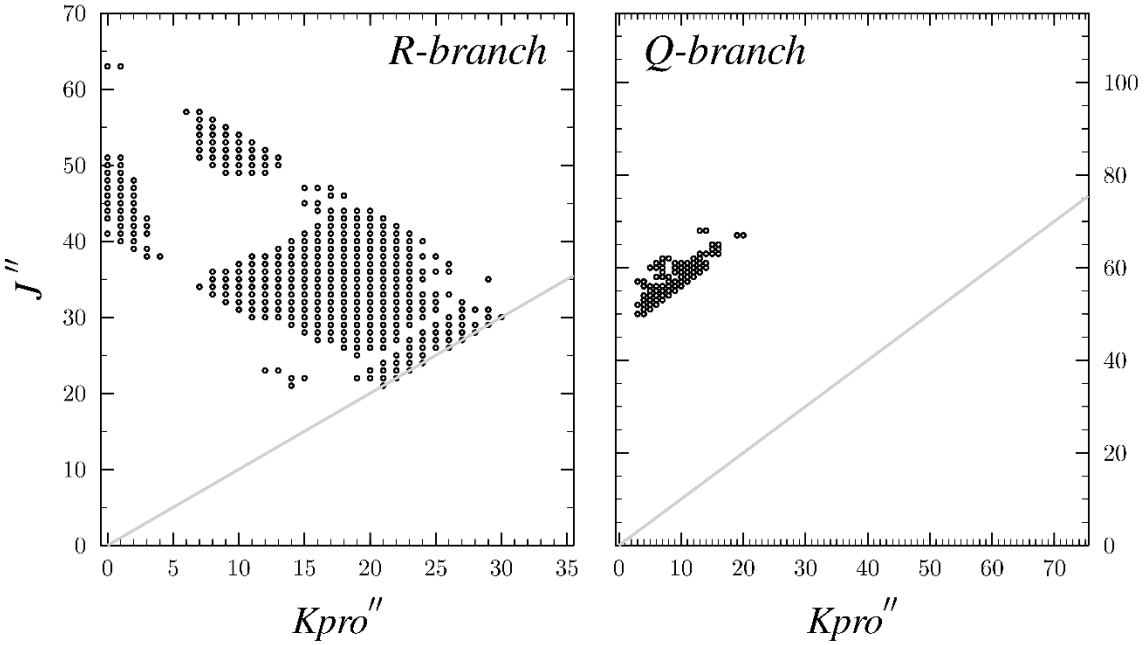
[2,4,5-^{2}H]



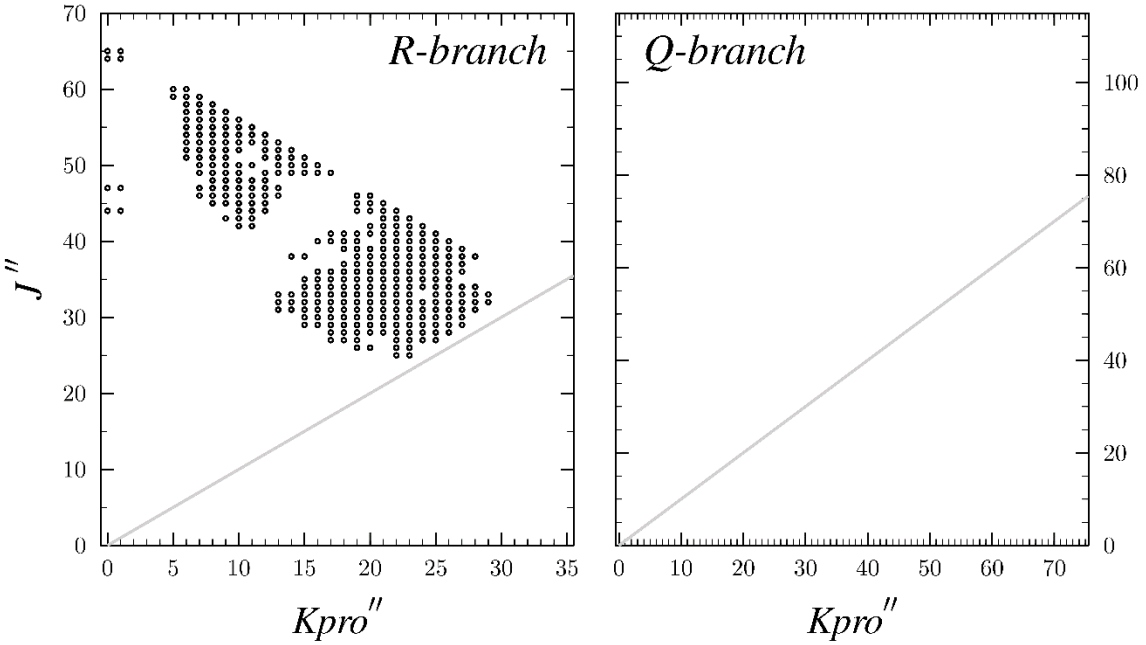
[2,4,6-^{2}H]



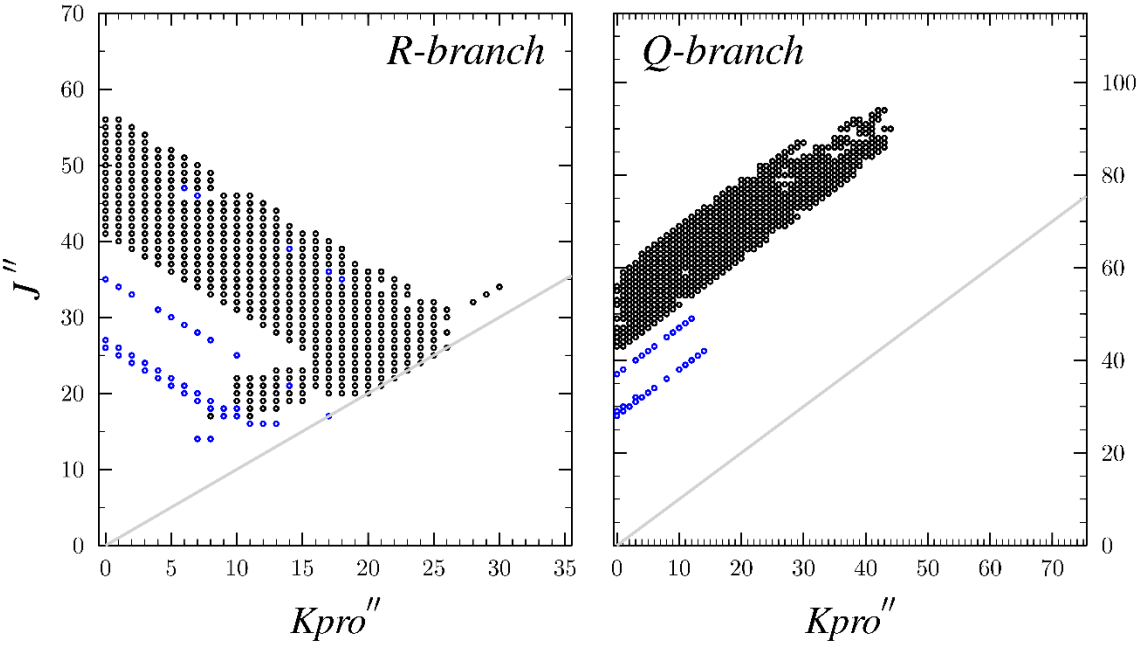
[4,5,6-^{2}H]



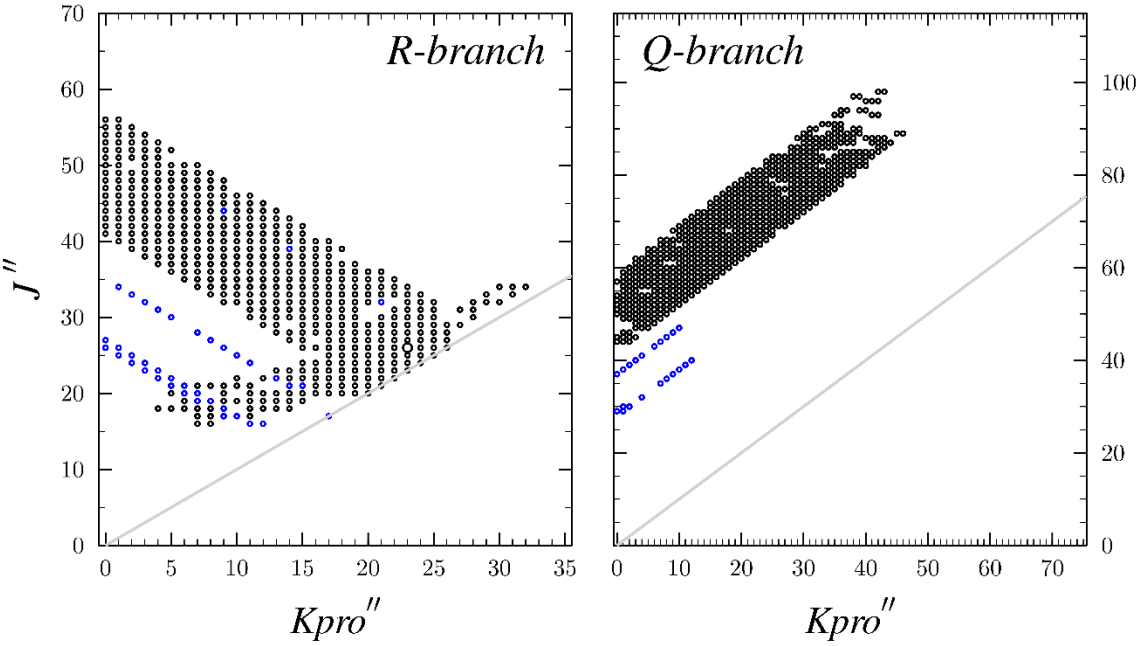
[2,4,5,6-^{2}H]



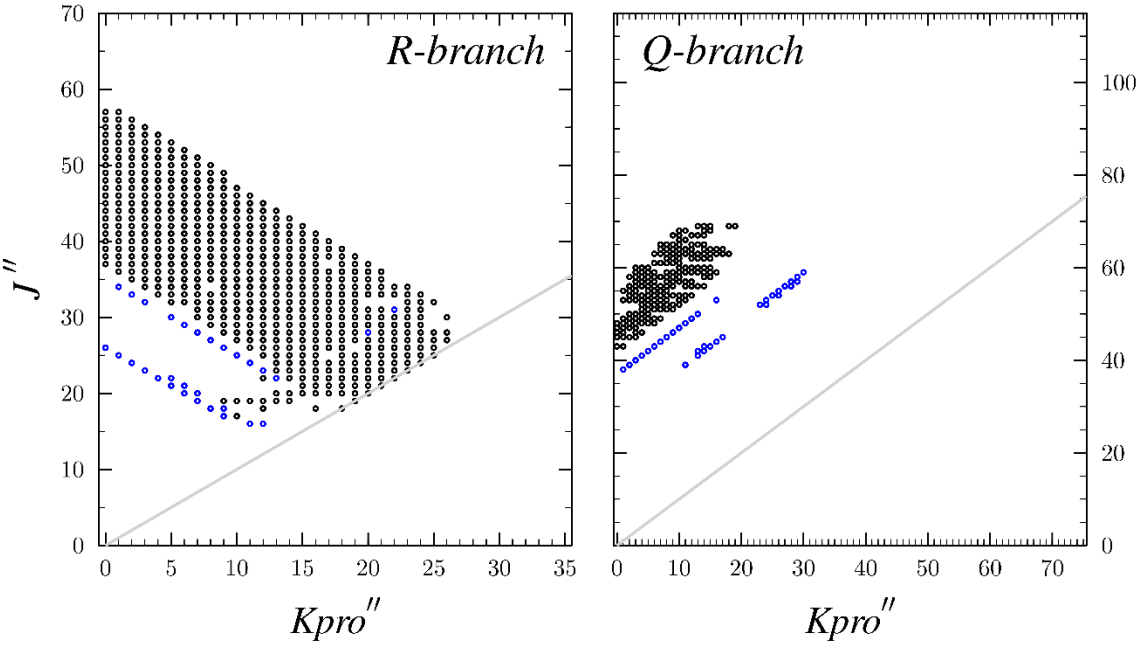
ν_{16}



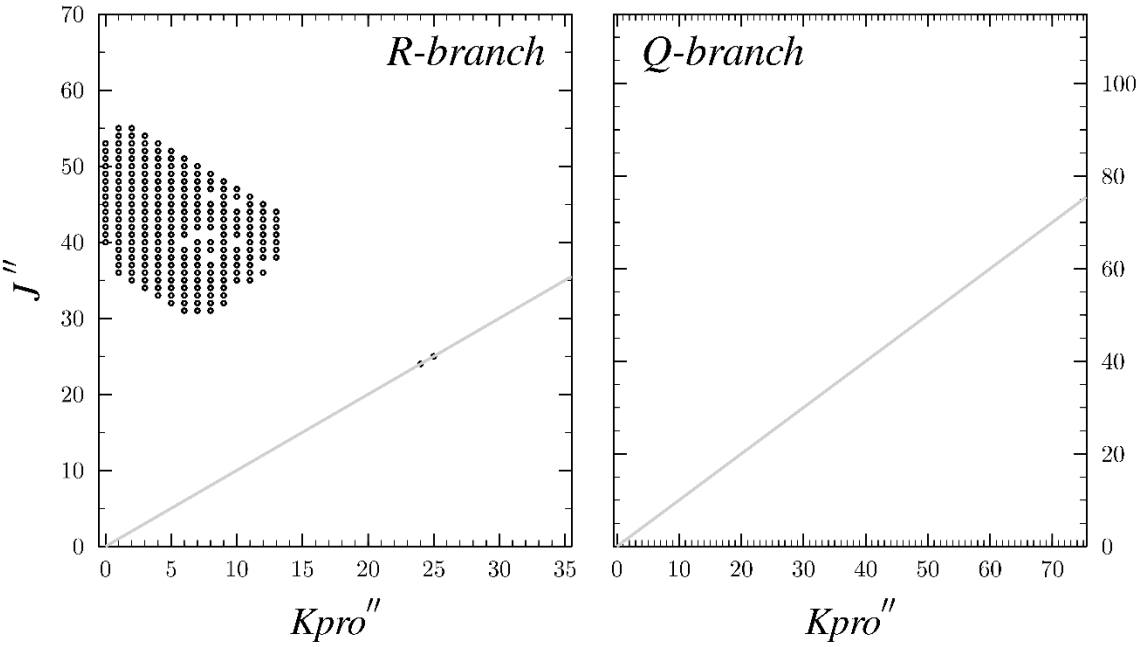
ν_{11}



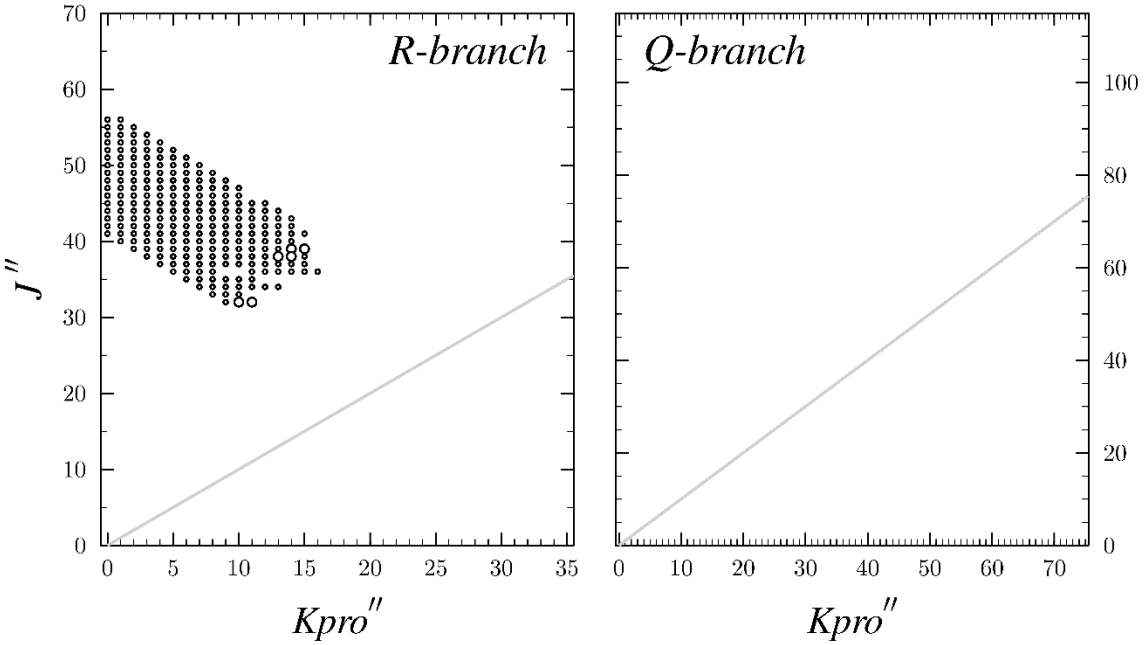
ν_{24}



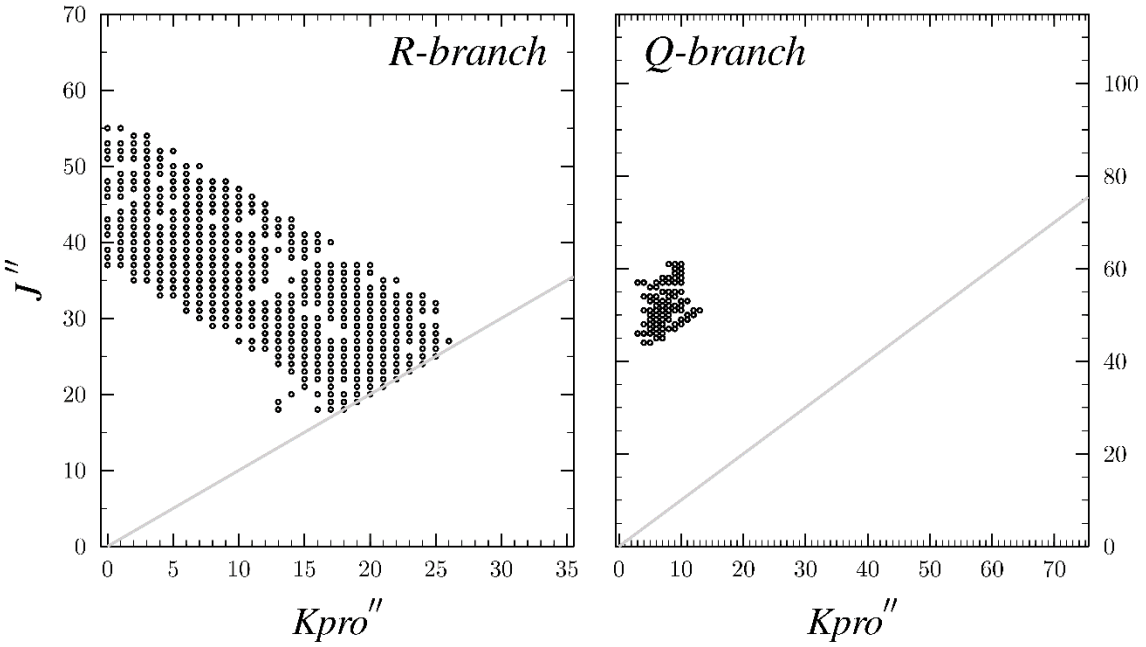
ν_9



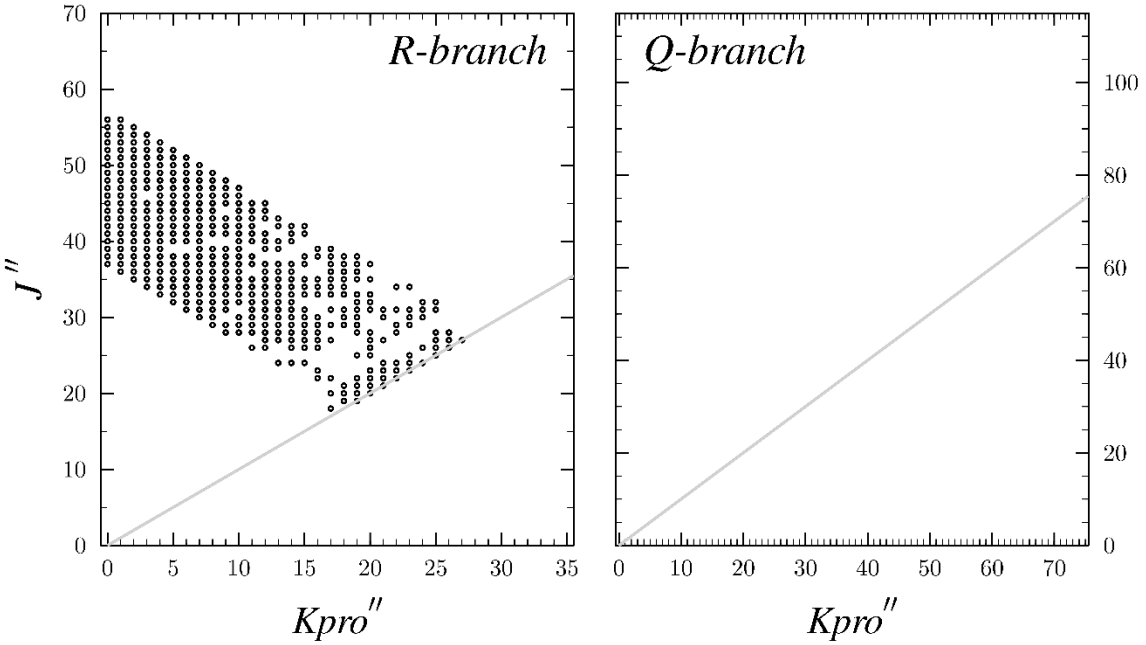
$2\nu_{16}$



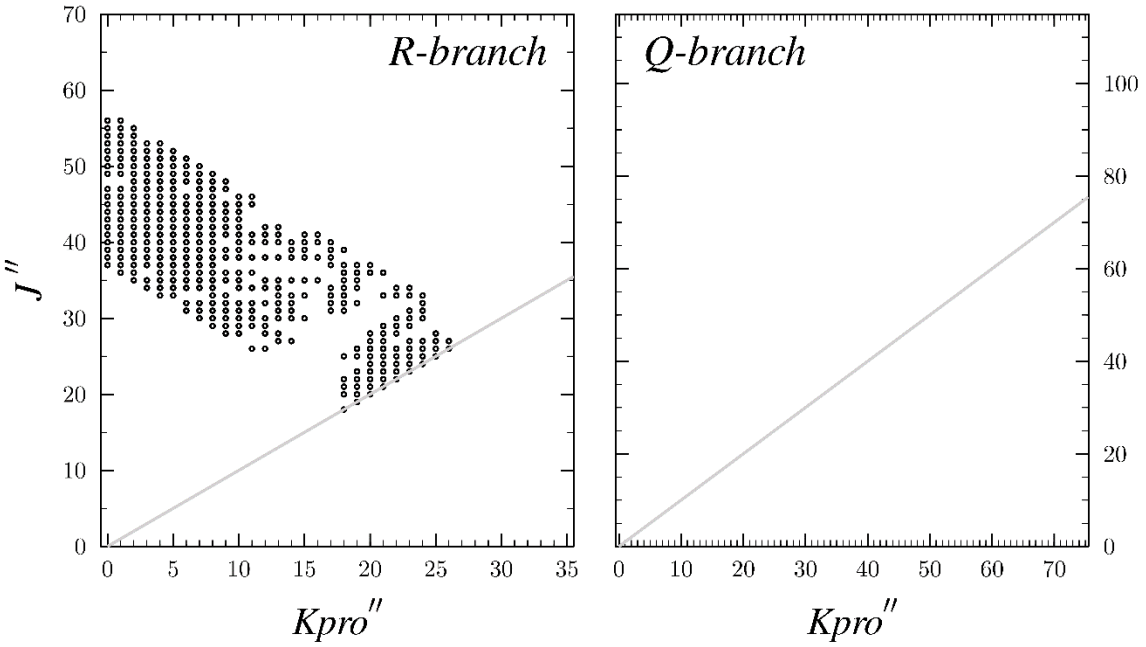
ν_{15}



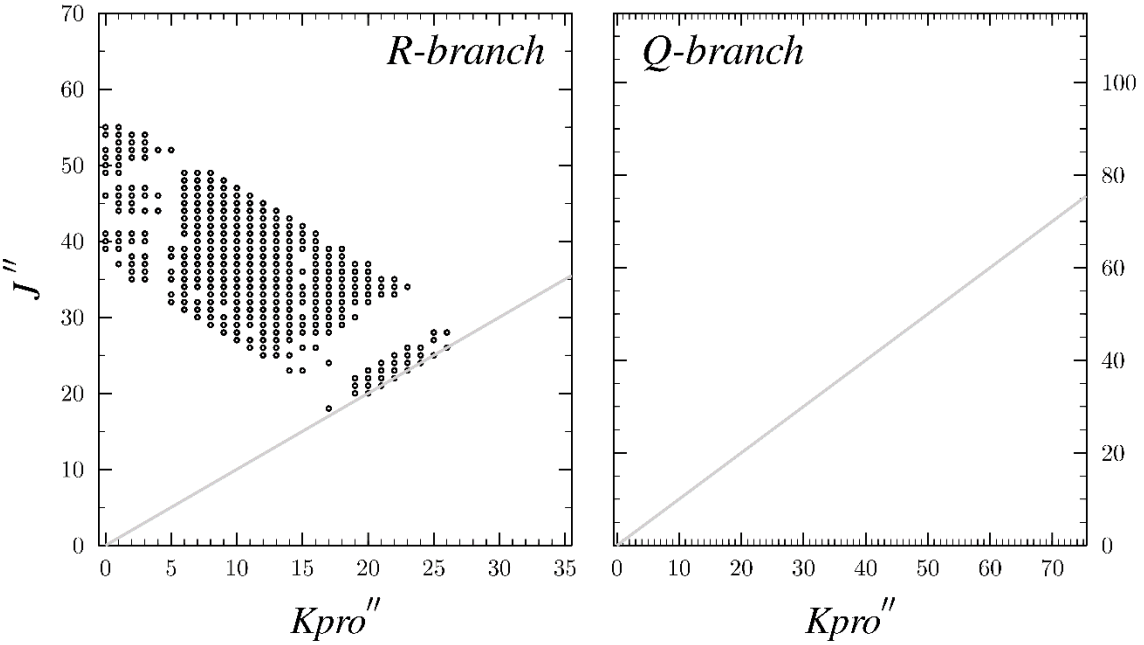
$\nu_{11} + \nu_{16}$



$2\nu_{11}$



ν_{14}


R-branch
Q-branch
J''
Kpro''
Kpro''
0
10
20
30
40
50
60
70
0
5
10
15
20
25
30
35
0
10
20
30
40
50
60
70
0
20
40
60
80
100

Pyrimidine Structure Optimization Summaries and Output Files

ZMAT

```
H
C 1   R1*
X 2   R     1  A90
C 2   R2*  3  A90  1  D180
X 4   R     2  A90  3  D0
H 4   R3*  5  A90  2  D180
N 2   R4*  1  A1*   3  D90
N 2   R4*  1  A1*   3  Dn90
C 4   R5*  6  A2*   5  D90
C 4   R5*  6  A2*   5  Dn90
H 9   R6*  4  A3*   6  D0
H 10 R6*  4  A3*   6  D0
```

CCSD(T) / cc-pCVTZ

```
R1    =         1.082700110757032
R     =         1.000000000000000
A90   =        90.000000000000000
R2    =         2.657996774591383
D180  =       180.000000000000000
D0    =         0.000000000000000
R3    =         1.080345677297168
R4    =         1.337154759664054
A1    =       116.162695889511212
D90   =        90.000000000000000
Dn90  =       -90.000000000000000
R5    =         1.390107436823451
A2    =       121.693948908432503
R6    =         1.083446388865860
A3    =       121.155411345472885
```

CCSD(T) / ANO1

R1	=	1.084007777054950
R	=	1.000000000000000
A90	=	90.000000000000000
R2	=	2.662315528542129
D180	=	180.000000000000000
D0	=	0.000000000000000
R3	=	1.081402695914324
R4	=	1.339543374212763
A1	=	116.146381016364757
D90	=	90.000000000000000
Dn90	=	-90.000000000000000
R5	=	1.392353927929510
A2	=	121.695573897121918
R6	=	1.084557868185603
A3	=	121.170547125133169

CCSD(T) / ANO2

R1	=	1.083390126868490
R	=	1.000000000000000
A90	=	90.000000000000000
R2	=	2.659392724447296
D180	=	180.000000000000000
D0	=	0.000000000000000
R3	=	1.080935182234683
R4	=	1.336757652053162
A1	=	116.244266976588193
D90	=	90.000000000000000
Dn90	=	-90.000000000000000
R5	=	1.390371613931211
A2	=	121.690180042602748
R6	=	1.083980596983981
A3	=	121.221801704642999

Electronic Contribution Summary and Output Files

CCSD(T) / ANO1

Rotational Constants without electronic part (in Mhz)

Bxx = 3082.22403, Byy = 6280.28118, Bzz = 6052.81039

---

a) using Delta I(el) = - me/Mp g Ie

Electronic Contribution to Rotational Constants (in MHz)

---

dBxx = 0.06702,dByy = -0.31940,dBzz = -0.36526

Total Rotational Constants (in MHz)

Bxx = 3082.29105, Byy = 6279.96178, Bzz = 6052.44513

---

b) using Delta B(el) = me/Mp g Be

Electronic Contribution to Rotational Constants (in MHz)

---

dBxx = 0.06702,dByy = -0.31941,dBzz = -0.36528

Total Rotational Constants (in MHz)

Bxx = 3082.29105, Byy = 6279.96176, Bzz = 6052.44511

Structure Determination Summaries Using xrefit

The least-squares fitting module of CFOUR, xrefit,

SYMBOL	BOND TO	LENGTH (ANGST)	ANGLE WRT	ANGLE (DEG)	DIHED WRT	ANGLE (DEG)
H						
C	1	R1				
X	2	R	1	A90		
C	2	R2	3	A90	1	D180
X	4	R	2	A90	3	D0
H	4	R3	5	A90	2	D180
N	2	R4	1	A1	3	D90
N	2	R4	1	A1	3	Dn90
C	4	R5	6	A2	5	D90
C	4	R5	6	A2	5	Dn90
H	9	R6	4	A3	6	D0
H	10	R6	4	A3	6	D0

Initial values for internal coordinates

Name	Value
R1	1.084008
R	1.000000
A90	90.000000
R2	2.662316
D180	180.000000
D0	0.000000
R3	1.081403
R4	1.339543
A1	116.146381
D90	90.000000
Dn90	-90.000000
R5	1.392354
A2	121.695574
R6	1.084558
A3	121.170547

R_0 Structure from determinable constants of 16 isotopologues

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0834593436
R2	2.6636328682
R3	1.0792213511
R4	1.3379420253
A1	116.2824806523
R5	1.3913949804
A2	121.6269851661
R6	1.0829929973
A3	121.4066456589

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.003032
2	0.003998
3	0.002802
4	0.002911
5	0.187708
6	0.003484
7	0.173471
8	0.002352
9	0.291535

R_e Structure from determinable constants of 16 isotopologues using coupled-cluster calculated α-corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0814523594
R2	2.6521099465
R3	1.0812716994
R4	1.3345749633
A1	116.2954985684
R5	1.3833911323
A2	121.5779382883
R6	1.0830877794
A3	121.4659934950

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.001473
2	0.001943
3	0.001363
4	0.001414
5	0.091402
6	0.001693
7	0.084880
8	0.001144
9	0.142193

R_e Structure from determinable constants of 16 isotopologues using DFT calculated α corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0806788799
R2	2.6560222851
R3	1.0787600668
R4	1.3351630321
A1	116.3774979807
R5	1.3868632817
A2	121.7236211543
R6	1.0832118196
A3	121.1741070500

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.002418
2	0.003188
3	0.002235
4	0.002326
5	0.150007
6	0.002785
7	0.138875
8	0.001879
9	0.232996

R_e Structure from determinable constants of 16 isotopologues using coupled-cluster calculated α and electron mass corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0813981692
R2	2.6520830232
R3	1.0812191149
R4	1.3345702101
A1	116.2940853867
R5	1.3833583332
A2	121.5768957611
R6	1.0830646746
A3	121.4674811966

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.000789
2	0.001041
3	0.000730
4	0.000757
5	0.048964
6	0.000907
7	0.045471
8	0.000613
9	0.076173

R_e Structure from determinable constants of 16 isotopologues using DFT calculated α and electron mass corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0807310282
R2	2.6559292668
R3	1.0788109482
R4	1.3351252621
A1	116.3748576626
R5	1.3868241742
A2	121.7209310660
R6	1.0831773528
A3	121.1767919640

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.002101
2	0.002771
3	0.001943
4	0.002022
5	0.130383
6	0.002420
7	0.120707
8	0.001633
9	0.202512

R_e Structure from A-reduced constants of 16 isotopologues using coupled-cluster calculated α and electron mass corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0813963320
R2	2.6520858642
R3	1.0812171367
R4	1.3345704877
A1	116.2941978236
R5	1.3833610981
A2	121.5769946969
R6	1.0830661488
A3	121.4670901398

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.000788
2	0.001040
3	0.000730
4	0.000757
5	0.048924
6	0.000906
7	0.045434
8	0.000612
9	0.076111

R_e Structure from S-reduced constants of 16 isotopologues using coupled-cluster calculated α and electron mass corrections

Final geometrical parameters (values in Å and deg):

Parameter	Rfinal
R1	1.0814000534
R2	2.6520807373
R3	1.0812203939
R4	1.3345703151
A1	116.2939814134
R5	1.3833556743
A2	121.5768269174
R6	1.0830637360
A3	121.4678304622

Uncertaintities in LSF parameters based on covariance matrix

parameter	uncertaintity
1	0.000790
2	0.001042
3	0.000731
4	0.000758
5	0.049015
6	0.000908
7	0.045519
8	0.000614
9	0.076253

Table S5. Summary of Structure Determinations

Parameters	$R_e^{a, d, f}$	$R_e^{a, e, f}$	$R_e^{b, d, f}$	$R_e^{c, d, f}$	$R_e^{a, d}$	$R_e^{a, e}$	R_0^{a}	R_s
$R_{C(2)-H}^{a}$ (Å)	1.0814 (7)	1.081 (2)	1.0814 (8)	1.0814 (8)	1.081 (1)	1.081 (2)	1.084 (3)	1.087
$R_{C(2)-C(5)}^{a}$ (Å)	2.6521 (10)	2.656 (3)	2.6521 (10)	2.6521 (10)	2.652 (2)	2.656 (3)	2.664 (4)	2.660
$R_{C(5)-H}^{a}$ (Å)	1.0812 (7)	1.079 (2)	1.0812 (7)	1.0812 (7)	1.081 (1)	1.079 (2)	1.079 (3)	1.082
$R_{C(2)-N(3)}^{a}$ (Å)	1.3346 (8)	1.335 (2)	1.3346 (8)	1.3346 (8)	1.335 (1)	1.335 (2)	1.338 (3)	1.338
$R_{C(4)-C(5)}^{a}$ (Å)	1.3834 (9)	1.387 (2)	1.3834 (9)	1.3834 (9)	1.383 (1)	1.387 (3)	1.391 (3)	1.393
$R_{C(4)-H}^{a}$ (Å)	1.0831 (6)	1.083 (2)	1.0831 (6)	1.0831 (6)	1.083 (1)	1.083 (2)	1.083 (2)	1.086
$R_{N(3)-C(4)}$ (Å)	1.3367 (10)	1.340 (3)	1.3366 (10)	1.3366 (10)	1.3367 (2)	1.341 (3)	1.348 (4)	1.344
$\theta_{N(1)-C(2)-H}^{a}$ (º)	116.29 (5)	116.4 (1)	116.29 (5)	116.29 (5)	116.30 (9)	116.4 (2)	116.3 (2)	116.39
$\theta_{C(4)-C(5)-H}^{a}$ (º)	121.58 (5)	121.7 (1)	121.58 (5)	121.58 (5)	121.58 (8)	121.7 (1)	121.6 (2)	121.79
$\theta_{C(5)-C(4)-H}^{a}$ (º)	121.47 (8)	121.2 (2)	121.47 (8)	121.47 (8)	121.5 (1)	121.2 (2)	121.4 (3)	120.96
$\theta_{C(2)-N(3)-C(4)}$ (º)	115.53 (5)	115.7 (1)	115.53 (5)	115.52 (5)	115.53 (9)	115.7 (2)	115.7 (2)	115.80
$\theta_{N(3)-C(4)-C(5)}$ (º)	122.35 (4)	122.4 (1)	122.35 (4)	122.35 (4)	122.35 (8)	122.4 (1)	122.3 (2)	122.38
$\theta_{C(4)-C(5)-C(6)}$ (º)	116.85 (9)	116.6 (2)	116.85 (9)	116.85 (9)	116.8 (2)	116.6 (3)	116.8 (3)	116.42

[a]Fit using determinable constants. [b]Fit using A-reduced constants. [c]Fit using S-reduced constants. [d]Corrections calculated using coupled-cluster methods. [e]Corrections calculated using DFT methods. [f]Fit includes corrections for electron mass.

Appendix B

Guanidine Computational Data

Guanidine *anti*-Conformer Structure Optimization Summaries

ZMAT

C
X 1 RX
N 1 R1* 2 A90
H 3 R2* 1 A1* 2 D90
N 1 R3* 3 A2* 4 D2*
H 5 R4* 1 A3* 3 D3*
H 5 R5* 6 A4* 3 D4*
N 1 R6* 3 A5* 4 D5*
H 8 R7* 1 A6* 3 D6*
H 8 R8* 9 A7* 3 D7*

CCSD(T) / cc-pCVDZ

R1	1.28771944
R2	1.02935737
R3	1.41236991
R4	1.02144895
R5	1.01950104
R6	1.40865836
R7	1.02047160
R8	1.02163767
A1	109.119253
A2	128.2274501
A3	110.6799046
A4	109.2140769
A5	120.3735416
A6	108.7396379
A7	109.4272273
D2	7.526577326
D3	-139.7537268
D4	-103.2527395
D5	-173.3227558
D6	-12.1772016
D7	126.9505243

CCSD(T) / ANO1

R1	1.279283411
R2	1.017747648
R3	1.400842842
R4	1.010328257
R5	1.008528629
R6	1.396423971
R7	1.009702786
R8	1.010228138
A1	110.2150494
A2	127.7920061
A3	112.7934468
A4	111.1745523
A5	120.4382859
A6	110.3000309
A7	111.5160897
D2	7.417254299
D3	-142.0389838
D4	-108.0648959
D5	-173.1933098
D6	-11.21773564
D7	131.6156834

CCSD(T) / cc-pCVTZ

R1	1.276762729
R2	1.016216646
R3	1.39851242
R4	1.00872404
R5	1.007060025
R6	1.394025293
R7	1.008157238
R8	1.008613061
A1	110.279571
A2	127.7003261
A3	112.789001
A4	111.2907202
A5	120.4772882
A6	110.3234431
A7	111.6512938
D2	7.289219611
D3	-142.5560874
D4	-108.4078853
D5	-173.4470332
D6	-11.10691619
D7	131.7003434

Guanidine *syn*-Conformer Structure Optimization Summaries

ZMAT

C
X 1 RX
N 1 R1* 2 A90
H 3 R2* 1 A1* 2 D90
N 1 R3* 3 A2* 4 D2*
H 5 R4* 1 A3* 3 D3*
H 5 R5* 6 A4* 3 D4*
N 1 R6* 3 A5* 4 D5*
H 8 R7* 1 A6* 3 D6*
H 8 R8* 9 A7* 3 D7*

CCSD(T) / cc-pCVDZ

R1	1.28858257
R2	1.029676532
R3	1.40878045
R4	1.01831662
R5	1.018127353
R6	1.409644219
R7	1.020765696
R8	1.021765578
A1	108.843553
A2	126.6642582
A3	113.8048453
A4	110.5796124
A5	119.9513339
A6	108.1895332
A7	108.9682421
D2	-1.351193985
D3	-163.997814
D4	-117.4499523
D5	171.5148081
D6	9.819595337
D7	-126.6995483

CCSD(T) / ANO1

R1	1.280216798
R2	1.0179492
R3	1.395836661
R4	1.007272434
R5	1.006636404
R6	1.396880647
R7	1.009786406
R8	1.009848468
A1	109.9777817
A2	126.4936229
A3	115.9498562
A4	112.874397
A5	119.9623802
A6	109.9236091
A7	111.322389
D2	-2.105859957
D3	-166.9455155
D4	-125.7643682
D5	171.8059301
D6	10.08078814
D7	-132.3770013

CCSD(T) / cc-pCVTZ

R1	1.277755227
R2	1.016429844
R3	1.393690064
R4	1.005729542
R5	1.005115468
R6	1.394438999
R7	1.00827606
R8	1.008260837
A1	110.0201681
A2	126.4047448
A3	115.9613921
A4	113.0040659
A5	120.0508917
A6	109.9143995
A7	111.4008857
D2	-2.306308319
D3	-167.5647108
D4	-126.2254358
D5	171.8116472
D6	9.874397474
D7	-132.2104075

Calculated Rotational and Distortion Constants

anti-Conformer

	(MHz)	(kHz)		
A_IIIr Dist constants			Ae	10617.961
ΔJ	0.0082477133	8.2477132870	Be	10181.839
ΔJK	0.0038545346	3.8545346320	Ce	5256.222
ΔK	-0.0115247061	-11.5247061300		
δJ	0.0002665194	0.2665194256	A0	10547.128
δK	0.0083238277	8.3238276740	B0	10100.278
			C0	5216.206
S_IIIr Dist Constants				
DJ	0.0084241532	8.4241532470	A^a	10617.97782
DJK	-0.0125833459	-12.5833458900	B^a	10181.8202
DK	0.0047367344	4.7367344340	C^a	5256.22303
dJ	-0.0002665194	-0.2665194256		
dK	0.0000882200	0.0882199803	A^s	10617.96152
			B^s	10181.8372
A_Ir Dist constants			C^s	5256.22215
ΔJ	0.0041461081	4.1461081124		
ΔJK	0.0161593502	16.1593501558		
ΔK	-0.0115247061	-11.5247061300		
δJ	0.0017842832	1.7842831617		
δK	-0.0055355366	-5.5355365734		
S_Ir Dist Constants				
DJ	0.0047673669	4.7673669446		
DJK	-0.0016129870	-1.6129869828		
DK	0.0047367344	4.7367344340		
dJ	0.0020949126	2.0949125768		
dK	0.0025858786	2.5858786161		

syn-Conformer

	In MHz	In kHz		In MHz
A_IIIr Dist constants			Ae	10545.132
ΔJ	0.00857793	8.57793104	Be	10244.511
ΔJK	0.00408609	4.08609433	Ce	5267.377
ΔK	-0.01199288	-11.99287522		
δJ	-0.00002041	-0.02041437	A0	10473.727
δK	0.03645472	36.45472196	B0	10160.49
			C0	5218.74
S_IIIr Dist Constants				
DJ	0.00911227	9.11226515	A^a	10545.20356
DJK	-0.01519888	-15.19887983	B^a	10244.43526
DK	0.00675776	6.75776483	C^a	5267.38082
dJ	0.00002041	0.02041437		
dK	0.00026717	0.26716705	A^s	10545.13172
			B^s	10244.50924
A_Ir Dist constants			C^s	5267.37815
ΔJ	0.00464495	4.64495497		
ΔJK	0.01588502	15.88502256		
ΔK	-0.01199288	-11.99287522		
δJ	0.00198690	1.98690241		
δK	-0.00415581	-4.15581350		
S_Ir Dist Constants				
DJ	0.00487129	4.87129328		
DJK	-0.00247596	-2.47596423		
DK	0.00675776	6.75776483		
dJ	0.00210007	2.10007156		
dK	0.00359911	3.59910656		

Alpha Calculation Summary

anti-Conformer

	Alpha contributions		(in MHz)
	A	B	C
	8.13252	12.98034	3.10744
	11.89063	10.98634	1.72144
	9.50289	11.4927	2.92175
	9.70361	11.10985	2.46355
	9.62829	6.32569	3.13019
	25.32106	27.8567	10.26526
	-0.6828	-0.43735	2.68352
	-3.40095	4.55984	3.35447
	21.9594	19.94856	21.32203
	2.36331	-3.78527	3.2493
	0.08927	-17.03193	6.92115
	2.59059	9.88257	7.89109
	15.2434	26.83361	11.66389
	14.72704	13.4392	0.89849
	9.74789	3.26295	1.8902
	0.01425	-2.63848	1.20145
	-12.28228	-7.00969	-1.74348
	-6.17435	-7.95139	-14.41296
	-24.92348	-6.10744	22.82094
	39.39235	23.13523	2.94419
	9.96822	31.1888	2.97878
1/2 Sum	71.40543	84.020415	48.636345

syn-Conformer

	Alpha contributions		(in MHz)
	A	B	C
	6.66008	21.59762	4.28687
	13.4165	6.27827	2.47828
	8.32454	17.77005	4.25314
	11.37272	8.52244	2.62187
	10.18136	6.67045	2.85978
	25.99296	28.17282	8.70198
	-1.95359	-2.33677	2.45635
	-2.21908	15.59935	2.94442
	23.10155	33.3036	17.69247
	2.23802	-1.01029	2.07806
	1.04574	-0.13723	3.77468
	-1.19185	10.11757	6.6875
	15.92358	25.98941	11.10427
	13.63923	7.00905	1.81026
	7.97056	7.99404	0.12461
	-3.91097	-6.61154	-2.43591
	-5.80877	-16.84335	-23.90401
	-15.59834	0.23823	28.73605
	4.82648	-32.5423	-1.35619
	24.53701	24.52087	0.85608
	3.11733	8.82001	4.26085
1/2 sum	70.83253	81.56115	40.015705