THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

Cite This: J. Phys. Chem. Lett. 2019, 10, 2408−2413

pubs.acs.org/JPCL

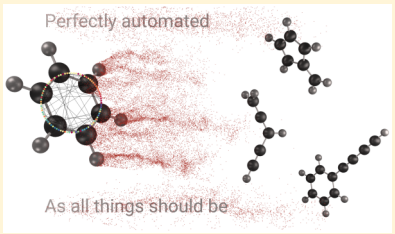
# Study of Benzene Fragmentation, Isomerization, and Growth Using Microwave Spectroscopy

Kin Long Kelvin Lee* and Michael McCarthy

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, United States

Supporting Information

**ABSTRACT:** Using a combination of broadband and cavity Fourier transform microwave spectroscopies, and newly developed analysis and assignment tools, the discharge products of benzene have been extensively studied in the 2−18 GHz frequency range. More than 450 spectral features with intensities greater than 6$\sigma$ of the noise RMS were identified, of which of roughly four-fifths (82%) constituting 90% of the total spectral intensity were assigned to 38 species previously detected in the radio band, and nine entirely new hydrocarbon molecules were identified. The new species include both branched and chain fragments of benzene, high energy $C_6H_6$ isomers, and larger molecules such as phenyldiacetylene and isomers of fulvenallene; taken together they account for roughly half of the number of observed transitions and 51% of the spectral line intensity. Transitions from vibrationally excited states of several molecules were also identified in the course of this investigation. A key aspect of the present analysis was implementation of a rapid and efficient workflow to assign spectral features from known molecules and to identify line progressions by pattern recognition techniques.






Benzene is of central importance in combustion chemistry because formation of the first aromatic ring is thought to be the rate-limiting step in polycyclic aromatic hydrocarbon (PAH) growth.[1] Once formed, still larger rings are hypothesized to form in a straightforward manner, although the precise mechanism remains a subject of considerable debate, with no fewer than four schemes proposed to date.[2−5] The recent astronomical detection of cyanobenzene (*c*-$C_6H_5CN$)[6] has led to renewed interest in the formation pathways that lead to this prototypical aromatic hydrocarbon, because this discovery implies benzene is formed at remarkably low density (<10 K) and pressure ($10^4$−$10^5$ particles/cm$^3$), since subsequent reaction with the CN radical to form cyanobenzene is known to be both exothermic and barrierless.[7−10] For these reasons, the fragmentation and isomerization of benzene are highly topical, as are intermediates formed en route to larger rings.

Rotational spectroscopy is a powerful tool for analyzing complex mixtures of polar species because its intrinsic high spectral resolution coupled with the highly diagnostic nature of rotational transitions enables isomers and isotopic species to be readily differentiated. In many instances, it provides complementary information to mass spectrometry. Rotational frequencies sensitively depend on rotational constants which in turn are inversely proportional to the principal moments of inertia and, consequently, are unique for each arrangement of atoms in a molecule. A number of recent microwave studies have sought to characterize mixtures with varying degrees of complexity, either in an automated or in a semiautomated fashion.[11−15]

In this Letter, we describe a comprehensive reaction screening study in which benzene was subjected to an electrical discharge and analyzed by a combination of broadband and cavity-enhanced Fourier transform (FT) microwave spectroscopies. Although a modest (∼25%) fraction of the observed features were readily assigned to molecules whose rotational spectra have been previously measured, the remaining features—responsible for more than half of the overall spectral line intensity (i.e., defined as the sum of individual peak intensities)—defied identification. To investigate the compounds responsible for these features, automated microwave double resonance spectroscopy (AMDOR),[11] microwave spectral taxonomy,[15] and pattern recognition methods were extensively used. Several approaches were then adopted to uniquely determine the stoichiometry and structure of each compound on the basis of the derived rotational constants. Benzene fragmentation and isomerization, as well as pathways that may lead to larger aromatic rings, are discussed in the context of the newly discovered molecules. An improved workflow for spectral analysis emphasizing speed, efficiency, and accuracy is also highlighted.

In the present study, a sample of benzene heavily diluted in an inert buffer gas was subjected to an electrical discharge just prior to expansion into a vacuum chamber using a pulsed nozzle source. Collisions with electrons and metastable rare gas atoms or their ions have enough energy to break chemical bonds, and owing predominately to fast neutral−neutral reactions that subsequently take place, a rich broth of products are quickly formed. Because of the large pressure differential behind the nozzle relative to the chamber, the gas mixture undergoes supersonic expansion at the end of the discharge

**Received:** February 28, 2019
**Accepted:** April 25, 2019
**Published:** April 25, 2019




DOI: 10.1021/acs.jpclett.9b00586
J. Phys. Chem. Lett. 2019, 10, 2408−2413

nozzle assembly, and the rotational temperatures reaches about 2 K near the center of the chamber where the microwave radiation intersects the gas. In the experiments described here, the experimental conditions were optimized to produce fulvene ($C_5H_4$=$CH_2$), a higher energy isomer of benzene, by monitoring one of its centimeter-wave lines[16] using a cavity-enhanced FT microwave spectrometer operating between 5 and 26 GHz. Subsequently, a broadband spectrum in the 2−18 GHz frequency range was recorded using a chirped pulse spectrometer in two segments (2−8 and 8−18 GHz). Follow-up cavity measurements, including taxonomy and double resonance experiments, were then performed on unassigned spectral features in the broadband spectrum. Switching between the two spectrometers is very rapid because both are located in the same vacuum chamber but lie along perpendicular axes relative to one another.

During and following acquisition of the broadband spectrum, a peak finding algorithm identifies and collates in real time the frequency and intensities of spectral features above a user-defined intensity threshold. In this analysis, a threshold of six times the noise RMS was chosen because it was found by trial and error to prevent erroneous detections, while flagging the largest number of real spectral features for subsequent analysis. A total of 452 features were identified in this way, which corresponds to a line density of approximately 40 lines per GHz, or a filling factor of about 1%, i.e., defined here as the number of active or "bright" resolution elements within the 2−18 GHz spectrum.

With a well-defined set of spectral frequencies in hand, the process of identifying the carrier of each followed (Figure 1).

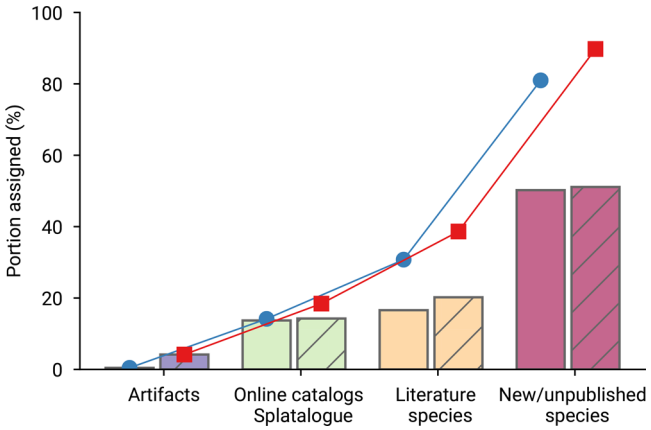



**Figure 1.** Fraction of lines and spectral intensity assigned at each stage of the assignment pipeline, with solid bars representing the number of lines and striped bars the sum of the spectral line peak intensities. Cumulative percentages are indicated by blue circles for lines and red squares for intensity.

Instrumental artifacts and any low-level contaminants were first removed from the line list. These are readily identified by collecting a chirp spectrum in the absence of the gas sample and a second one with precursor present but without application of the electrical discharge. Only two features were observed in these two spectra, both arising from clock spurs (at 8 and 16 GHz), indicating that contaminants (i) do not possess rotational transitions within the frequency range of the measurement, (ii) are nonpolar, and/or (iii) are present at trace levels. On the assumption that the remaining features are molecular in origin, an assignment pipeline written in Python uses the "astroquery" package to query online spectral line databases such as the Cologne Database for Molecular Spectroscopy (CDMS)[17] and the Jet Propulsion Laboratories (JPL).[18] The frequency of each feature is separately and sequentially queried through the aggregate Splatalogue search engine, which only returns transitions that lie within a user-specified tolerance (0.2 MHz in the present case).

Automated assignment of each feature is carried out by systematically rejecting unlikely transitions based on two criteria: those arising from species which have an elemental composition grossly at odds with that of the precursor (e.g., sulfur-, phosphorus-bearing) and transitions originating from unreasonably high lower rotational state energies. The former discriminates on the basis of chemical composition, while the latter is justified given the extremely low rotational temperature of the gas expansion. The remaining transitions are subsequently ranked with a normalized weighting factor that depends on the absolute difference between the measured and predicted frequency $\Delta\nu$ and the theoretical intensity of the transition $I$:[18]

$$w_n = \frac{I}{|\Delta\nu|} \tag{1}$$

$$W_n = \frac{w_n}{\sum_n w_n} \tag{2}$$

where $w_n$ and $W_n$ are unnormalized and normalized weights, respectively, for each candidate line $n$. The most likely assignments (yielding the largest $W_n$) are therefore those that lie very close to the measured frequency and are predicted to be intense. In instances where multiple transitions of the same molecule may be present, the strongest transition is chosen. Using this automated procedure, 62 features in the line list, or about 14% of the total, were assigned to closed-shell hydrocarbons such as *o*-benzyne (*o*-$C_6H_4$) and methyl polyynes (e.g., $CH_3$—C≡C—C≡CH and $CH_3$—C≡C—C≡C—C≡CH), radicals (e.g., $C_nH$, where $n$ = 3−6), and cumulene carbenes (e.g., $H_2C_5$). Several common contaminants were also identified in this step: they include cyanides such as cyanoacetylene ($HC_3N$) and benzonitrile (*c*-$C_6H_5CN$) and oxygen-bearing species such as propynal (HC≡C—CHO). Lines of these species are present presumably because of residual contamination from atmospheric gases (i.e., $N_2$, $O_2$, $H_2O$), which either are nonpolar or do not possess strong transitions in the frequency range of the chirped spectrum. Nevertheless the contaminant features account for only a small fraction of the total number of lines (19 or 4.2%) and very little (2.2%) of the total spectral intensity in the chirped spectrum.

Because online databases focus primarily on astronomical species, other hydrocarbons previously measured by microwave spectroscopy at high spectral resolution are not commonly included. In these instances, published constants were used to generate catalogs[19] for species that might be present in our discharge expansion. An additional 62 features were assigned in this way, including lines of two $C_6H_4$ isomers, *cis*-hex-3-ene-1,5-diyne (HC≡C$(CH)_2$C≡CH) and vinyl diacetylene ($H_2C$=CH—C≡C—C≡CH), fulvene, and phenylacetylene ($C_6H_5$—C≡CH).

Somewhat remarkably, approximately 340 features—or about 75% of the total—defied obvious assignment. In an attempt to assign the remaining features and ultimately ascertain their identity, microwave spectral taxonomy



DOI: 10.1021/acs.jpclett.9b00586
J. Phys. Chem. Lett. 2019, 10, 2408−2413

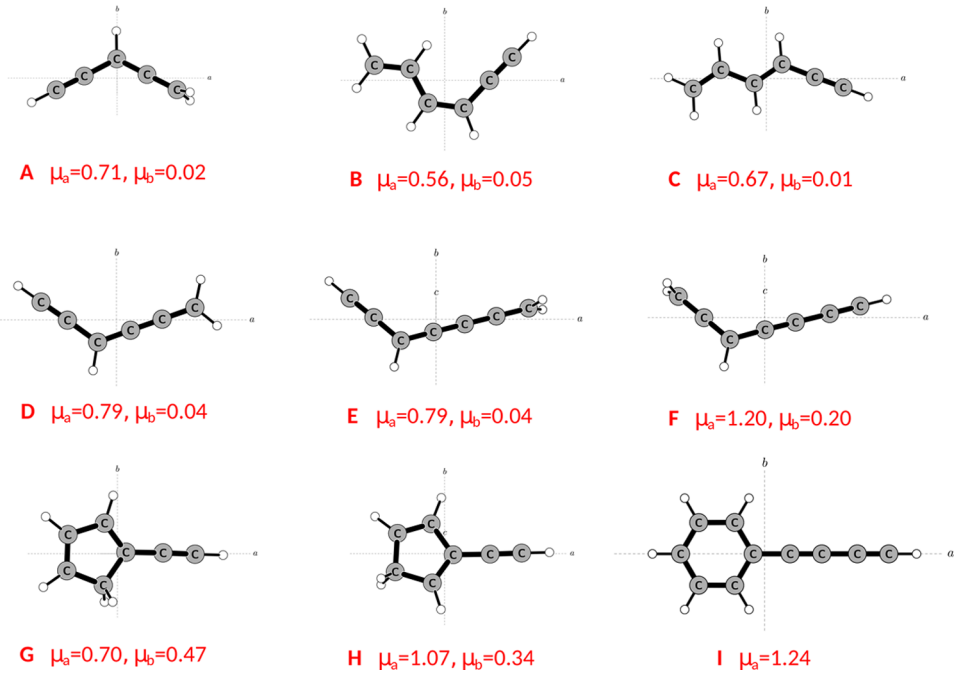



**Figure 2.** Structures of nine new hydrocarbon molecules (labeled A–I) identified in a benzene discharge. Dipole moments (in Debye), calculated at the M05/6-31G(d) level of theory, are shown in red. The principal axes are overlaid for each structure.

(MST)[15] categorization tests, specifically elemental composition, dipole moment, and susceptibility to an external magnetic field, were performed. These assays, however, yielded very little new information. First, almost by default, the carriers of all the lines must contain both elements, since a hydrocarbon precursor (benzene) was used, and pure carbon or hydrogen clusters are known to form highly symmetrical (nonpolar) structures. Not surprisingly, our elemental tests confirmed this prediction. Second, with few exceptions, hydrocarbons typically have modest, but similar, dipole moments, since the two elements have comparable electronegativities. Third, magnet tests were equally unenlightening, since the intensity of the unassigned features were uniformly insensitive to an applied field, implying they all arise from closed-shell molecules. As a consequence, AMDOR experiments played a key role in disentangling the spectrum, in which binary double resonance tests are used to identify transitions that share a common energy level and, hence, arise from the same carrier. Because the number of pair combinations scales as $n(n-1)/2$, where $n$ is the number of lines,[15] however, nearly 23 000 binary tests, equivalent to at least 63 h of acquisition time, would be needed to perform an exhaustive cross-correlation.

To perform the AMDOR analysis more efficiently, the line list was analyzed using pattern recognition algorithms to identify lines that might arise from prolate asymmetric tops, since their rotational lines are spaced at nearly harmonic frequency intervals and are, thus, well described by a simple effective linear molecule Hamiltonian with two parameters ($B_{\mathrm{eff}}$ and $D_{\mathrm{eff}}$). From this analysis, a much smaller number of lines and pair combinations (~3800) warranted further testing by AMDOR. The number of pair combinations was reduced even further (~700) by affinity propagation (AP) cluster analysis[20] as implemented in the Scikit-Learn library,[21] which groups sets of transitions that have similar $B_{\mathrm{eff}}$ values—the rationale being that line sequences with similar rotational constants likely arise from different $K_a$ states and/or vibrationally excited states of the same carrier. Since the AP cluster model does not require the number of clusters to be known *a priori*, the model is tuned using a Euclidean mean silhouette metric ($s$) for similarity[22] to compromise between too many cluster centers and too few: the former confers no experimental simplification, while the latter forces sets to cluster even when the constants differ by values that have little or no experimental precedent. The metric classifies progressions in their clusters as good matches ($s \approx 1$) or highly dissimilar from the rest of the cluster ($s \approx -1$). Our procedure requires iteratively minimizing the number of clusters, while also minimizing ill-fitting progression and cluster matches (i.e., low $s$). Ultimately, more than a dozen sequences of lines linked by DR were identified using this combined approach. By trial and error, it was then possible to assign sequences with similar $B_{\mathrm{eff}}$ values to different $K_a$ ladders of the same carrier and derive precise sets of rotational constants. In many cases, additional lines from higher $K_a$ levels were subsequently assigned in the chirped spectrum or at higher frequency using the cavity instrument. As the number of unassigned lines began to dwindle, exhaustive DR measurements were routinely performed on the remaining features, yielding still more sequences of linked transitions. By this iterative procedure, ultimately about 75% of the approximately 340 initially unassigned features were found to arise from more than a dozen carriers or their variants.

Uniquely identifying the molecule associated with each set of rotational constants is the final and presently the most challenging step in the analysis. Electronic structure theory was used extensively to predict the equilibrium rotational constants of molecules closely related to those already identified in the spectrum. For example, since phenylacetylene is present, its derivatives (e.g., longer chains) and isomers may be present as well. Since one unassigned carrier has a nearly identical $A$ rotational constant to phenylacetylene, has a small inertial defect (i.e., planar), and does not possess a $b$-type spectrum but has smaller $B$ and $C$ constants, a promising candidate is



DOI: 10.1021/acs.jpclett.9b00586
J. Phys. Chem. Lett. 2019, 10, 2408–2413

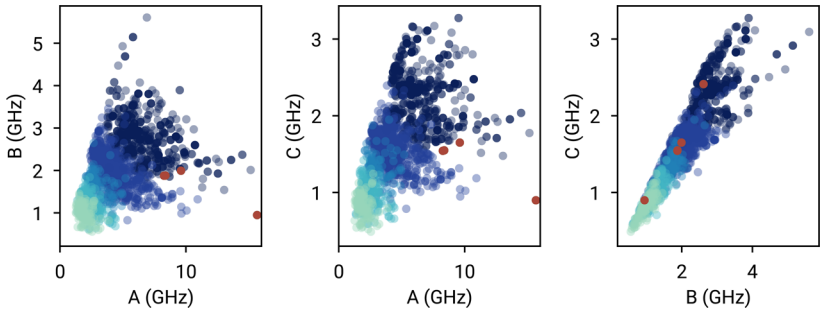



**Figure 3.** Pair plots of *A*, *B*, and *C* rotational constants for species generated with a graph-based approach (green to blue) compared to those for four of the newly identified molecules shown in Figure 2 (red). 1786 different structures were optimized at the M05/STO-3G level and vary in size and degree of saturation. The color scale ranges from 6 carbon atoms (dark blue) to 10 carbon atoms (light green).

phenyldiacetylene (*c*-$C_6H_5C_4H$, structure **I**, Figure 2). The extremely close agreement of the theoretical rotational constants with those derived from the experimental data lend considerable confidence to this identification. The same approach and arguments were used to identify a number of other new hydrocarbon chains, including both higher energy chain isomers of benzene: *E*- and *Z*-hexa-1,3-dien-5-yne (Structures **B** and **C**, Figure 2) and 1- and 2-ethynylcyclopentadiene (Structures **G** and **H**), two higher energy isomers of fulvenallene ($C_5H_4$═C═$CH_2$). New molecules identified in the microwave analysis are shown in Figure 2, and a comparison of the experimental and theoretical rotational constants is given in the Supporting Information. Additionally, this work yielded several new vibrational states of methyldiacetylene and phenylacetylene, and improved spectroscopic constants for species such as fulvene and fulvenallene, among others.

It is worth noting that more than one-half of the new species are larger than the precursor benzene, a finding that suggests carbon insertion reactions[23] occur facilely under our experimental conditions. It is intriguing, for example, that phenyldiacetylene is the largest molecule detected so far in the benzene discharge but also the most conspicuous. Crossed molecular beam studies[24] previously concluded that a viable pathway to form this ring chain was via the reaction between phenyl radical and diacetylene. In light of these results, isotopic studies using samples enriched in $^{13}C$ may provide insight into the formation pathway of this and other larger molecules.

With the exception of the benzene isomers, a fairly close correlation was found between isomer abundance and relative stability as calculated theoretically. For this reason, detection of other higher-energy isomers would appear simply to be an issue of detection sensitivity. Benzocyclopropene, a higher energy isomer of fulvenallene,[25] for example, might be detectable by exploiting the higher sensitivity of a cavity spectrometer, since rotational lines of $C_7H_6$ isomers with comparable stability (i.e., Structures **G** and **H** in Figure 2) have already been observed with good signal-to-noise ratios.

While the present approach has ultimately been successful in identifying a number of new molecules, an inherent weakness is its over-reliance on chemical intuition to determine viable candidates from a set of rotational constants. Toward the end of our analysis, we also experimented with a molecular structure generator based on graph theory to rapidly generate many possible candidates. The graph representation of molecules is a well-documented approach to structure generation, with both open-source (e.g., Open Molecule Generator[26]) and commercial options available (e.g., MOLGEN[27,28]). The algorithm uses routines from the NetworkX library[29] and begins with every non-hydrogen atom connected by at least one bond to prevent fragments (i.e., disconnected graphs). Each heavy atom is assigned a valency between one and its maximum uncharged value (e.g., four for carbon, three for nitrogen), and bonds between other heavy atoms are formed with preference given to low valency atoms. Finally, hydrogen atoms are added to the structure according to the remaining valencies.

The resulting structures therefore have varying degrees of saturation and connectivity and, to prevent redundancy, can be checked for isomorphism with previously generated molecular graphs. The three-dimensional Cartesian coordinates are then generated and optimized with the Fruchterman and Reingold method[30] and are subsequently refined with low-level electronic structure methods with a small basis. In the present tests, hydrocarbon species from 6 to 10 carbon atoms were generated, ranging from the smallest, $C_6H_4$, to the largest, $C_{10}H_{22}$. Of these, chemically intuitive structures such as the $C_6H_6$ isomers fulvene and cyclohexa-1,2,4-triene arose. As shown in Figure 3, the resulting *A*, *B*, and *C* constants can be used to help elucidate the identity of an unassigned carrier by establishing the structural diversity possible for a given number of heavy atoms.

With a modestly sized set of structures, those yielding rotational constants closest to that of the unassigned molecule can be subsequently refined with more sophisticated methods, which should reduce the number of possible candidates even further. This approach, while still in its infancy, shows promise as way of identifying molecules via their rotational constants without the need for chemical intuition. When coupled with the information gleaned from assay tests (elemental composition, dipole moments, and susceptibility to magnetic fields), it might prove to be a robust method for identifying completely unknown molecules and, if so, one that further illustrates the power of microwave spectroscopy as a tool for complex mixture analysis. In the current work, we have successfully disentangled and identified the fragmentation, isomerization, and larger products when benzene was subjected to an electrical discharge. A natural extension would be an analogous study where benzene is mixed with $O_2$ or $N_2$ to yield species of interest to combustion and astrochemistry.



DOI: 10.1021/acs.jpclett.9b00586
*J. Phys. Chem. Lett.* 2019, 10, 2408−2413

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.9b00586.

Complete list of molecules and number of transitions identified in the 2−18 GHz chirped spectrum, along with a comparison between the experimental and theoretical rotational constants for the structures shown in Figure 2 (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*(K.L.K.L.) E-mail: kinlee@cfa.harvard.edu.

### ORCID

Kin Long Kelvin Lee: 0000-0002-1903-9242

Michael McCarthy: 0000-0001-9142-0008


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank M.-A. Martin-Drumel for helpful comments and acknowledge NSF grants AST-1615847 and CHE-1566266 and NASA grants NNX13AE59G and 80NSSC18K0396 for financial support.

## ■ REFERENCES


(1) Cherchneff, I.; Barker, J. R.; Tielens, A. G. G. M. Polycyclic Aromatic Hydrocarbon Formation in Carbon-Rich Stellar Envelopes. *Astrophys. J.* **1992**, *401*, 269−287.

(2) Constantinidis, P.; Schmitt, H.-C.; Fischer, I.; Yan, B.; Rijs, A. M. Formation of Polycyclic Aromatic hydrocarbons from Bimolecular Reactions of Phenyl Radicals at High Temperatures. *Phys. Chem. Chem. Phys.* **2015**, *17*, 29064−29071.

(3) Shukla, B.; Miyoshi, A.; Koshi, M. Role of Methyl Radicals in the Growth of PAHs. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 534−544.

(4) Frenklach, M.; Clary, D. W.; Gardiner, W. C.; Stein, S. E. Detailed Kinetic Modeling of Soot Formation in Shock-Tube Pyrolysis of Acetylene. *Symp. (Int.) Combust., [Proc.]* **1985**, *20*, 887−901 Twentieth Symposium (International) on Combustion .

(5) Johansson, K. O.; Head-Gordon, M. P.; Schrader, P. E.; Wilson, K. R.; Michelsen, H. A. Resonance-Stabilized Hydrocarbon-Radical Chain Reactions May Explain Soot Inception and Growth. *Science* **2018**, *361*, 997−1000.

(6) McGuire, B. A.; Burkhardt, A. M.; Kalenskii, S.; Shingledecker, C. N.; Remijan, A. J.; Herbst, E.; McCarthy, M. C. Detection of the Aromatic Molecule Benzonitrile (c-C6H5CN) in the Interstellar Medium. *Science* **2018**, *359*, 202−205.

(7) Balucani, N.; Asvany, O.; Chang, A. H. H.; Lin, S. H.; Lee, Y. T.; Kaiser, R. I.; Bettinger, H. F.; Schleyer, P. v. R.; Schaefer, H. F., III Crossed Beam Reaction of Cyano Radicals with Hydrocarbon Molecules. I. Chemical Dynamics of Cyanobenzene ($C_6H_5CN$; $X^1A_1$) and Perdeutero Cyanobenzene ($C_6D_5CN$; $X^1A_1$) Formation from Reaction of CN ($X^2\Sigma^+$) with Benzene $C_6H_6$ ($X^1A_{1g}$), and $d_6$-benzene $C_6D_6$ ($X^1A_{1g}$). *J. Chem. Phys.* **1999**, *111*, 7457−7471.

(8) Woon, D. E. Modeling Chemical Growth Processes in Titan's Atmosphere: 1. Theoretical Rates for Reactions Between Benzene and the Ethynyl ($C_2H$) and Cyano (CN) Radicals at Low Temperature and Pressure. *Chem. Phys.* **2006**, *331*, 67−76.

(9) Trevitt, A. J.; Goulay, F.; Taatjes, C. A.; Osborn, D. L.; Leone, S. R. Reactions of the CN Radical with Benzene and Toluene: Product Detection and Low-Temperature Kinetics. *J. Phys. Chem. A* **2010**, *114*, 1749−1755.

(10) Lee, K. L. K.; McGuire, B. A.; McCarthy, M. C. Gas-Phase Synthetic Pathways to Benzene and Benzonitrile: a Combined Microwave and Thermochemical Investigation. *Phys. Chem. Chem. Phys.* **2019**, *21*, 2946.

(11) Martin-Drumel, M.-A.; McCarthy, M. C.; Patterson, D.; McGuire, B. A.; Crabtree, K. N. Automated Microwave Double Resonance Spectroscopy: A Tool to Identify and Characterize Chemical Compounds. *J. Chem. Phys.* **2016**, *144*, 124202.

(12) Seifert, N. A.; Finneran, I. A.; Perez, C.; Zaleski, D. P.; Neill, J. L.; Steber, A. L.; Suenram, R. D.; Lesarri, A.; Shipman, S. T.; Pate, B. H. AUTOFIT, an Automated Fitting Tool for Broadband Rotational Spectra, and Applications to 1-Hexanal. *J. Mol. Spectrosc.* **2015**, *312*, 13−21.

(13) McCarthy, M. C.; Zou, L.; Martin-Drumel, M.-A. To Kink or Not: A Search for Long-Chain Cumulenones using Microwave Spectral Taxonomy. *J. Chem. Phys.* **2017**, *146*, 154301.

(14) Zaleski, D. P.; Prozument, K. Automated Assignment of Rotational Spectra Using Artificial Neural Networks. *J. Chem. Phys.* **2018**, *149*, 104106.

(15) Crabtree, K. N.; Martin-Drumel, M.-A.; Brown, G. G.; Gaster, S. A.; Hall, T. M.; McCarthy, M. C. Microwave Spectral Taxonomy: A Semi-Automated Combination of Chirped-Pulse and Cavity Fourier-Transform Microwave Spectroscopy. *J. Chem. Phys.* **2016**, *144*, 124201.

(16) Baron, P. A.; Brown, R. D.; Burden, F. R.; Domaille, P. J.; Kent, J. E. The Microwave Spectrum and Structure of Fulvene. *J. Mol. Spectrosc.* **1972**, *43*, 401−410.

(17) Müller, H. S. P.; Schlöder, F.; Stutzki, J.; Winnewisser, G. The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists. *J. Mol. Struct.* **2005**, *742*, 215−227.

(18) Pickett, H. M.; Poynter, R. L.; Cohen, E. A.; Delitsky, M. L.; Pearson, J. C.; Müller, H. S. P. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transfer* **1998**, *60*, 883−890.

(19) Pickett, H. M. The Fitting and Prediction of Vibration-Rotation Spectra with Spin Interactions. *J. Mol. Spectrosc.* **1991**, *148*, 371−377.

(20) Frey, B. J.; Dueck, D. Clustering by passing Messages Between Data Points. *Science* **2007**, *315*, 972−976.

(21) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(22) Rousseeuw, P. J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53−65.

(23) da Silva, G. Reaction of Benzene with Atomic Carbon: Pathways to Fulvenallene and the Fulvenallenyl Radical in Extraterrestrial Atmospheres and the Interstellar Medium. *J. Phys. Chem. A* **2014**, *118*, 3967−3972.

(24) Parker, D. S. N.; Zhang, F.; Kim, Y. S.; Kaiser, R. I.; Landera, A.; Mebel, A. M. On the Formation of Phenyldiacetylene ($C_6H_5CCCCH$) and D5-Phenyldiacetylene ($C_6D_5CCCCH$) Studied Under Single Collision Conditions. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2997−3003.

(25) Wong, M. W.; Wentrup, C. Interconversions of Phenylcarbene, Cycloheptatetraene, Fulvenallene, and Benzocyclopropene. A Theoretical Study of the $C_7H_6$ Energy Surface. *J. Org. Chem.* **1996**, *61*, 7022−7029.

(26) Peironcely, J. E.; Rojas-Chertó, M.; Fichera, D.; Reijmers, T.; Coulier, L.; Faulon, J.-L.; Hankemeier, T. OMG: Open Molecule Generator. *J. Cheminf.* **2012**, *4*, 21.

(27) Meringer, M.; Schymanski, E. Small Molecule Identification with MOLGEN and Mass Spectrometry. *Metabolites* **2013**, *3*, 440−462.

(28) Gugisch, R.; Kerber, A.; Kohnert, A.; Laue, R.; Meringer, M.; Rcker, C.; Wassermann, A. MOLGEN 5.0, A Molecular Structure Generator. In *Advances in Mathematical Chemistry and Applications*; Bentham Books: 2014, Vol. *1*, pp 113−138.




DOI: 10.1021/acs.jpclett.9b00586
*J. Phys. Chem. Lett.* 2019, 10, 2408−2413

(29) Hagberg, A. A.; Schult, D. A.; Swart, P. J. Exploring Network Structure, Dynamics, and Function using NetworkX. *Proceedings of the 7th Python in Science Conference (SciPy2008)*; 2018; pp 11−15.
(30) Fruchterman, T. M. J.; Reingold, E. M. Graph Drawing by Force-Directed Placement. *Software: Practice and Experience* **1991**, *21*, 1129−1164.



DOI: 10.1021/acs.jpclett.9b00586
*J. Phys. Chem. Lett.* 2019, 10, 2408−2413