nature astronomy

ARTICLES

https://doi.org/10.1038/s41550-020-01213-y

Interstellar detection of the highly polar five-membered ring cyanocyclopentadiene

Michael C. McCarthy[1]✉, Kin Long Kelvin Lee[1], Ryan A. Loomis[2], Andrew M. Burkhardt[1], Christopher N. Shingledecker[3,4,5], Steven B. Charnley[6], Martin A. Cordiner[6,7], Eric Herbst[8,9], Sergei Kalenskii[10], Eric R. Willis[8], Ci Xue[8], Anthony J. Remijan[2] and Brett A. McGuire[1,2,11]✉

Much like six-membered rings, five-membered rings are ubiquitous in organic chemistry, frequently serving as the building blocks for larger molecules, including many of biochemical importance. From a combination of laboratory rotational spectroscopy and a sensitive spectral line survey in the radio band toward the starless cloud core TMC-1, we report the astronomical detection of 1-cyano-1,3-cyclopentadiene (1-cyano-CPD, c-C_5H_5CN), a highly polar, cyano derivative of cyclopentadiene. The derived abundance of 1-cyano-CPD is far greater than predicted from astrochemical models that well reproduce the abundance of many carbon chains. This finding implies that either an important production mechanism or a large reservoir of aromatic material may need to be considered. The apparent absence of its closely related isomer, 2-cyano-1,3-cyclopentadiene, may arise from that isomer's lower stability or may be indicative of a more selective pathway for formation of the 1-cyano isomer, perhaps one starting from acyclic precursors. The absence of N-heterocycles such as pyrrole and pyridine is discussed in light of the astronomical finding of 1-cyano-CPD.

The recent astronomical detection of benzonitrile, c-C_6H_5CN (ref. [1]), has at least partially resolved a long-standing conundrum in astrochemistry—the apparent absence of five- and six-membered rings, which are the building blocks of organic chemistry on earth. For example, of the more than 135 million compounds registered with the Chemical Abstract Service, it is estimated[2,3] that nearly 80% contain at least one of these two rings. In contrast, nearly 20 acyclic compounds with a comparable number of carbon atoms to benzene, c-C_6H_6, were known astronomically[4] before the discovery of benzonitrile. Although benzene is non-polar, and hence cannot be detected via its rotational transitions, its five-membered ring analogue cyclopentadiene (CPD, c-C_5H_6) is weakly polar, with a small permanent dipole moment ($\mu_b = 0.420$ D; refs. [5,6]) and a well-known rotational spectrum[5]. Nevertheless, as with benzene, replacing a single H atom in CPD with a CN group imparts the three possible cyanocyclopentadiene (cyano-CPD) variants with substantial dipole moments, regardless of the substitution site, meaning that their rotational lines are roughly 100 times more conspicuous than those of the hydrocarbon parent in space at the same abundance.

Unlike benzene, CPD is a highly reactive diene, which readily dimerizes at room temperature via a Diels–Alder reaction[7]. It is widely used in synthetic organic chemistry, often for stereoselective purposes, while its anion c-$C_5H_5^-$ is central to organometallic chemistry owing to the stability of metallocenes, such as derivatives of ferrocene $Fe(C_5H_5)_2$ (ref. [8]), which are regularly used for the catalytic synthesis of asymmetric and chiral molecules. Functionalized derivatives of CPD have been the subject of many studies (ref. [9] and references therein); substitution at the 1-carbon position is thermodynamically favoured[10] because it is fully conjugated with the π-electron system, as opposed to the other two positions, which are either cross-conjugated (2-carbon) or not conjugated (5-carbon).

During the course of a laboratory study to identify products formed in an electrical discharge starting from benzene and molecular nitrogen[11], evidence was found for the two cyano-CPDs shown in Fig. 1. Using cavity Fourier transform microwave spectroscopy, the rotational spectrum of each was subsequently measured at very high spectral resolution (~0.1 ppm) between 5 and 40 GHz, work that yielded highly accurate values for the three rotational constants, several of the leading quartic centrifugal distortion terms and the two nitrogen quadrupole tensor elements (Supplementary Tables 1 and 2). Confirmation of the elemental composition and structure for both molecules is provided by the extremely close agreement of the rotational constants given in Table 1 with those calculated theoretically and derived from previous measurements, although the earlier spectral analyses[12,13] were performed at considerably lower resolution (40–100 kHz or 2–5 ppm) than the present work (2 kHz or 0.1 ppm). With the new measurements, the rotational spectrum of both molecules can now be calculated to better than 0.05 km s^{-1} at 20 GHz in terms of equivalent radial velocity, more than sufficient for an astronomical search in the coldest, most quiescent molecular clouds, where lines may be as narrow as 0.3–0.4 km s^{-1} full-width at half-maximum.

Concurrent with the laboratory work, a large-scale, high-sensitivity spectral line survey, Green Bank Telescope Observations of TMC-1: Hunting for Aromatic Molecules

[1]Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA. [2]National Radio Astronomy Observatory, Charlottesville, VA, USA. [3]Department of Physics and Astronomy, Benedictine College, Atchison, KS, USA. [4]Center for Astrochemical Studies, Max Planck Institute for Extraterrestrial Physics, Garching, Germany. [5]Institute for Theoretical Chemistry, University of Stuttgart, Stuttgart, Germany. [6]Astrochemistry Laboratory and the Goddard Center for Astrobiology, NASA Goddard Space Flight Center, Greenbelt, MD, USA. [7]Institute for Astrophysics and Computational Sciences, The Catholic University of America, Washington DC, USA. [8]Department of Chemistry, University of Virginia, Charlottesville, VA, USA. [9]Department of Astronomy, University of Virginia, Charlottesville, VA, USA. [10]Astro Space Center, Lebedev Physical Institute, Russian Academy of Sciences, Moscow, Russia. [11]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ✉e-mail: mccarthy@cfa.harvard.edu; brettmc@mit.edu

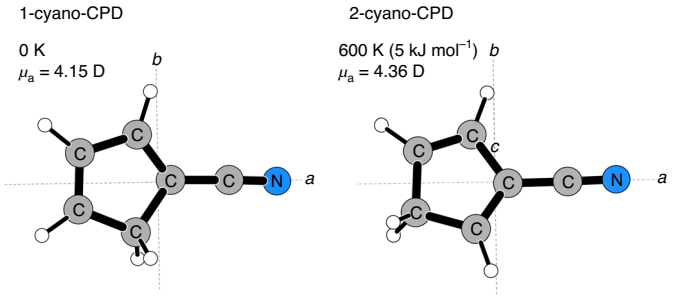


Fig. 1 | Geometric structures of the two low-lying cyano-CPD isomers, along with their relative stabilities calculated theoretically at the G3// B3LYP level of theory. Owing to its substantially lower stability, the third isomer, 5-cyano-CPD, is not shown. Dipole moments are from ref. [13].

Table 1 | Comparison of experimental and theoretical rotational constants of 1- and 2-cyano-CPD

	1-cyano-CPD			
Constant	**This work**	**Ref. [12]**	**Ref. [13]**	**Theoretical**
A_0	8,352.98(2)	8,356(5)	8,353.97(9)	8,424.05
B_0	1,904.2514(3)	1,904.24(2)	1,904.24(1)	1,915.81
C_0	1,565.3659(3)	1,565.36(2)	1,565.38(1)	1,575.55
	2-cyano-CPD			
Constant	**This work**		**Ref. [13]**	**Theoretical**
A_0	8,235.66(4)		8,235.0(13)	8,279.26
B_0	1,902.0718(2)		1,902.07(2)	1,916.03
C_0	1,559.6502(2)		1,559.67(2)	1,570.48

Constants are given in megahertz, with values in parentheses corresponding to 1σ uncertainty. The constants in this study were derived used an *A*-reduced Hamiltonian in the I^r representation. A complete set of experimental spectroscopic constants and associated uncertainties is given in Supplementary Tables 1 and 2. Theoretical values are calculated at the ωB97X-D/cc-pVQZ level of theory, with anharmonic vibrational corrections computed at the same level using second-order vibrational perturbation theory.

(GOTHAM), predominantly in the K (18–27 GHz) and K_a bands (26–40 GHz), has been underway since autumn 2018 toward the molecule-rich starless cloud core TMC-1 using the 100-m Robert Byrd Green Bank Telescope—the same source and radio telescope as used to detect benzonitrile[1]. Although only ~30% complete at the time of submission, where there is frequency coverage it is roughly an order of magnitude more sensitive than a survey toward the same source using the Nobeyama 45-m telescope[14]. The higher sensitivity arises primarily from three factors: the larger collection area of the 100-m dish; a beam size better matched to the source; and uniformly adopting a velocity resolution (0.05 km s^{-1}) that is appropriate for this narrow line source. The resolution in the Nobeyama survey of 0.22–1.26 km s^{-1} was frequently two to four times too low. An overview of the GOTHAM survey, full observational details and a discussion of the data reduction are given in ref. [15].

Results

A detailed description of the data analysis procedures and statistics is presented by R.A.L. et al. (submitted manuscript) and summarized in Methods. Briefly, a Markov chain Monte Carlo (MCMC) fit is performed to the data using transitions of the target species that have predicted flux ≥5% of the strongest line. In agreement with previous work[16,17], we detect and fit four distinct velocity components in the source (after accounting for hyperfine structure) with four column densities and four source sizes. A uniform linewidth and excitation temperature is used to reduce the number of parameters being fitted. The MCMC fit to 1-cyano-1,3-cyclopentadiene (1-cyano-CPD) showed significant detection in in all four velocity components. This is shown visually in Fig. 2, where we see compelling evidence for 1-cyano-CPD, whereas the isomer 2-cyano-1,3-cyclopentadiene (2-cyano-CPD) is not strongly detected. At our current detection sensitivity, we estimate an upper limit for 2-cyano-CPD roughly one-third that of 1-cyano-CPD, taking into account relatively small differences in the dipole moments and partition functions between the two. As such, it is not yet possible to establish with certainty if 1-cyano-CPD is formed selectively or if the apparent absence of 2-cyano-CPD is simply a question of sensitivity. Although it may have no bearing on the astronomical observations, 2-cyano-CPD was found to be approximately four times less abundant than 1-cyano-CPD in our laboratory study[11] under identical experimental conditions; in the earlier study by Sakaizumi et al.[13], the ratio was 2:1, with 1-cyano-CPD again being more abundant.

To quantify the relative stabilities of the three cyano-CPD isomers, energetics have been calculated with the B3LYP variant of the G3 thermochemical method (G3//B3LYP), which has been shown to yield reliable and highly accurate (~4-kJ-mol^{-1}) energetics[18,19]. As indicated in Fig. 1, 1-cyano-CPD is the most stable of the three, with 2-cyano-CPD computed to lie higher in energy by only 5 kJ mol^{-1} (600 K), which is consistent with the energy ordering from the NMR findings of Wentrup and Crow[10] and conjugation arguments. 5-cyano-1,3-cyclopentadiene (5-cyano-CPD) is considerably less stable (26 kJ mol^{-1}, or ~3,130 K), which presumably explains why this isomer was not detected in our laboratory study or previous work; for this reason, it is not considered further. Both 1- and 2-cyano-CPD have comparably large dipole moments along their *a*-inertial axis (4.15(15) D and 4.36(25) D, respectively; ref. [13]).

Astrochemical implications

Astrochemical modelling has been carried out to estimate the abundance of 1-cyano-CPD in TMC-1. However, it is apparent from Fig. 3 that the abundance inferred from observations vastly exceeds what our simulations predict, implying a far richer aromatic inventory for TMC-1 than previously expected. This result is surprising, since the same models are in very good agreement with observations for many carbon-chain molecules at reasonable timescales for the source (10^4–10^5 yr) (refs. [20,15]), with the long-carbon-chain $HC_{11}N$ being a possible exception (R.A.L. et al., submitted manuscript), where the models actually overpredict the abundance relative to observation. The best agreement for 1-cyano-CPD is achieved at unrealistically early times in a simulation (<10^4 yr) when a substantial initial abundance of CPD is assumed (~0.5% of the carbon budget). Not surprisingly, the abundances of 1- and 2-cyano-CPD are nearly identical since, in the absence of a laboratory measurement, it was necessary to assume equal branching for cyanation of CPD. Even so, our models produce slightly more 1-cyano-CPD than 2-cyano-CPD due to the marginally larger dipole moment, and therefore faster rate of destruction with ions, of the latter isomer[21].

1-cyano-CPD is now the fourth organic ring, along with benzonitrile and 1- and 2-cyanonaphthalene (B.A.M. et al., submitted manuscript), that has been observed in TMC-1 with an abundance that exceeds, sometimes greatly so, predictions from our most up-to-date models. Although not directly comparable, the same situation appears to be true for benzonitrile in other prestellar cores (A.M.B. et al., submitted manuscript). In light of these findings, other production pathways for cyclic molecules may need to be explored or a large unseen reservoir of aromatic material, such as benzene, may need to be invoked. The surprisingly high abundances of cyclic species, deviations from previous astrochemical models, and promising potential pathways are further discussed by A. M. Burkhardt et al. (submitted manuscript).

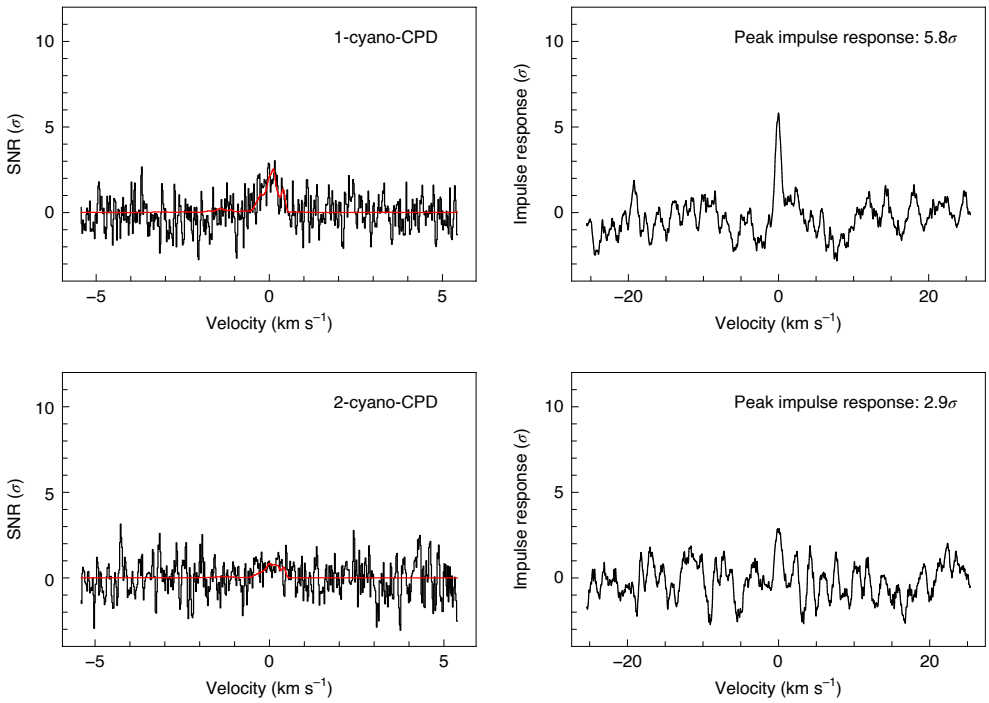


Fig. 2 | Velocity-stacked spectra of 1-cyano-CPD and 2-cyano-CPD and the impulse response function of the stacked spectra. Left: velocity-stacked spectra in black, with the corresponding stack of the simulation using the best-fit parameters to the individual lines in red. The data have been uniformly sampled to a resolution of 0.02 km s^{-1}. The intensity scale is the signal-to-noise ratio of the spectrum at any given velocity. Right: the impulse response function of the stacked spectrum generated using the simulated line profile as a matched filter. The intensity scale is the signal-to-noise ratio of the response function when centred at a given velocity. The peak of the impulse response function provides a minimum significance for the detection of 5.8σ. The bottom row shows the same, but for the 2-cyano-CPD upper limit (2.9σ). See the work of R.A.L. et al. (submitted manuscript) for details.

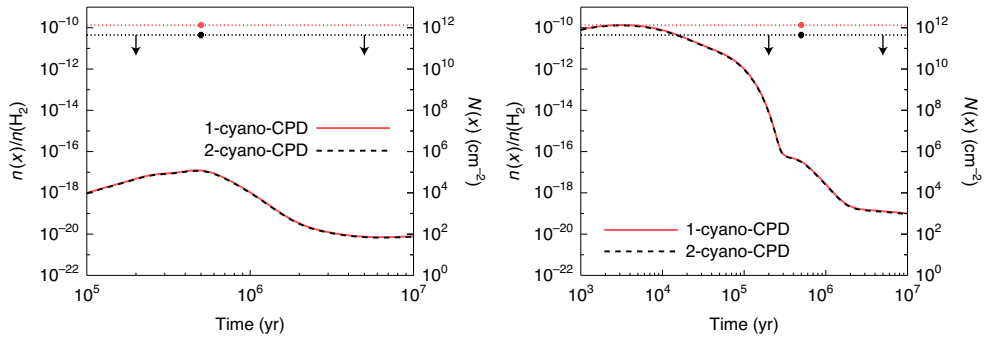


Fig. 3 | Abundance predictions for 1-cyano-CPD and 2-cyano-CPD from two chemical models in comparison with those derived from observations of TMC-1. Left: model results from a full chemical treatment. Abundances derived from the MCMC analysis of the observational data are represented by dot-dashed lines. Right: results from the same model but including an initial abundance of CPD given by $X(\text{c-}C_5H_6)_{t=0}=1.5\times10^{-7}$, or about 0.5% of the total carbon budget, chosen to best reproduce the abundances obtained from our MCMC analysis.

In this context, it is intriguing that no evidence has been found in TMC-1 to date for well-known N-heterocycles (defined here as a cyclic compound in which one or more of the heavy atoms of the ring is an atom other than carbon) of comparable size to 1-cyano-CPD and benzonitrile, such as pyrrole (C_4H_4NH, $\mu=1.80$) or pyridine (C_5H_5N, $\mu=2.19$; see Supplementary Fig. 3), differences in their polarities aside[22]. Indeed, upper limits of their column densities from the GOTHAM survey are of the same order as the measured column densities for 1-cyano-CPD and the cyanonaphthalene isomers (B.A.M. et al., submitted manuscript). One possibility is that the lack of nitrogen incorporation into the heavy atom ring may indicate that much of the chemically active nitrogen is not available in the form of NH or NH_2, since the reaction of these radicals with butadiene[23] is thought to be an important pathway to pyrrole; an analogous argument may hold for pyridine. A second possibility is that cyanation of pre-existing aromatic material to produce 1-cyano-CPD and other CN-functionalized rings is efficient. Nevertheless, in either scenario, the preponderance of nitrile-terminated chains and now rings highlights the key role the CN radical apparently plays as a chemically active sink for nitrogen in TMC-1.

Methods

Spectral stacking routine. Full details of the methodology used to detect and quantify molecules in our spectra are provided by R.A.L. et al. (submitted manuscript), including assumptions made in simulating the spectra and in determining physical parameters of the source, priors used in the fitting analysis,

and an in-depth examination of the robustness of the methodology to interlopers and false positives. Briefly, we first perform an MCMC fit to the lines of strongly detected and optically thin cyano-octatetra-yne (HC_9N) and benzonitrile in our data. Model spectra for these fiducial molecules, and for all other species, are generated using the formalisms outlined by Turner[24], which includes corrections for optical depth, and adjusted for the effects of beam dilution. The specific transition parameters for each species are obtained from spectral line catalogues primarily pulled from publicly accessible databases (https://spec.jpl.nasa.gov and https://cdms.astro.uni-koeln.de/classic/). In some cases, we have generated these catalogues directly from the parameters provided from laboratory work outlined in individual publications. In almost all cases, the substantial number of transitions used in the analysis (hundreds to thousands) makes it impractical to provide a table of parameters in the text. Instead, the interested reader is referred to the catalogue files in Supplementary Information, which contain all of the required information in a machine-readable format.

We detect four distinct velocity components (v_{lsr}) for nearly all molecules, in agreement with previous observations of the source[17]. We also simultaneously fit for column densities (N_T), source sizes (θ_s), excitation temperatures (T_{ex}) and linewidth (ΔV). The derived parameters for HC_9N and benzonitrile are then used as Gaussian priors for MCMC fits to other linear and cyclic species of interest, respectively. In general, we find that the linewidths (~0.1–0.3 km s^{-1}) and excitation temperatures (5–7 K) are consistent across velocity components and molecular species. Both this trend, and the derived values, are in excellent agreement with previous observations of the source[14,17,25].

For nearly all species, there are many more transitions covered by our spectra than are visible above the local root mean square noise level of the observations. We therefore extract a small portion of the observations centred around each spectral line, disregarding any windows that have a spectral feature greater than 5σ, to avoid interloping signals from other species. A average of these spectra weighted by signal-to-noise ratio is then performed on the basis of the expected intensity of the line (derived from the MCMC parameters) and the local root mean square noise of the observations. For the purposes of this analysis, largely due to hyperfine splitting, we treat the signals on a per-line basis rather than a per-transition basis. This result is that the stacked feature is somewhat broadened, as the hyperfine components and velocity components are not collapsed, but there is no overcounting of flux. This results in a substantial increase in overall signal-to-noise ratio, with the spectrum now encapsulating the total information content of all observed lines, rather than only that from the brightest lines. Finally, the model spectra are stacked using identical weights, and this stacked model is used as a matched filter, which is cross-correlated with the stacked observations. The resulting impulse response spectrum provides a lower-limit statistical significance to the detection. Because the filter contains the same broadened hyperfine and velocity structure as the stack, there is no loss in significance. Below, we provide a number of details pertinent to the MCMC fitting and stacking analyses performed for the specific molecules discussed in this work.

Extended Data Fig. 1 shows the total number of transitions (including hyperfine components) of the molecules analysed in this Article that were covered by GOTHAM observations at the time of analysis and were above our predicted flux threshold of 5%, as discussed by R.A.L. et al. (submitted manuscript). Also included are the number of transitions, if any, that were coincident with interfering transitions of other species, and the total number of lines used after excluding interlopers. Observational data windowed around these transitions, spectroscopic properties of each transition, and the partition function used in the MCMC analysis are provided in the Harvard Dataverse repository[26].

Reaction pathways and energetics. Possible formation pathways for 1-cyano- and 2-cyano-CPD under interstellar conditions have been investigated by calculating the reaction between CPD and the CN radical, by analogy to benzonitrile formation[27–29]. For this and other potential surface calculations, the growing string method[30–32] has been used: it is computationally efficient, allows various reactants to be rapidly surveyed and reaction barriers identified, and yields approximate reaction enthalpies between stationary points. Stationary points are then subsequently refined with the G3//B3LYP method. As shown in Extended Data Fig. 2, the c-C_5H_6 + CN reaction to yield either 1-cyano or 2-cyano-CPD is found to be both exothermic and barrierless, irrespective of the stability of the the two isomers.

On the operative assumption that 1-cyano-CPD is formed from CPD, the reaction pathways that might produce this five-membered hydrocarbon ring in space have also been investigated, since they apparently have not been considered in current chemical models for dark clouds. Two possible reactions are shown in Extended Data Fig. 3. The red trace follows the reaction between ethylene ($H_2C{=}CH_2$) and propargyl radical ($HC{\equiv}CCH_2$), which combines barrierlessly to form an acyclic intermediate. Before CPD formation, however, this intermediate must undergo ring-closure, which requires surmounting a barrier in excess of its initial energy (~33 kJ mol^{-1}, or 4,000 K), implying that this pathway is not viable in TMC-1. The blue trace follows the reaction between *gauche*-butadiene and CH radical, which is calculated to produce CPD exothermically and barrierlessly via subsequent hydrogen atom loss. Due to large energy barriers, the formation of CPD via reaction of allyl radical and acetylene[33] can also be neglected. It should be emphasized that reactions with high activation barrier may take place on grains by the bombardment of cosmic rays and internal ultraviolet photons[34,35]. Though we have not considered the effects of such reactions involving excited, suprathermal species here, it is possible that the contribution of these processes to the abundances of gas-phase aromatic molecules could be non-trivial, even in cold cores such as TMC-1.

Astrochemical modelling. We have used the NAUTILUS v1.1 program[36], together with a modified version the KIDA 2014 network[37], to predict the abundance of 1-cyano-CPD and other species in our GOTHAM observations (R.A.L. et al., submitted manuscript, ref. [20], B.A.M. et al., submitted manuscript). This network includes the barrierless, exothermic reaction of CPD with CN radical to yield 1-cyano-CPD and 2-cyano-CPD, which we assume has a rate coefficient of ~3×10^{-10} cm^{-3} s^{-1}, which is reasonable assuming that the reaction occurs with every collision[29,38]. Pathways that yield CPD from acyclic precursors are also explicitly included in our model, as are destruction reactions with ions[39] and depletion onto grains.

Standard physical conditions relevant to TMC-1 were used in our simulations, for example, a gas temperature of $T_{gas} = T_{dust} = 10$ K, gas density of 2×10^4 cm^{-3} and standard cosmic ray ionization rate of 1.3×10^{-17} s^{-1} (refs. [1,40]). For initial elemental abundances, we use the values of Hincelin et al.[41], except for the initial atomic oxygen, where we have chosen an initial value of $X(O)_{t=0} \approx 1.5\times10^{-4}$, resulting in a slightly carbon-rich C/O ≈ 1.1 by optimizing to best reproduce the observed cyanopolyyne abundances (R.A.L. et al., submitted manuscript). To test the possible influence of CPD inherited from earlier stages of cloud evolution, we have also run a second simulation including an initial abundance of CPD $X(c\text{-}C_5H_6)_{t=0} = 1.5\times10^{-7}$, which was chosen to best reproduce the abundances obtained from our MCMC analysis. This is equivalent to approximately 0.5% of the entire carbon budget in these simulations.

Data availability

The datasets analysed during the current study are available in the Green Bank Telescope archive (https://archive.nrao.edu/archive/advquery.jsp). A user manual for their reduction and analysis is available as well (https://greenbankobservatory.org/science/gbt-observers/visitor-facilities-policies/data-reduction-gbt-using-idl/). The complete, reduced survey data in the X band are available as supplementary information in ref. [15]. The individual portions of reduced spectra used in the analysis of the individual species presented here are available in the Harvard Dataverse archive[26].

Code availability

All the codes used in the MCMC fitting and stacking analysis presented in this paper are open source and publicly available at https://github.com/ryanaloomis/TMC1_mcmc_fitting.

Received: 16 March 2020; Accepted: 11 August 2020;
Published online: 28 September 2020

References

1. McGuire, B. A. et al. Detection of the aromatic molecule benzonitrile (c-C_6H_5CN) in the interstellar medium. *Science* **359**, 202–205 (2018).
2. Lipkus, A. H. et al. Structural diversity of organic chemistry. A scaffold analysis of the CAS registry. *J. Org. Chem.* **73**, 4443–4451 (2008).
3. Ruddigkeit, L., van Deursen, R., Blum, L. C. & Reymond, J.-L. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J. Chem. Inf. Model.* **52**, 2864–2875 (2012).
4. McGuire, B. A. 2018 census of interstellar, circumstellar, extragalactic, protoplanetary disk, and exoplanetary molecules. *Astrophys. J. Suppl. Ser.* **239**, 17 (2018).
5. Laurie, V. W. Microwave spectrum and dipole moment of cyclopentadiene. *J. Chem. Phys.* **24**, 635–636 (1956).
6. Damiani, D., Ferretti, L. & Gallinella, E. Structure of cyclopentadiene from microwave spectra of several deuterated species. *Chem. Phys. Lett.* **37**, 265–269 (1976).
7. Turro, N. J. & Hammond, G. S. The photosensitited dimerization of cyclopentadiene. *J. Am. Chem. Soc.* **84**, 2841–2842 (1962).
8. Werner, H. At least 60 years of ferrocene: the discovery and rediscovery of the sandwich complexes. *Angew. Chem. Int. Ed.* **51**, 6052–6058 (2012).
9. Dalkılıç, E. & Daştan, A. Synthesis of cyclopentadiene derivatives by retro-Diels–Alder reaction of norbornadiene derivatives. *Tetrahedron* **71**, 1966–1970 (2015).
10. Wentrup, C. & Crow, W. Structures of cyanocyclopentadienes and related compounds. *Tetrahedron* **26**, 4375–4386 (1970).
11. McCarthy, M. C. et al. Exhaustive product analysis of three benzene discharges by microwave spectroscopy. *J. Phys. Chem. A* **124**, 5170–5181 (2020).

12. Ford, R. G. & Seitzman, H. A. The microwave spectrum and dipole moment of 1-cyano-1,3-cyclopentadiene. *J. Mol. Spectrosc.* **69**, 326–329 (1978).
13. Sakaizumi, T., Kikuchi, H., Ohashi, O. & Yamaguchi, I. The microwave spectra, dipole moments, and two isomers of cyclopentadiene-1-carbonitrile. *Bull. Chem. Soc. Jpn.* **60**, 3903–3909 (1987).
14. Kaifu, N. et al. A 8.8–50 GHz complete spectral line survey toward TMC-1 I. Survey data. *Publ. Astron. Soc. Jpn.* **56**, 69–173 (2004).
15. McGuire, B. A. et al. Early science from GOTHAM: project overview, methods, and the detection of interstellar propargyl cyanide ($HCCCH_2CN$) in TMC-1. *Astrophys. J. Lett.* **900**, L10 (2020).
16. Dobashi, K. et al. Spectral tomography for the line-of-sight structures of the Taurus Molecular Cloud 1. *Astrophys. J.* **864**, 82 (2018).
17. Dobashi, K. et al. Discovery of CCS velocity-coherent substructures in the Taurus Molecular Cloud 1. *Astrophys. J.* **879**, 88 (2019).
18. Simmie, J. M. & Somers, K. P. Benchmarking compound methods (CBS-QB3, CBS-APNO, G3, G4, W1BD) against the active thermochemical tables: a litmus test for cost-effective molecular formation enthalpies. *J. Phys. Chem. A* **119**, 7235–7246 (2015).
19. 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.* **21**, 2946–2956 (2019).
20. Xue, C. et al. Detection of interstellar HC_4NC and an investigation of isocyanopolyyne chemistry in TMC-1 conditions. *Astrophys. J. Lett.* **900**, L9 (2020).
21. Shingledecker, C. N., Molpeceres, G., Rivilla, V. M., Majumdar, L. & Kaestner, J. Isomers in interstellar environments (I): the case of Z- and E-cyanomethanimine. *Astrophys. J.* (in the press).
22. Nelson, R. D. Jr, Lide, D. R. & Maryott, A. A. *Selected Values of Electric Dipole Moments for Molecules in the Gas Phase* National Standard Reference Data Series—National Bureau of Standards 10 (US Department of Commerce National Bureau of Standards, 1967).
23. Lee, K. L. K., McGuire, B., Burkhardt, A. & McCarthy, M. C. Interstellar aromatic chemistry: a combined laboratory, observational, and theoretical perspective. In *Laboratory Astrophysics (IAU S350): from Observations to Interpretation* (eds Salama, F. & Linnartz, H.) (Cambridge Univ. Press, in the press).
24. Turner, B. E. A molecular line survey of Sagittarius B2 and Orion-KL from 70 to 115 GHz. II—Analysis of the data. *Astrophys. J. Suppl. Ser.* **76**, 617–686 (1991).
25. Gratier, P. et al. A new reference chemical composition for TMC-1. *Astrophys. J. Suppl. Ser.* **225**, 1110 (2016).
26. GOTHAM Collaboration. Spectral stacking data for Phase 1 science release of GOTHAM. *Harvard Dataverse* V4 https://doi.org/10.7910/DVN/PG7BHO (2020).
27. Balucani, N. et al. Formation of nitriles in the interstellar medium via reactions of cyano radicals, $CN(X\ ^2\Sigma^+)$, with unsaturated hydrocarbons. *Astrophys. J.* **545**, 892–906 (2000).
28. Woods, P. M., Millar, T. J., Zijlstra, A. A. & Herbst, E. The synthesis of benzene in the proto-planetary nebula CRL 618. *Astrophys. J.* **574**, L167–L170 (2002).
29. Cooke, I. R., Gupta, D., Messinger, J. P. & Sims, I. R. Benzonitrile as a proxy for benzene in the cold ISM: low temperature rate coefficients for $CN + C_6H_6$. *Astrophys. J. Lett.* **891**, L41 (2020).
30. Zimmerman, P. M. Growing string method with interpolation and optimization in internal coordinates: method and examples. *J. Chem. Phys.* **138**, 184102 (2013).
31. Zimmerman, P. M. Automated discovery of chemically reasonable elementary reaction steps. *J. Comput. Chem.* **34**, 1385–1392 (2013).
32. Jafari, M. & Zimmerman, P. M. Reliable and efficient reaction path and transition state finding for surface reactions with the growing string method. *J. Comput. Chem.* **38**, 645–658 (2017).
33. Bouwman, J., Bodi, A., Oomens, J. & Hemberger, P. On the formation of cyclopentadiene in the $C_3H_5 + C_2H_2$ reaction. *Phys. Chem. Chem. Phys.* **17**, 20508–20514 (2015).
34. Shingledecker, C. N., Tennis, J., Gal, R. L. & Herbst, E. On cosmic-ray-driven grain chemistry in cold core models. *Astrophys. J.* **861**, 20 (2018).
35. Shingledecker, C. N. & Herbst, E. A general method for the inclusion of radiation chemistry in astrochemical models. *Phys. Chem. Chem. Phys.* **20**, 5359–5367 (2018).
36. Ruaud, M., Wakelam, V. & Hersant, F. Gas and grain chemical composition in cold cores as predicted by the Nautilus three-phase model. *Mon. Not. R. Astron. Soc.* **459**, 3756–3767 (2016).
37. Wakelam, V. et al. The 2014 KIDA network for interstellar chemistry. *Astrophys. J. Suppl. Ser.* **217**, 20 (2015).
38. 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* **114**, 1749–1755 (2010).
39. Woon, D. E. & Herbst, E. Quantum predictions of the properties of known and postulated neutral interstellar molecules. *Astrophys. J. Suppl. Ser.* **185**, 273–288 (2009).
40. Loomis, R. A. et al. Non-detection of $HC_{11}N$ towards TMC-1: constraining the chemistry of large carbon-chain molecules. *Mon. Not. R. Astron. Soc.* **463**, 4175–4183 (2016).
41. Hincelin, U. et al. Oxygen depletion in dense molecular clouds: a clue to a low O_2 abundance? *Astron. Astrophys.* **530**, A61 (2011).

Acknowledgements

M.C.M. and K.L.K.L. acknowledge support from National Science Foundation (NSF) grant AST-1908576 and NASA grant 80NSSC18K0396. A.M.B. acknowledges support from the Smithsonian Institution as a Submillimeter Array Fellow. C.N.S. thanks the Alexander von Humboldt Stiftung/Foundation for their generous support, as well as V. Wakelam for use of the NAUTILUS v1.1 code. S.B.C. and M.A.C. were supported by the NASA Astrobiology Institute through the Goddard Center for Astrobiology. E.H. thanks the NSF for support through grant AST-1906489. C.X. is a Grote Reber Fellow, and support for this work was provided by the NSF through the Grote Reber Fellowship Program administered by Associated Universities, Inc./National Radio Astronomy Observatory and the Virginia Space Grant Consortium. Support for B.A.M. was provided by NASA through Hubble Fellowship grant HST-HF2-51396 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. The Green Bank Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc.

Author contributions

M.C.M. and K.L.K.L. performed the laboratory experiments and theoretical calculations and wrote the paper with the help of A.M.B. and C.N.S. A.M.B, B.A.M., A.J.R. and R.A.L. performed the astronomical observations and subsequent analysis. E.H. determined and/or estimated rate coefficients and is the originator of many of the chemical simulations. A.M.B. and C.N.S. contributed or undertook the astronomical modelling and simulations. E.R.W., M.A.C., S.B.C., S.K., C.X., B.A.M., A.J.R. and R.A.L. contributed to the design of the GOTHAM survey, and helped revise the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-020-01213-y.

Supplementary information is available for this paper at https://doi.org/10.1038/s41550-020-01213-y.

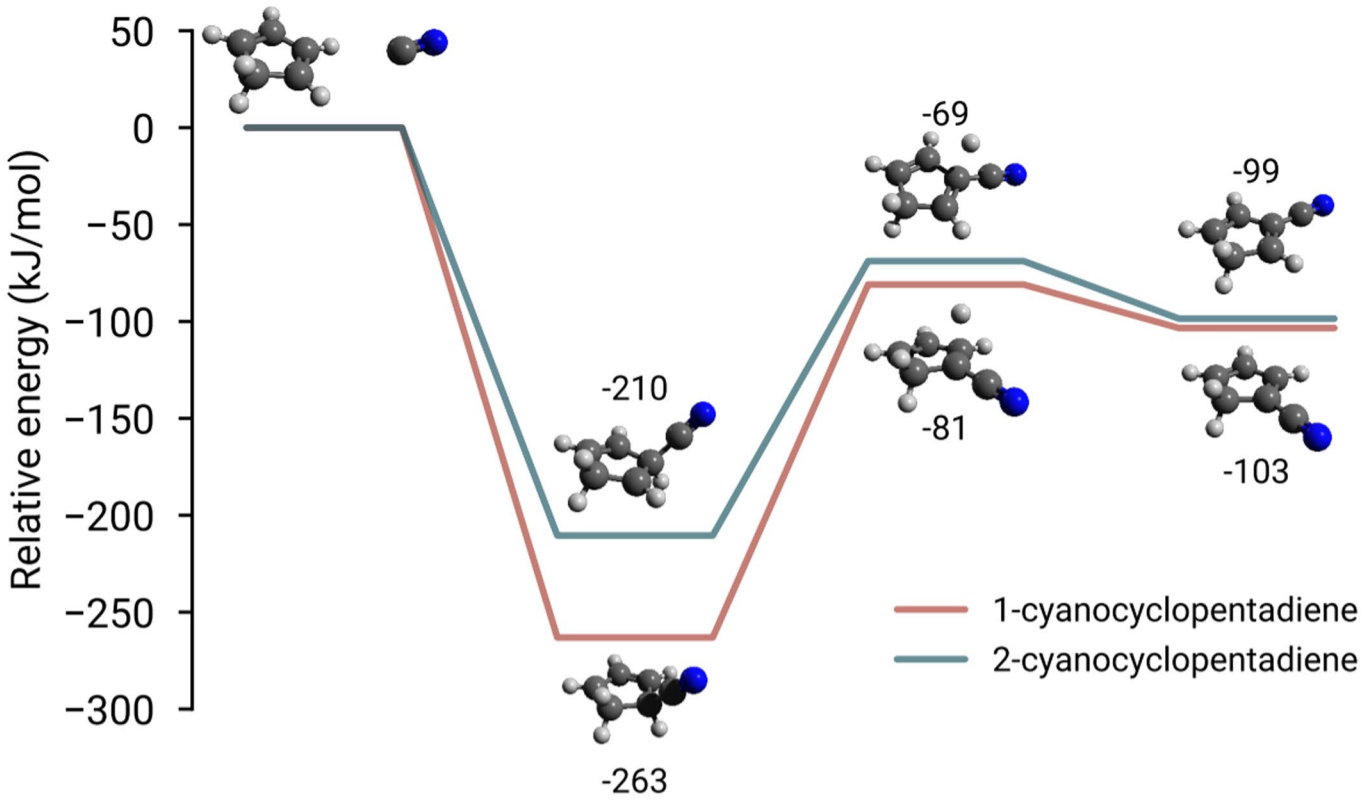
Correspondence and requests for materials should be addressed to M.C.M. or B.A.M.

Reprints and permissions information is available at www.nature.com/reprints.

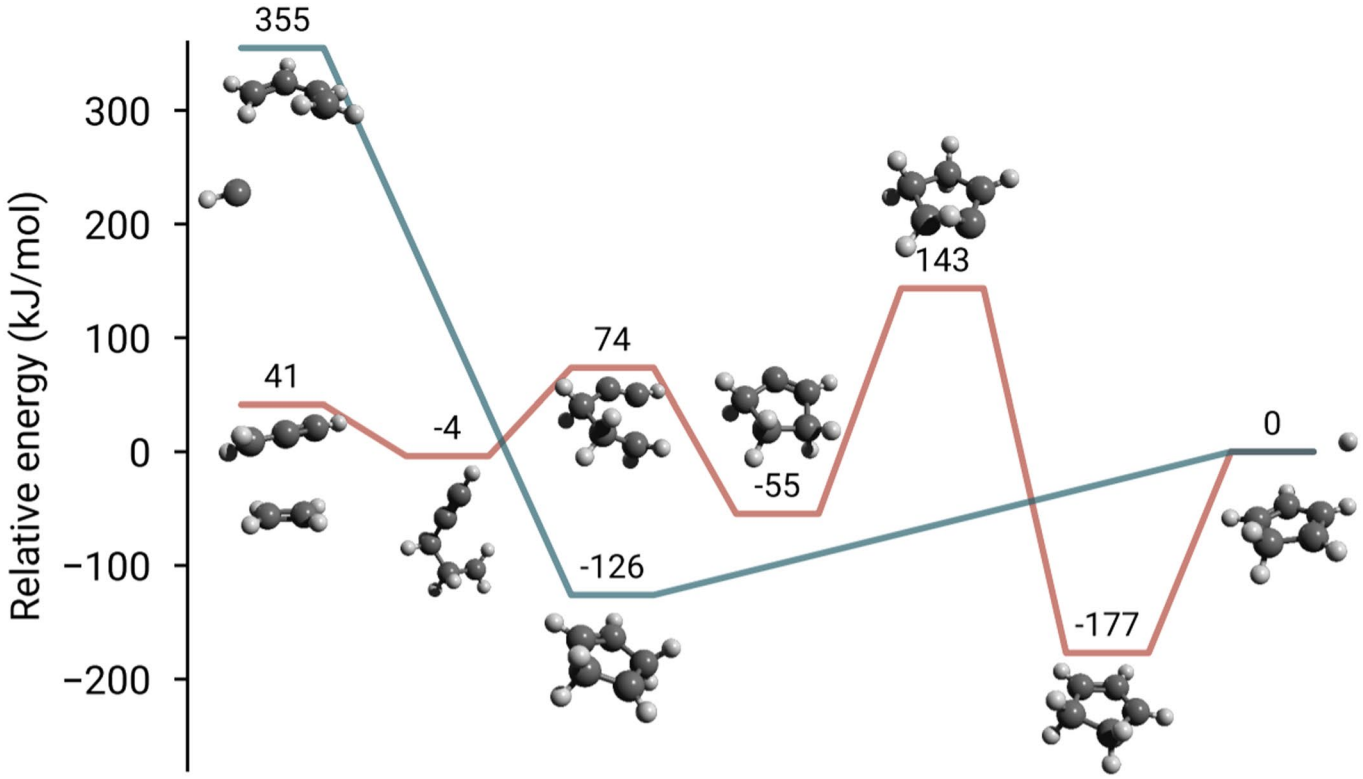
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Molecule	Transitions Covered By GOTHAM	Interfering Lines In Data	Total Transitions Used in MCMC
1-Cyano-CPD	111	0	111
2-Cyano-CPD	108	0	108
cyclopentadiene	21	0	21
pyridine	76	0	76
pyrrole	34	0	34

Extended Data Fig. 1 | Spectral data. Total number of transitions of a given species within the range of the GOTHAM data, number of interfering lines, and total number included in MCMC fit.



Extended Data Fig. 2 | Thermochemistry (0 K) for the reaction between CN radical and CPD. The calculated has been performed at the G3//B3LYP level of theory, and energies in kJ/mol are given relative to the reactant asymptote. The reaction bifurcates as CN attacks CPD barrierlessly, forming a radical intermediate. Subsequent hydrogen atom loss yields 1-cyano- and 2-cyano-CPD.



Extended Data Fig. 3 | Potential energy surface for the formation of CPD at 0 K. Reaction energies are given relative to the product (CPD + H atom) asymptote. The red trace corresponds to reaction between propargyl radical ($HCCCH_2$) and ethylene ($H_2C{=}CH_2$), and the blue trace represents gauche-butadiene ($H_2C(CH)_2CH_2$) reacting with CH radical.