PCCP

Polycyclic (Anti)aromatic Hydrocarbons: Interstellar Formation and Spectroscopic Characterization of Biphenylene and Benzopentalene

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-02-2025-000630.R2
Article Type:	Paper
Date Submitted by the Author:	22-Apr-2025
Complete List of Authors:	Flint, Athena; University of Mississippi, Department of Chemistry & Biochemistry Esposito, Vincent; NASA Ames Research Center Fortenberry, Ryan; University of Mississippi, Department of Chemistry & Biochemistry

PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Polycyclic (Anti)aromatic Hydrocarbons: Interstellar Formation and Spectroscopic Characterization of Biphenylene and Benzopentalene†

Athena R. Flint,[a] Vincent J. Esposito,[b] and Ryan C. Fortenberry*[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Formation of biphenylene, $(C_6H_4)_2$, and its isomer benzopentalene, $C_{12}H_8$, may act as a consumption route for *ortho*-benzyne (*o*-C_6H_4) in interstellar clouds such as TMC-1. MRCI-F12 and CCSD(T)-F12 potential energy surfaces show that *o*-C_6H_4 dimerization is possible through a C_{2h}-symmetry single-bond association to a $(C_6H_4)_2$ precursor before isomerization to $(C_6H_4)_2$ and subsequently $C_{12}H_8$. Formation of a bimolecular product set from either species is energetically hindered, allowing $(C_6H_4)_2$ and $C_{12}H_8$ to stabilize radiatively. To remedy the dearth of spectroscopic data for these species, anharmonic frequencies from explicitly-correlated quartic force fields (QFFs) for *o*-C_6H_4 and *c*-C_4H_4 are used to reparameterize the semiempirical method PM6 for use in lower-cost QFFs for $(C_6H_4)_2$ and $C_{12}H_8$. In both cases, at least one reparameterized PM6-QFF spectrum results in the prominent C-H stretch and symmetric C-H out-of-plane-bend features to be accurately predicted with respect to gas-phase experiment or the B3LYP/N07D anharmonic absorption spectrum. B3LYP/N07D accurately recreates the experimental infrared spectrum of $(C_6H_4)_2$, showing the utility of this method for spectral prediction of small and midsize polycyclic hydrocarbons on the whole. For larger systems, reparameterized PM6-QFF spectra can reproduce the most important infrared features for a species. B3LYP/N07D cascade emission spectra show that the 730 cm^{-1} C-H symmetric out-of-plane bending feature dominates the emission spectrum of $(C_6H_4)_2$, while the spectrum of $C_{12}H_8$ becomes characterized by the collective set of C-H out-of-plane bends. As such, infrared emission spectra of $(C_6H_4)_2$ will likely be overshadowed by C_2H_2. Derivatives such as $C_6H_4C_6H_3CN$ are better targets for infrared observation.

1 Introduction

Ortho-benzyne (*o*-C_6H_4) is an alkynic derivative of benzene (C_6H_6) that is granted greater reactivity than its parent molecule due to the enhanced strain and high electron density at the alkyne site(s).[1,2] As such, it has long been of interest in the synthesis of larger hydrocarbons.[3–16] The reactivity of *o*-C_6H_4 is especially relevant in the context of gas-phase interstellar chemistry, as it has been observed in the cold molecular cloud TMC-1[17] and may be of importance in the synthesis of the large cosmic population of polycyclic aromatic hydrocarbons (PAHs). Interestingly, the reported concentration of *o*-C_6H_4 is somewhat low ($[5 \pm 1] \times 10^{-11}$ relative to H_2) given the numerous gas-phase[18,19] and fragmentation[20,21] processes reported to create it. *o*-C_6H_4 consumption through the formation of larger species is likely a factor in the low concentrations observed in TMC-1.

Further chemical reactions involving *o*-C_6H_4 in the complex but frigid chemistry in TMC-1, however, likely require barrierless, exothermic process. The low temperatures in the cold interstellar medium (ISM) induce a bottleneck that inhibit most barriered reactions.[22] Only a few of the processes that could contribute to *o*-C_6H_4 depletion are shown to be barrierless.[7,8,11,14] As such, there is likely additional *o*-C_6H_4 chemistry that has yet to be explored in the context of TMC-1 as well as the larger ISM.

Any reactions involving *o*-C_6H_4 in TMC-1 will be limited to the available chemical inventory. Interestingly, *o*-C_6H_4 has been known to be able to dimerize for some time,[3–5] resulting in the tricyclic species biphenylene, $(C_6H_4)_2$. $(C_6H_4)_2$ has an unusual electronic structure, as it is formally antiaromatic via Hückel's Rule, but this antiaromaticity is largely constrained to the cyclobutadiene-like group in the center. The 6-membered-ring moieties on either end of $(C_6H_4)_2$ exhibit some aromatic

[a] *Department of Chemistry and Biochemistry, University of Mississippi, University, Mississippi 38677-1848, USA; E-mail: r410@olemiss.edu*
[b] *NASA Ames Research Center, MS N245-6, Moffett Field, CA 94035, USA*

† Supplementary Information available: S1-S36: Cartesian coordinates, F12-DZ T1 and D1 diagnostic values, and harmonic frequencies for all reaction pathway structures; S37-S50: full frequency lists (harmonic, fundamental, and corrected fundamental) for all QFFs; S51: generated semiempirical parameters; S52-S55: rotational constants for all QFFs; S56-S68: Coriolis and Fermi resonances for all QFFS. See DOI: 10.1039/cXCP00000x/

character, resulting in a stability that is uncommon for antiaromatic species.[23] The formation of $(C_6H_4)_2$, while shown to occur quickly in the gas phase,[4] has not been studied under low-temperature conditions similar to those within TMC-1. These previous experiments show the dimerization rate to be invariant with respect to temperature in the applied range (90 - 200 °C), but reactions at these temperatures may obscure an association barrier on the potential energy surface, requiring computational investigation of the reaction pathway.

However, the ability of *o*-C_6H_4 to dimerize in the ISM does not guarantee that $(C_6H_4)_2$ is observable, which is necessary to confirm if $(C_6H_4)_2$ is available for further PAH growth. $(C_6H_4)_2$ possesses no permanent dipole moment, largely limiting its observation to infrared methods. Furthermore, vibrational data on $(C_6H_4)_2$ is lacking,[24,25] and the generation of such data with high-accuracy quantum-chemical methods **such as explicitly-correlated coupled-cluster theory** is rendered nearly impossible due to the size of $(C_6H_4)_2$. **The geometric scaling of anharmonic vibrational frequency computations combined with the scaling of accurate electronic treatment methods with basis set size results in a computational cost that increases drastically with the size of the molecule.** $(C_6H_4)_2$ is also shown to isomerize to benzopentalene ($C_{12}H_8$),[26] a process that could occur prior to the radiative stabilization of $(C_6H_4)_2$. However, there is no spectroscopic data, computational or otherwise, in the literature for $C_{12}H_8$, a species that falls victim to the same data-generation issues as mentioned above due to its size and lowered symmetry.

This work aims to exploit highly-accurate multireference and explicitly-correlated quantum chemical calculations to identify the nature of the potential surface along the *o*-C_6H_4-*o*-C_6H_4 reaction coordinate. This will determine if $(C_6H_4)_2$ is a viable interstellar reaction product, adding another piece to the chemical inventory of the ISM. The reaction pathways for the net-exothermic bimolecular dissociation of $(C_6H_4)_2$ into benzene and triacetylene as well as diethynylbenzene and acetylene are investigated to determine if such processes may inhibit the radiative stabilization of $(C_6H_4)_2$ or $C_{12}H_8$.[27] **Vibrational** absorption spectra are also generated for both $(C_6H_4)_2$ and $C_{12}H_8$ through modified semiempirical methods in order to provide accurate spectral data at a lower computational cost. Data for $(C_6H_4)_2$ will be compared to the experimental gas-phase IR spectrum to ascertain which semiempirical approach is the most accurate. **This information will reveal the extendability of these modified semiempirical approaches to even larger polycyclic hydrocarbons for which *ab initio* or DFT approaches become infeasible.** Additionally, the emission spectra of these species are computed, which are especially important for the observation of interstellar polycyclic hydrocarbons that are expected to have differences in their absorption and emission profiles due to the absorption of UV photons that are then lost through a vibrational cascade.[28] These reaction pathways and spectroscopic data will outline new targets for observation with **the James Webb Space Telescope** and expand the understanding of interstellar carbon chemistry as well as PAH formation.

2 Computational Details

2.1 Reaction Pathways

Most minimum and transition state structures on the reaction pathway are optimized at the B3LYP/aug-cc-pVDZ[29–34] level of theory (in MOLPRO 2024.1[35–37] and GAUSSIAN16[38]) prior to harmonic frequency calculations at the same level. To improve the accuracy of the reported energies, single-point energies are computed for each optimized structure at the CCSD(T)-F12/cc-pVDZ-F12[39–41] (F12-DZ) level of theory in MOLPRO 2024.1. Here, an explicit-correlation (F12) correction is applied to the "gold standard" CCSD(T) method. CCSD(T)-F12, when combined with the requisite F12-optimized Dunning basis set of size X, gives results comparable in quality to those produced by CCSD(T)/cc-pV(X+2)Z calculations.[42] This approach has been validated previously to bypass possible potential surface inaccuracies associated with the use of B3LYP for stationary point optimization.[19] Such calculations for *o*-C_6H_4, C_2H_2, and *l*-C_6H_2, due to their small size, are done only at the F12-DZ level of theory.

2.2 Potential Energy Surfaces

One-dimensional relaxed potential energy surfaces (PESs) are used to create an energetic profile of reaction coordinates that lack or have difficult-to-locate transition states. The reaction coordinate is followed by fixing the associated variable (often an interatomic distance) to a series of values within a predetermined range and letting all other structural variables optimize. The resulting energies are plotted against the corresponding location of that point along the reaction coordinate. In this work, all PESs are computed over an interatomic distance and are examined for bond lengths starting from 0.1-0.5 Å less than the equilibrium value to a value of 5.0 Å with a step size of 0.1 Å. Two levels of theory are used to create the PESs, and both employ MOLPRO 2024.1. For more well-behaved systems, F12-DZ//B3LYP/aug-cc-pVDZ can be applied. In the case of the *o*-C_6H_4 dimerization of D_{2h} symmetry, single-reference descriptions of the potential surface(s) result in an incomplete description of the process, necessitating that a multireference scheme is employed. Here, the PES is computed using Davidson-corrected, explicitly-correlated internally contracted multireference configuration interaction (MRCI-F12+Q) single-point energy calculations done on MCSCF-optimized geometries (MRCI-F12+Q/cc-pVDZ-F12//MCSCF/aug-cc-pVDZ)[33,34,40,43–47] with a (6e,6o) active space for both the ground state as well as selected excited states relevant to the association. This includes 1 orbital each in a_g, b_{2u}, b_{1g}, b_{1u}, b_{2g}, and b_{3g} symmetries. The MCSCF wavefunction is state-averaged with the two lowest-lying singlet states in each of the D_{2h} irreducible representations except for A_g which had three. Molecular orbital (MO) calculations are used for some points in PESs to better understand the nature of processes occurring on the potential surface. All MO calculations are performed with RHF (for singlet species) or ROHF (for triplet species) with the aug-cc-pVDZ basis set in GAUSSIAN16.

2.3 Quartic Force Fields

Accurate anharmonic rovibrational data can be generated computationally through a quartic force field (QFF). A QFF is a fourth-order Taylor series expansion of the potential portion of the internuclear Watson Hamiltonian.[48,49] All QFFs in this work are computed through the software PBQFF,[50] a package that streamlines the calculation of necessary single-point-energies, subsequent generation of force constants through central finite differences, and usage of a modified version of SPECTRO[51] to generate frequencies through vibrational perturbation theory.[52–54]

Spectra for $(C_6H_4)_2$ and $C_{12}H_8$ are generated through QFFs that utilize reparameterized versions of the semiempirical method PM6[55] within MOPAC[56] as previously implemented by Westbrook *et al.*[57 58,59]. This reparameterization is done through an optimization of the majority of the default PM6 parameters using the Levenburg-Marquardt algorithm to minimize the difference between PM6-computed anharmonic frequencies and a set of reference values, **which can be experimental data or from high-accuracy calculations such as those done by explicitly-correltated coupled-cluster theory**. Optimized parameters include the one-center, one-electron integrals (U), two-center, one-electron resonance integrals (β), Slater orbital exponents (Ξ), and the one-center, two-electron repulsion (G) and exchange (H) integrals for both carbon and hydrogen (where applicable, for the latter). An additional term (a) corresponding to the amplitude of the Gaussian function that corrects for the non-polarized core electrons in carbon is also optimized. **These methods are shown to result in a vast improvement upon the computation of frequencies for certain vibrational modes, such as those for C-H stretching modes.**[57] Often, species are chosen for the reparameterization process that are chemically similar to the target molecules for the modified PM6 QFF. For this work, *o*-C_6H_4 and cyclobutadiene (*c*-C_4H_4) are chosen as training species. **For very large species, this method has the ability to drastically lower the necessary computational time for anharmonic rovibrational data collection while potentially providing a greater degree of accuracy compared to methods of a similar computational cost by fitting the existing parameters to data for chemically relevant species.**

Several QFFs at different displacement sizes and levels of theory for both *o*-C_6H_4 and *c*-C_4H_4 are computed in order to determine the best fit to existing Ne[60] and Ar[61–63] matrix data for the IR-active vibrational modes. For *o*-C_6H_4, QFFs are computed at the F12-DZ level of theory, with and without core-electron correlation included, at the 0.010 and 0.015 Å step sizes. For *c*-C_4H_4, F12-DZ and F12-TZ QFFs, both including core-electron correlation, are produced at the 0.005 and 0.010 Å step sizes. The mean absolute error (MAE) between the computed and corresponding IR-active gas-phase experimental frequencies is used to select the most accurate frequency set. The full set of computed frequencies is used for the semiempirical refitting. Six sets of parameters are generated. Parameters are trained on the frequency sets of *o*-C_6H_4 and *c*-C_4H_4 alone as well as in tandem. For each scenario two initial parameter sets are used: the default PM6 parameters and parameters previously trained[57] on C_2H_4.

QFFs for $(C_6H_4)_2$ and $C_{12}H_8$ are computed for each parameter set with a step size of 0.005 Å. B3LYP/N07D anharmonic intensities computed in GAUSSIAN16 are paired with the fundamental frequencies, first overtone of each fundamental, and all two-quanta combination bands. Subsequently, a broadening of 3.0 cm^{-1} is applied to generate the final spectra.

2.4 Anharmonic Cascade Emission Spectra

Simulation of the anharmonic cascade emission spectra of benzopentalene and biphenylene begins with optimization of the geometry and computation of the normal modes and harmonic vibrational frequencies at the B3LYP[30]/N07D[64] level of theory using Gaussian 16[38]. The N07D basis set is derived from the 6-31G basis with additional diffuse and polarization functions added that have been shown to produce accurate geometries and vibrational spectra of large aromatic molecules.[65] The geometry optimization uses an energy convergence threshold of 1×10^{-12} and a custom integration grid consisting of 200 radial shells and 974 angular points per shell.[66] Following this, the QFF of each molecule is computed at the same level of theory. Second order vibrational perturbation theory (VPT2[49,52,54,67,68]) is then used to compute the anharmonic frequencies with the software program SPECTRO.[69] SPECTRO utilizes resonance polyad matrices[70–72] in the VPT2 treatment, allowing for the advanced treatment of resonance interactions through symmetry-separated matrices based on frequency separation. The polyad matrices also allow for the redistribution of intensities between coupled states, which has been shown to be important in densely populated vibrational spectra.[28,66,72–74] States with a separation of less than or equal to 200 cm^{-1} are included in the resonance polyads.[75,76]

In space, PAH emission occurs in the infrared following the absorption of a UV photon that causes electronic excitation, followed by internal conversion to the electronic ground state, and intramolecular vibrational energy redistribution (IVR) to populate the vibrational modes.[77] Single IR photons of discreet energy are emitted in subsequent fashion until the internal energy is depleted.

The process of IR emission is modeled using a straightforward cascade emission process that has been described previously in the literature.[28,78–80] In short, the probability of emitting an IR photon at a given internal energy is proportional to the magnitude of the vibrational frequency of the corresponding normal mode multiplied by the energy-dependent emission at the given internal energy. An IR photon from the anharmonic frequency calculations described above is chosen based on this probability, resulting in energy loss equal to the photon energy. This process uses VPT2 to compute the anharmonicity constants via the equation:

$$E(n) = \sum_k \omega_k (n_k + \frac{1}{2}) + \sum_{k\leq l} x_{kl} (n_k + \frac{1}{2}) \times (n_l + \frac{1}{2}), \qquad (1)$$

where ω represents the harmonic frequency and n represents the number of quanta in a vibrational mode. Transition energies between adjacent vibrational energy levels of a given mode are then

calculated by:

$$\Delta E^{'(k)}(\{n\}) = \omega_k + 2x_{kk}(n_k) + \frac{1}{2}\sum_{i\neq k} x_{ik} + \frac{1}{2}\sum_{i\neq k} x_{ik} n_i, \qquad (2)$$

where n_k is relative to the upper state and the n_i represents the 'spectator' mode that is populated but is not involved in the transition. The total remaining internal energy within the molecule is incorporated by the vibrational modes that are populated, n_i. At each new internal energy following photon emission, the energy-dependent emission spectrum is recalculated and given a new set of emission probabilities. This is repeated until the molecule has relaxed to its vibrational ground state. A Monte Carlo approach is utilized that repeats the entire process to improve accuracy and yield the final cascade emission spectrum. The details of the cascade emission simulations have been validated in the previous papers.[28,72,75,81,82]

3 Results and Discussion

3.1 Reaction Pathways

3.1.1 *ortho*-Benzyne Dimerization

The *o*-C_6H_4 dimerization is first modeled as a D_{2h}-symmetry approach, where the intermolecular distance used is that between the two pairs of acetylenic carbons (r). Multireference PESs of the $(C_6H_4)_2$ dimer ground state ($1\ ^1A_g$, black) and first singlet excited state ($2\ ^1A_g$, blue), as shown in Fig. 1, reveal the presence of an avoided crossing at $r = r_{AC} \approx 2.6$ Å between the $1\ ^1A_g$ and $2\ ^1A_g$ states with the triplet ground state ($1\ ^3B_{3g}$, green) lying between them. The $1\ ^1A_g$ state adiabatically correlates to the ground state singlet products *o*-C_6H_4 ($X\ ^1A_1$) + *o*-C_6H_4 ($X\ ^1A_1$) and diabatically correlates to the first excited singlet products *o*-C_6H_4 ($X\ ^1A_1$) + *o*-C_6H_4 ($1\ ^1A_2$)/($1\ ^1B_1$). An explanation regarding the identity of the first singlet excited state is provided later in this section. As the two *o*-C_6H_4 molecules approach, the $1\ ^1A_g$ state is repulsive, with a higher orbital occupation and lower energy for the C-C antibonding (σ^*) orbital than for the respective bonding (σ) orbital; higher occupancy in the σ orbital is necessary to form the biphenylene cyclic-C_4 moiety. In contrast, the $2\ ^1A_g$ state populates the σ orbital during the association. The higher relative energy of this orbital at large intermolecular distances is due to the loss of intermolecular interaction that serves to stabilize that orbital in the excited state. Procession through the avoided crossing effectively swaps the frontier orbital occupations between the $1\ ^1A_g$ and $2\ ^1A_g$ states. Within the region of the reaction coordinate, $r_e \leq r < r_{AC}$, the $1\ ^1A_g$ state is able to engage in bonding. The repulsive behavior caused by the occupancy of σ^* is taken up by the $2\ ^1A_g$ state and likely continues into higher-lying excited states at $r \approx 2.0$ Å.

For the D_{2h}-symmetry complex to associate to form $(C_6H_4)_2$ as depicted in Fig. 1, it would need to be excited into the $2\ ^1A_g$ state or proceed over the avoided-crossing-induced barrier on the $1\ ^1A_g$ surface. The energy needed to excite $2\ ^1A_g \leftarrow 1\ ^1A_g$ at $r = 5.0$ Å is 96.9 kcal mol^{-1}/4.20 eV at the MRCI-F12+Q/cc-pVDZ-F12//MCSCF/aug-cc-pVDZ level of theory, which corresponds roughly to excitation of one equivalent of *o*-C_6H_4 into its lowest excited state. The exact identity of this excited state is highly sensitive to the choice of active space used to locate it in an MCSCF/aug-cc-pVDZ calculation, with the ordering of the energies of the $1\ ^1A_2$ and $1\ ^1B_1$ states dependent upon the number of virtual B_1 orbitals included in the active space. Direct procession over the barrier on the $1\ ^1A_g$ surface requires 19.2 kcal mol^{-1}/0.83 eV of energy. In an interstellar context, both processes are likely rare. The energy cost of either excitation is high enough to negate this highly symmetric association as a significant contributor to the formation of $(C_6H_4)_2$ in the ISM.

More promising, however, the dimerization of *o*-C_6H_4 can also be modeled as a C_{2h} approach, shown in Fig. 2, that only allows *one* acetylenic carbon on each *o*-C_6H_4 molecule to interact. **Here, the C_2(x) and C_2(y) axes of symmetry have been removed in order to allow the two *o*-C_6H_4 monomers to approach at a different orientation.** The PESs for the lowest-energy spin-paired-singlet configuration and five triplet configurations including the repulsive $1\ ^3A_g$ (cyan) state are computed for these PESs. The singlet minimum in Fig. 2 is found to have a remarkably low singlet-triplet splitting of 2.3 kcal mol^{-1}, **with the bulk of the difference being carried in the zero-point correction of the triplet, which is 1.7 kcal mol^{-1} greater than that of the singlet. The low singlet-triplet splitting** is likely a result of how the dimerization at C_{2h} symmetry impacts the orbitals of *o*-C_6H_4 - interaction of one acetylenic carbon per molecule to form the resulting σ bond leaves two electrons to remain in the same plane on the two carbons that are not involved in bonding. **Another near-degeneracy is observed between the $1\ ^3B_u$ and $1\ ^3A_g$ states. As the intermolecular distance for these states approaches 5.0 Å, the relative energy of the structures with respect to the $1\ ^1A_g$ approaches the *o*-C_6H_4 singlet-triplet splitting, indicating that both states form through interaction between the $\tilde{X}\ ^1A_1$ and $\tilde{a}\ ^3B_2$ states of *o*-C_6H_4. Further interaction between the two species under the constraint of C_{2h} symmetry breaks this degeneracy, and results in the $1\ ^3B_u$ undergoing favorable bonding interactions while the $1\ ^3A_g$ exhibits repulsive behavior.**

On the $1\ ^1A_g$ (black) PES, as the intermolecular separation **decreases to** ~2.5 Å, the F12-DZ single-point energies begin to experience strong multireference effects**, which are most intense at the minimum, indicated by a large doubles amplitude for the excitation between the highest-occupied molecular orbital (HOMO) and the lowest-occupied molecular orbital (LUMO) in the CCSD calculation. At greater intermolecular separations, these effects are largely absent. Interestingly, the opposite is true for the $1\ ^3B_u$ (magenta) state - at large intermolecular distances, the scan experiences multireference effects which correlate with the presence of the nearly-degenerate $1\ ^3A_g$ state as discussed above. However, these are largely alleviated as the intermolecular distance decreases towards the portion of the scan that is nearly degenerate with that of the $1\ ^1A_g$ state. As indicated by the structure of the minimum and the low singlet triplet-splitting, the structure of the complex at small intermolecular distances is well-described by the $1\ ^3B_u$ PES. This is supported by the frontier molecular orbitals of the two minima as shown in Fig. 3. Both the HOMO and LUMO of the singlet, as well as the two HOMOs (both singly occupied) of the low-lying triplet localize much**

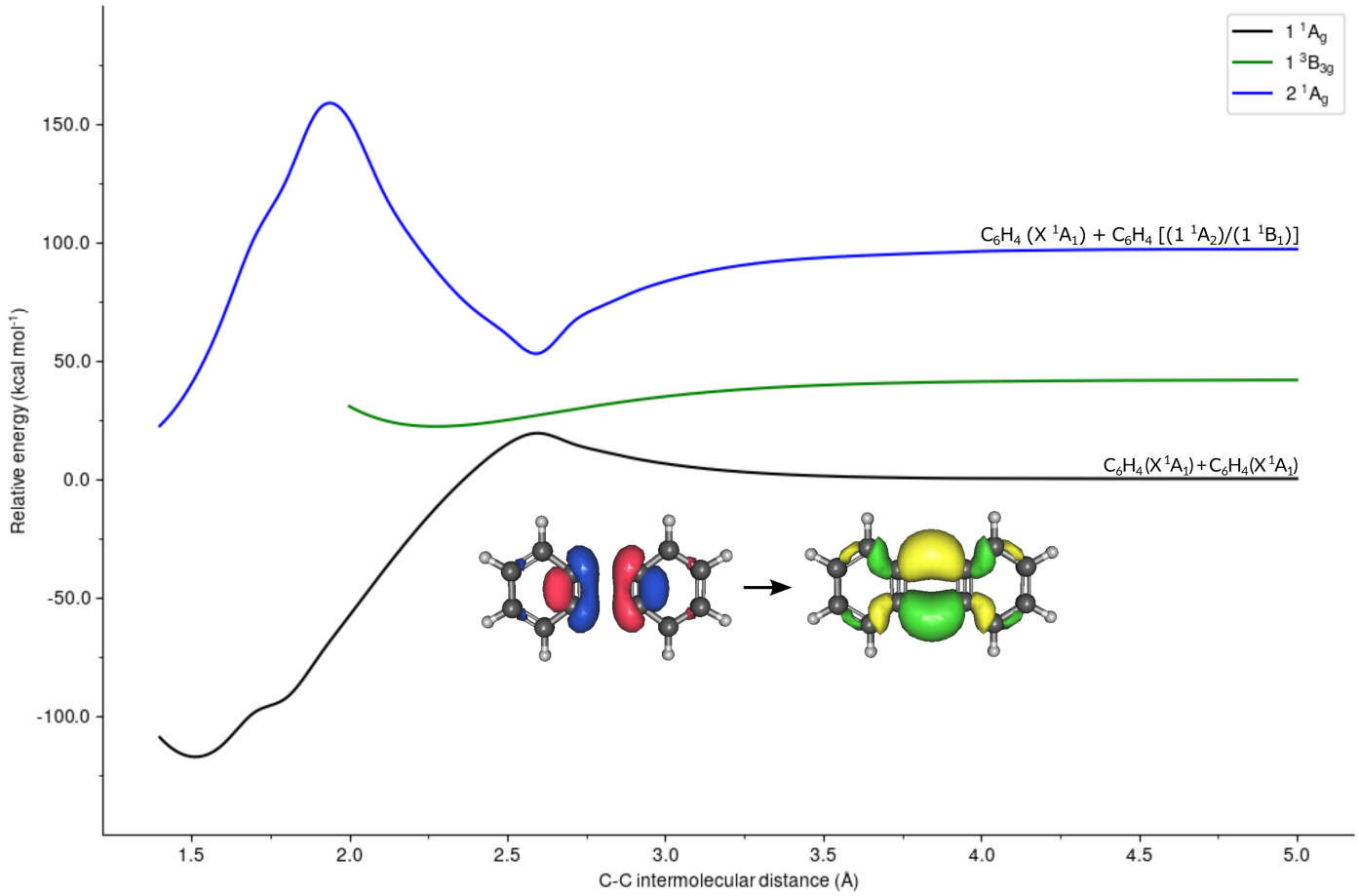

Fig. 1 MRCI-F12+Q/cc-pVDZ-F12//MCSCF/aug-cc-pVDZ potential energy surfaces for the two lowest-energy ^1A_g complexes and lowest-energy triplet complex for the dimerization of *o*-C_6H_4 along a D_{2h} reaction coordinate. Energies are relative to that of the 1 ^1A_g state (black) at 5.0 Å intermolecular separation. Displayed orbitals reflect the less-occupied (bottom) and more-occupied (top) frontier orbitals prior to the crossing point

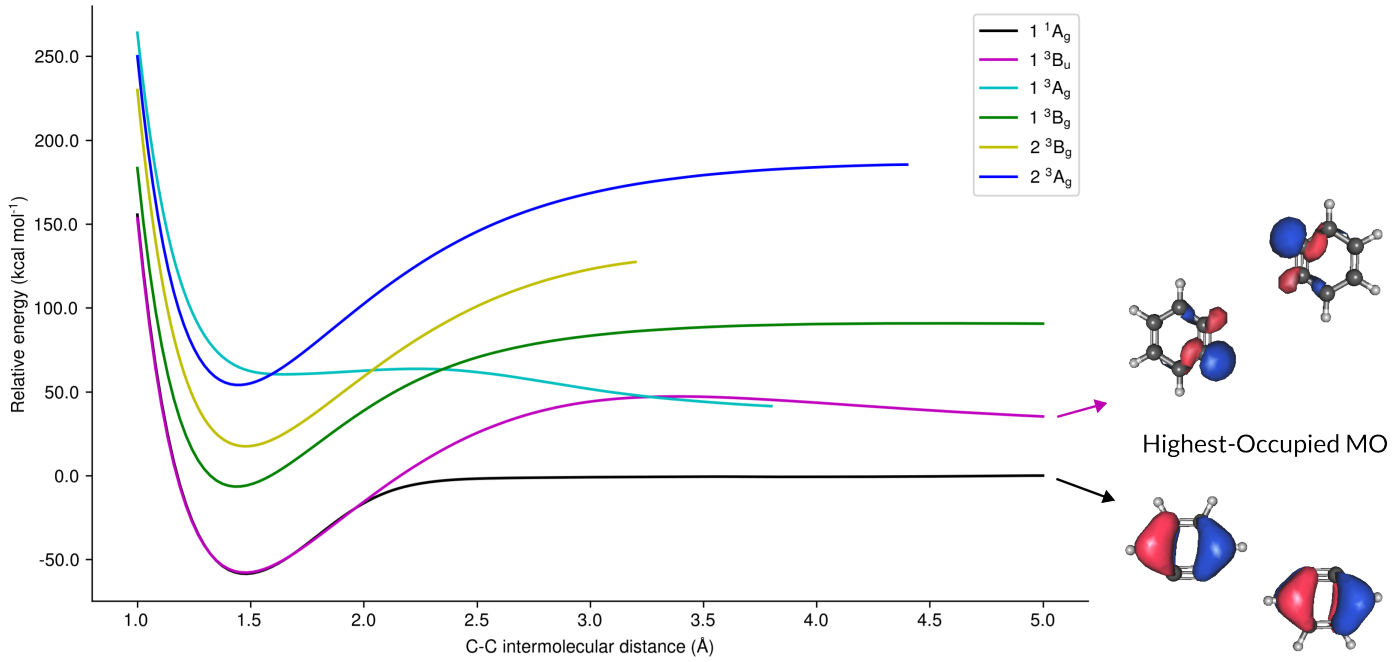

Fig. 2 Potential energy surfaces for the dimerization of o-C_6H_4 under the constraint of C_{2h} symmetry. Energies are relative to that of the singlet complex at 5.0 Å intermolecular separation.

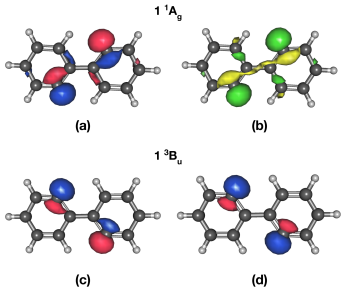

Fig. 3 Top: the HOMO (a) and the LUMO (b) of the 1 ^1A_g state of the C_{2h} o-C_6H_4 dimer. Bottom: the HOMO-1 (c) and HOMO (d) of the 1 ^3B_u state of the C_{2h} o-C_6H_4 dimer.

of the density onto the nonbonding carbon sites. As is evident from Fig. 3(b), population of the singlet LUMO would contribute to stronger bonding along the reaction coordinate. In fact, at lower isovalues of the orbital in Fig. 3(d), the electron density opposite the blue lobes is shown to extend through the created σ bond much like that in Fig. 3(b). As the CCSD calculation indicates, population of such an orbital results in increased stability. Additionally, there exists a section of the PESs in Fig. 2 within the range 1.7 Å $\leq r_{C-C} \leq$ 1.9 Å where the electronic energy E_e of the triplet becomes lower than that of the singlet. The overall behavior of both electronic states suggests that an intersystem crossing (ISC) from the 1 ^1A_g state to the 1 ^3B_u state near r_{C-C} = 2.0 Å is likely, and as will be discussed below, favorable for continuation along the reaction pathway.

3.1.2 Formation of Biphenylene and Benzopentalene

Since the formation of the o-C_6H_4 dimer appears to be practical, the formation of biphenylene and its isomer benzopentalene are possible. The reaction pathway shown in Fig. 4 maps out the initial formation of $(C_6H_4)_2$ and its isomerization to $C_{12}H_8$. Additionally, possible routes to the two bimolecular product sets are also computed and appear to be exothermic with respect to the reactants.

The large potential well depth of both species is immediately obvious, which indicates both to be good candidates for radiative stabilization **in cold environs.**[83] Following the initial dimerization of o-C_6H_4 into I1 through ISC (see Fig. 2), rotation around the central C-C bond creates $(C_6H_4)_2$, which must again proceed through **ISC** in order to enter the lower-energy singlet state. **Attempts to locate a transition state for the isomerization between the previously-discussed 1 ^1A_g state of I1 and $\tilde{X}$ ^1A_g $(C_6H_4)_2$ at the B3LYP/aug-cc-pVDZ level of theory fail, prohibiting the alternate pathway that would see I1 form in the 1 ^3B_u state and subsequently proceed through a second ISC. After the formation of $\tilde{X}$ ^1A_g $(C_6H_4)_2$,** two separate, two-step processes are identified for the isomerization into $C_{12}H_8$. The first involves a hydride shift (TS1) prior to a ring contraction (TS4) while the other undergoes those steps (TS3, TS5) in the opposite order through the intermediate I3. Owing to the lower energies of both barriers and the intermediate, the latter sequence will be the primary method of isomerization. The barrier height of TS3 relative to $(C_6H_4)_2$ of 79.7 kcal mol^{-1}, however, is sizable, and will likely impact relative concentrations of $(C_6H_4)_2$ and $C_{12}H_8$ produced through radiative stabilization on the PES.

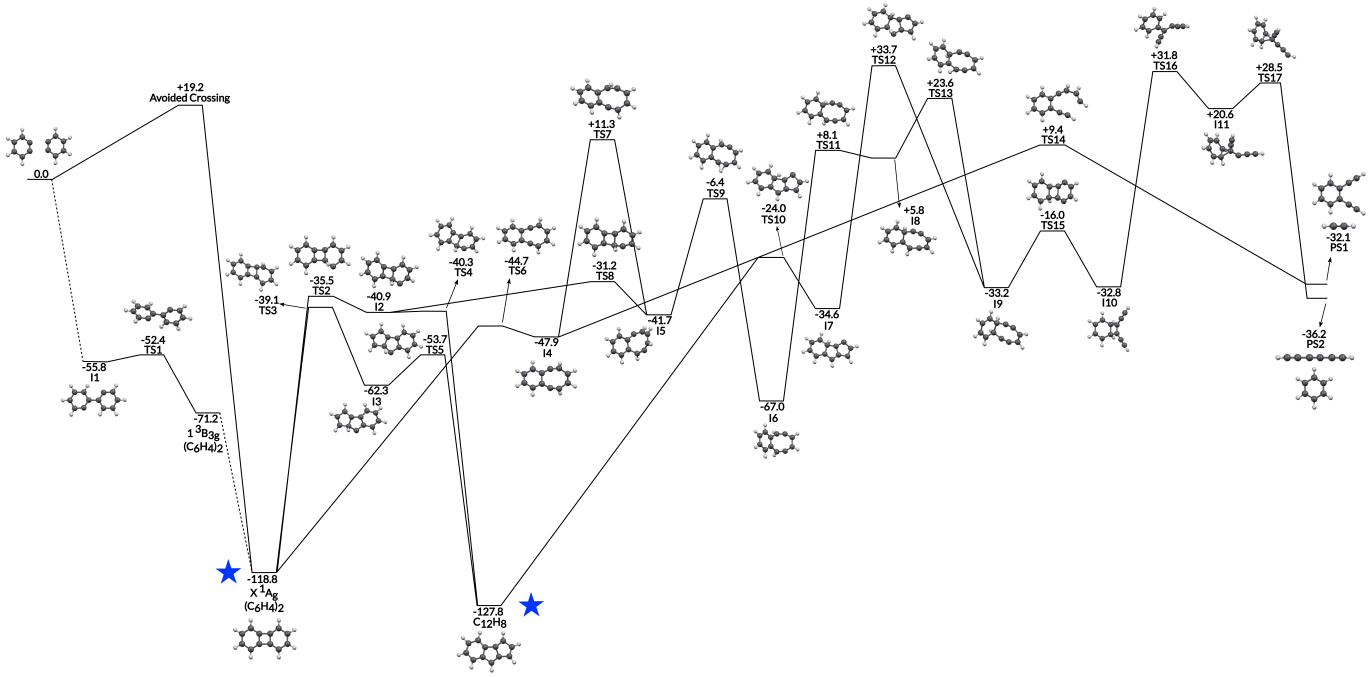

Fig. 4 Reaction pathway for the dimerization of o-C_6H_4 to form $(C_6H_4)_2$, $C_{12}H_8$ (indicated by blue stars), and the bimolecular product sets $C_6H_4(C_2H)_2$ + C_2H_2 and C_6H_6 + l-C_6H_2. Dashed lines indicate transition between electronic states of different spin multiplicity. All values in kcal mol^{-1}.

Following the formation of $(C_6H_4)_2$ and possible isomerization to $C_{12}H_8$, there are no accessible routes to either possible product set without a significant energy input. $(C_6H_4)_2$ can undergo ring expansion (TS6) to I4, but the +9.4 kcal mol^{-1} concerted ring-breaking and C_2H_2 loss (TS14) is likely completely inaccessible under the physical conditions present in TMC-1. Formation of the second product set, C_6H_6 and l-C_6H_2, is even more unlikely, with every route to these species requiring passage over 3-5 non-submerged barriers. As such, the greatest inhibitor to isolation of $(C_6H_4)_2$ or $C_{12}H_8$ upon the successful dimerization of o-C_6H_4 is the potential redissociation into the reactants. As mentioned above, the well depth of both species (as well as the low temperature of TMC-1) will facilitate radiative stabilization. Both $(C_6H_4)_2$ and $C_{12}H_8$ should be considered targets for astronomical observation within TMC-1 and other interstellar environments.

3.2 Spectroscopy

3.2.1 Optimization of Semiempirical Parameters

In order to examine if the previously-discussed species can be observed in TMC-1, protoplanetary disks, or any other astronomical environment, spectral reference data must be available for comparison. To capture the effect of the unique electronic structures of $(C_6H_4)_2$ and $C_{12}H_8$ on their respective spectra generated by PM6, QFFs of o-C_6H_4 and c-C_4H_4 are performed to generate training data for the PM6 reparameterization algorithm, as discussed in §2.3. The computed vibrational frequencies for o-C_6H_4 and c-C_4H_4, as well as available experimental data, are tabulated in Tables 1 and 2, respectively. A slight increase in step size from 0.010 Å to 0.015 Å, as well as the inclusion of core correlation in the calculations, results in a $\sim$0.5 cm^{-1} increase in accuracy. However, the C-H stretching frequencies - among the most intense - improve from a MAE of 8 cm^{-1} to 5 cm^{-1} with the inclusion of core correlation. All QFFs predict ν_{15}, the second-most-intense transition, within 1 cm^{-1} of experiment. ν_{15} corresponds to the concerted C-H out-of-plane motion; the analogous motion of $(C_6H_4)_2$ is the most intense transition.

Three of the four QFFs for c-C_4H_4 compute the IR-active frequencies to within 6 cm^{-1} of experiment. This remarkable accuracy does not translate to the Raman-visible frequencies. While the frequencies of the a_g and b_{3g} C-H stretching modes are predicted within 12 cm^{-1} across all QFFs, the remaining frequencies differ by 30-160 cm^{-1}. This apparent inaccuracy of the frequencies of the Raman-active modes, when considering the accuracy of the IR-active modes in comparison to matrix isolation experiments, is unexpected and may indicate flaws in the original Raman spectrum, which has never been repeated to our knowledge. The availability of this new, high-accuracy computational data for peak assignment within the Raman spectrum may motivate the collection of new experimental data.[63]

The o-C_6H_4 and c-C_4H_4 QFFs with the lowest MAE are taken as the true frequencies for the PM6 reparameterization. As discussed in §2.3, the reparameterization algorithm can be implemented using multiple sets of frequencies at once, often sacrificing accuracy of the prediction of frequencies across a single training molecule in order to capture a greater range of chemical behavior within the parameters. The reparameterization program also allows for the choice of the parameters used to compute training molecule frequencies during the first pass of the algorithm rather than forcing the PM6 default parameters for this purpose. Changing the initial parameter set allows for the program to find alternative - and potentially more well-fit - solutions during the multivari-

Table 1 Computed QFF F12-DZ frequencies for o-C_6H_4 with and without core correlation included at two different step sizes. Experimental frequencies collected using Ne matrix isolation are provided where applicable. Reported MAEs account for the discrepancy between theory and experiment for the frequencies of the IR-active vibrational modes. All values in cm^{-1}

Symmetry	Mode	F12-DZ		F12-DZ + core		Exp.[60]
		0.010 Å	0.015 Å	0.010 Å	0.015 Å	
a_1	ν_1	3088.0	3090.0	3092.9	3094.8	3094
	ν_2	3054.6	3055.2	3060.4	3061.0	3071
	ν_3	1909.5	1909.9	1917.3	1918.5	1846
	ν_4	1436.3	1436.1	1440.8	1440.5	1415
	ν_5	1261.4	1261.1	1264.4	1265.1	1271
	ν_6	1131.5	1131.2	1135.5	1135.0	1055
	ν_7	1042.8	1043.3	1046.4	1046.7	1039
	ν_8	965.3	965.8	968.8	969.2	982
	ν_9	604.3	604.2	608.0	607.2	589
a_2	ν_{10}	1029.8	1019.0	1033.9	1024.0	—
	ν_{11}	858.1	858.0	858.4	858.6	—
	ν_{12}	639.5	636.5	638.8	637.2	—
	ν_{13}	443.1	444.1	444.5	446.9	—
b_1	ν_{14}	927.5	925.6	926.5	924.7	838
	ν_{15}	736.1	736.1	736.2	736.3	737
	ν_{16}	386.5	386.1	388.1	387.5	388
b_2	ν_{17}	3085.5	3087.8	3090.2	3092.6	3086
	ν_{18}	3039.1	3040.0	3045.2	3046.2	3049
	ν_{19}	1435.1	1435.6	1439.6	1440.4	1451
	ν_{20}	1397.3	1396.0	1402.5	1401.5	1394
	ν_{21}	1238.9	1239.6	1241.6	1242.7	1307
	ν_{22}	1082.5	1082.7	1086.1	1086.0	1094
	ν_{23}	840.2	840.9	843.8	845.1	849
	ν_{24}	456.0	457.8	461.3	463.7	472
Mean absolute error		22.8	22.3	22.1	21.7	

Table 2 Computed QFF F12-DZ and F12-TZ frequencies for c-C_4H_4 with core correlation included at two different step sizes. Experimental frequencies collected using Ne and/or Ar matrix isolation are provided where applicable. Reported MAEs account for the discrepancy between theory and experiment for the frequencies of the IR-active vibrational modes. All values in cm^{-1}.

Symmetry	Mode	F12-DZ + core		F12-TZ + core		Exp. (IR)	Exp. (Raman)[84]
		0.005 Å	0.010 Å	0.005 Å	0.010 Å		
a_g	ν_1	3132.1	3131.1	3128.8	3134.3	—	3140
	ν_2	1510.3	1508.4	1510.4	1516.5	—	1678
	ν_3	1094.5	1090.8	1078.7	1082.7	—	1059
	ν_4	928.0	927.1	911.4	930.5	—	989
b_{1g}	ν_5	850.7	845.0	822.1	834.3	—	—
b_{2g}	ν_6	606.1	594.8	583.6	579.9	—	531
b_{3g}	ν_7	3081.2	3081.6	3080.0	3082.8	—	3093
	ν_8	1174.6	1173.9	1155.1	1172.2	—	—
	ν_9	859.4	853.0	843.1	852.3	—	723
a_u	ν_{10}	936.1	922.0	816.6	840.8	—	—
	ν_{11}	538.8	524.4	476.6	513.2	—	—
b_{1u}	ν_{12}	3121.3	3120.8	3121.6	3123.8	3124[63]	—
	ν_{13}	1539.5	1538.8	1535.6	1543.0	1527[62]	—
	ν_{14}	1031.6	1030.0	1009.5	1029.1	1028[61]	—
b_{2u}	ν_{15}	3094.9	3094.6	3093.0	3096.5	3105[61]	—
	ν_{16}	1249.3	1245.3	1232.2	1245.0	1244[62]	—
	ν_{17}	715.9	713.4	687.4	716.6	719[61]	—
b_{3u}	ν_{18}	573.8	568.7	533.7	565.4	576[62]	—
Mean absolute error (IR)		5.6	5.9	18.2	5.7		

Training Set	PM6 Default			C_2H_4-optimized		
	RMSD (cm^{-1})	Max. Error (cm^{-1})	Index	RMSD (cm^{-1})	Max. Error (cm^{-1})	Index
o-C_6H_4	131.2	333.0	(a)	26.5	68.1	(b)
c-C_4H_4	69.9	215.2	(c)	43.3	161.7	(d)
o-C_6H_4 + c-C_4H_4	64.9	158.8	(e)	54.8	143.4	(f)

Table 3 RMSD and maximum error across all frequencies for the semiempirical parameter refitting on the o-C_6H_4, c-C_4H_4, and combined core-correlated F12-DZ training sets. Results are separated between training initialization on the default PM6 parameters and parameters for C_2H_4 previously computed by Westbrook *et al.*[57]. All values in cm^{-1}.

ate error minimization. The two most accurate sets of frequencies computed for *o*-C_6H_4 and *c*-C_4H_4 as listed in Tables 1 and 2 are subsequently transformed into three groups: two groups each containing training frequencies for the individual species and one group combining both frequency sets. Each group produces two new parameter sets through reparameterizations initialized on the default PM6 parameters and previously-reparameterized values computed for C_2H_4, resulting in six total sets of parameters.

The results of each fitting are compiled in Table 3 and are defined as index/parameter set (a)-(f). Training the default PM6 parameters on only the frequencies of *o*-C_6H_4 results in the largest single-frequency error as well as the largest root-mean-squared deviation (RMSD = $[(\nu_n - \nu_n')^2/n]^{1/2}$ where ν_n' is the predicted value). Among the largest contributors to the large RMSD of parameter set (a) are the errors in the prediction of the C-H stretching frequencies. Parameter set (a) has an MAE of 147.5 cm^{-1} among the four C-H stretching frequencies alone. High error within these frequencies also plagues parameter set (c), although to a lesser degree, with an MAE of 46.6 cm^{-1} throughout the C-H stretching frequencies. All other parameter sets predict the C-H stretching frequencies rather accurately, with MAEs ranging from 6.8-12.3 cm^{-1}. The maximum error incurred by the training of parameter sets (c) and (d) is attached to the lowest core-correlated F12-DZ frequency (573.8 cm^{-1}) of *c*-C_4H_4, and may affect resulting spectra less, as few vibrational frequencies are expected at such long wavelengths. Outside of the error in the lowest frequency of *c*-C_4H_4, the largest single-frequency errors in parameter sets (c) and (d) are 84.8 and 42.7 cm^{-1}, respectively. In the mixed-molecule parameter sets (e) and (f), the largest single-frequency error is also attached to the 573.8 cm^{-1} frequency of *c*-C_4H_4. Omitting this frequency, several other frequencies are predicted with errors above 100 cm^{-1}. Interestingly, parameter set (e) does not suffer the large errors in C-H frequency prediction as do parameter sets (a) and (c). Additionally, there is a shift in how the large-intensity concerted C-H out-of-plane bends are predicted during the mixed-molecule reparameterizations. Parameter sets (a) - (d) predict this frequency for their respective species with a consistent 22.9-36.0 cm^{-1} absolute error. In contrast, parameter sets (e) and (f) predict the concerted C-H out-of-plane bend for *o*-C_6H_4 with absolute errors of 1.0 and 15.4 cm^{-1}, respectively, while the frequency for the same motion of *c*-C_4H_4 is predicted with an error of 154.0 cm^{-1} for parameter set (e) and 60.5 cm^{-1} for parameter set (f). These latter two parameter sets may result in the prediction of incredibly accurate (or inaccurate) concerted out-of-plane bend frequencies depending on the exact structure of the species under study. While the errors measured during the reparameterization process for *o*-C_6H_4 and *c*-C_4H_4 are larger than those measured between explicitly-correlated QFF frequencies and matrix-isolation values, they may enable accurate computation of frequencies for large polycyclic hydrocarbons, as examined in the cases of $(C_6H_4)_2$ and $C_{12}H_4$ below.

3.2.2 Vibrational Spectra

The PM6-QFF spectra for $(C_6H_4)_2$ generated from each of the six parameter sets are shown in Fig. 5 in cyan. Figs. 5(a)-(f) contain the gas-phase experimental spectrum made available by NIST[24,25] (black) as well as the spectrum generated from the B3LYP/N07D anharmonic frequency calculation (magenta). With respect to experiment, the B3LYP/N07D spectrum overpredicts by only 3 cm^{-1} the location of the major C-H stretch at 3074 cm^{-1} as well as the concerted C-H out-of-plane bend at 730 cm^{-1}. However, B3LYP/N07D does not capture the shoulder of the C-H stretching feature at 3026 cm^{-1}. The majority of the gas-phase features from 1500-750 cm^{-1} are also predicted fairly accurately by B3LYP/N07D, such as the 1426 cm^{-1} mixed C-H bend/C-C stretch mode (B3LYP/N07D: 1433 cm^{-1}), the C-C breathing mode at 958 cm^{-1} (B3LYP/N07D: 973 cm^{-1}) and the nonconcerted C-H out-of-plane bend at 910 cm^{-1} (B3LYP/N07D: 913 cm^{-1}). The exceptions to the overall remarkable performance of B3LYP/N07D are the predicted shape of the feature at 1146 cm^{-1}, which is bisected by B3LYP/N07D into two peaks at 1164 and 1132 cm^{-1}, and the exact location of the experimental 1254 cm^{-1} feature, which is relocated to 1286 cm^{-1} in the B3LYP results. The accuracy of the C-H bend overtones from 2000-1500 cm^{-1} is striking, with the three highest-frequency (and highest-intensity) overtones at 1922 cm^{-1}, 1874 cm^{-1}, and 1826 cm^{-1} predicted to within 1 cm^{-1} of experiment by B3LYP/N07D. For small-to-medium polycyclic hydrocarbons like $(C_6H_4)_2$ that exhibit some degree of antiaromatic behavior, anharmonic frequency calculations at the B3LYP/N07D level of theory appear to be a good approximation for the true gas-phase spectrum. The strong agreement between B3LYP/N07D and experimental spectra has been observed for other polycyclic species such as phenanthrene, pyrene, and pentacene,[73] as well as the standard and deuterated forms of the monocyclic species phenylacetylene.[74,75]

The frequencies computed from reoptimized PM6 parameters initialized on the parameters previously optimized for C_2H_4 (panels (b), (d), (f) in Fig. 5) match the location of the C-H stretch feature to within 30 cm^{-1}, as do the frequencies resulting from parameter set (e). Of these, parameter sets (b) and (d) allow the most accurate prediction for the major portion of this feature at 3072 cm^{-1} and 3069 cm^{-1}, respectively. However, as is the case for the B3LYP/N07D calculation, no parameter set results in accurate prediction of the C-H stretch "shoulder" feature. **Most parameter sets [(b)-(f)] predict what appears to be a shoulder detached from the main C-H stretching feature, and, with the exception of parameter set (c), all QFFs indicate these shoulders to be the result of weak combination bands. As the identities of the vibrations that give rise to the shoulder observed in the experimental spectrum are unknown, the absence of the shoulder at the correct position or intensity in all of the computed spectra could be attributed to several sources of error. In particular, the C-H stretching fundamental modes may be too highly clustered, the location of the relevant combination bands may be predicted incorrectly due to a propagation of error from a fundamental band that contributes to a combination band, or there may be an underprediction of the relevant infrared intensities.**

Parameter sets (a), (d), and (f) all predict the location of the intense 730 cm^{-1} peak to within 13 cm^{-1}, with frequencies from the other parameter sets erring by over 50 cm^{-1}. The C-H bend overtones from 2000-1500 cm^{-1} are most accurately computed

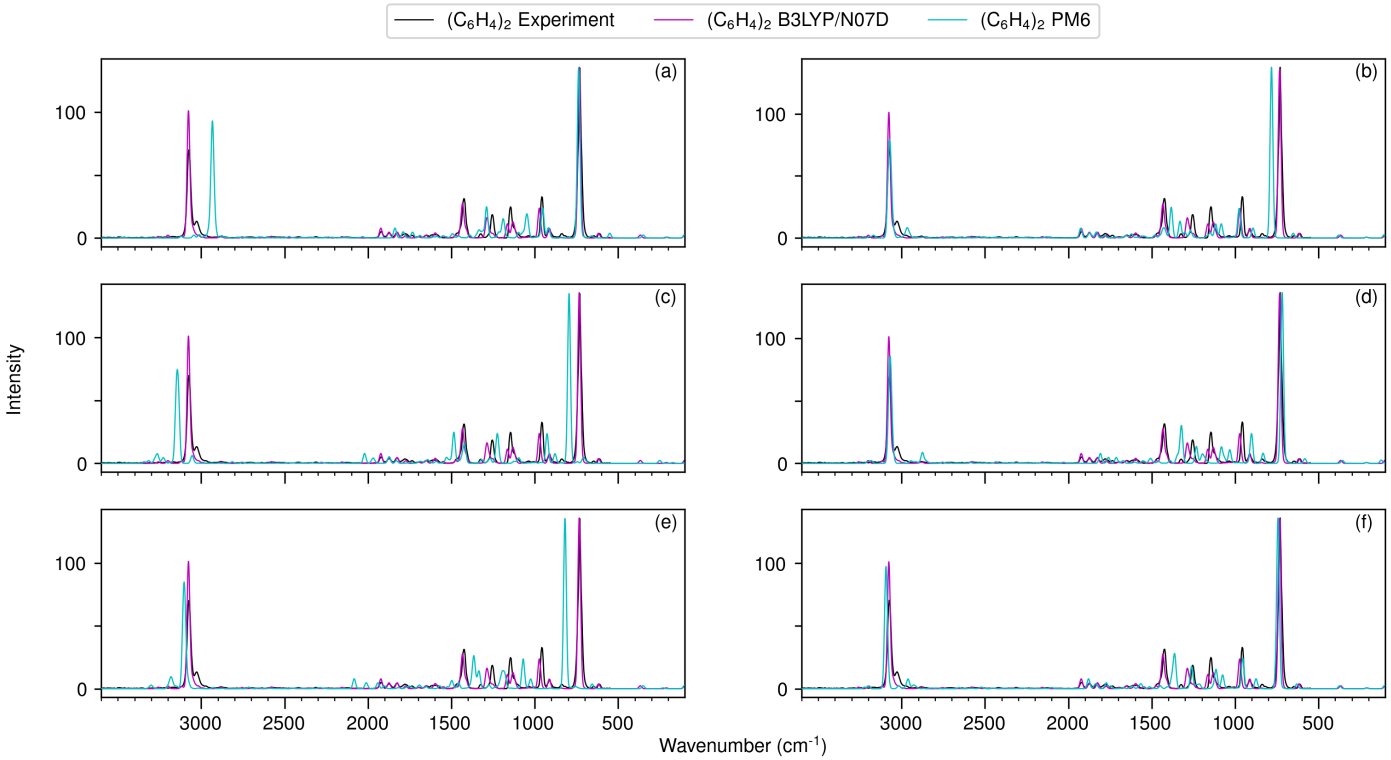

Fig. 5 Simulated spectra for biphenylene from 3600-100 cm^{-1}. Each spectrum contains both the experimental and B3LYP/N07D anharmonic reference spectrum. The intensity values of the experimental spectrum are scaled such that the peaks from each PM6 QFF spectra and the 730 cm^{-1} experimental feature appear equal in intensity. The data points for the experimental spectrum are from NIST [24,25]. Intensity values for the B3LYP/N07D and PM6 spectra are in units of kcal mol^{-1}. The index in the top right corner of each spectrum indicates the parameter set (see Table 3) used for the PM6 QFF calculation.

by parameter set (b), with all others resulting in the collective dislocation of these peaks relative to the experimental spectrum to some degree. No parameter set computes the entire 1500-750 cm^{-1} region as accurately as B3LYP/N07D. However, parameter sets (a) and (f) predict the frequencies of the C-C breathing mode to within 6 and 4 cm^{-1}, respectively, and parameter set (f) predicts the 1254 cm^{-1} C-C stretching motion at 1261 cm^{-1}, a more accurate prediction than B3LYP/N07D. In general, many of the PM6 QFFs fall victim to compression or expansion of the majority of the 2000-750 cm^{-1} region or reduced frequency density that leads to bisected or trisected features, with the features arising from parameter set (f) being the least prone to this issue. Promisingly, the two most intense features within the experimental spectrum are predicted relatively accurately in several of the QFFs, with parameter sets (d) and (f) resulting in frequencies with the lowest combined error for these peaks. QFFs created from optimized parameter sets initialized with previously-optimized C_2H_4 parameters predict the 3074 cm^{-1} and 730 cm^{-1} more accurately than their default-initialized counterparts. The MAEs for these peaks computed using parameter sets (b), (d), and (f) are 9 and 26 cm^{-1}, respectively, while they grow to 80 and 53 cm^{-1} for parameter sets (a), (c), and (e). **Use of these reparameterized versions of PM6 for calculations on much larger polycyclic antiaromatic hydrocarbons has the potential to retrieve reasonably accurate infrared spectra if the C-H stretching feature and the concerted C-H out-of-plane bend remain the dominant features of the spectrum and may be particularly relevant to the prediction of emission spectra of much larger species where low-frequency modes, such as the C-H out-of-plane bend, are expected to be dominant.**

As $C_{12}H_8$ has no inversion center, it harbors a permanent dipole. However, as is common in PAHs, the dipole moment computed at the B3LYP/aug-cc-pVDZ level of theory is rather small, only 0.28 D. Vibrational data is thus necessary if $C_{12}H_8$ is to be observed in the ISM or elsewhere. The PM6 QFFs for $C_{12}H_8$ generated from each of the six parameter sets are shown in Fig. 6. As the experimental spectrum of $(C_6H_4)_2$ is well-replicated with the spectrum simulated with B3LYP/N07D anharmonic frequencies, the B3LYP/N07D frequencies are again displayed here as a proxy for the experimental spectrum. **Comparatively, the spectra displayed in Fig. 6 have a greater number of features than those within Fig. 5. This is primarily due to the greater number of infrared-active modes for species of C_s symmetry versus those of D_{2h} symmetry. While this has the possible consequence of complicating the accurate generation of spectra, several features are well-predicted, among them the C-H stretching feature.** Four of the six parameter sets predict the 3071 cm^{-1} B3LYP/N07D C-H stretching peak within 21 cm^{-1}. Additionally, parameter sets (e) and (f) capture the higher- and lower-wavelength, respectively, shoulder peaks computed by B3LYP/N07D. Due to the lower symmetry of $C_{12}H_8$ relative to $(C_6H_4)_2$, however, the 2000-500 cm^{-1} region of the spectra in Fig. 6 is much more densely populated. The aggregated C-C stretching feature, centered by B3LYP/N07D at 1622 cm^{-1}, is often redshifted in the PM6 calculations, and the degree of this shifting is an indicator of the overall accuracy of the predicted frequencies from 2000-500 cm^{-1}. Two parameter sets, (c) and (e), predict the C-C strteching feature within 23 cm^{-1}, while (b) and (f) predict frequencies that err by just over 40 cm^{-1} each. Parameter sets (a) and (d), however, seem to have displaced this feature and most subsequent features as a result, and thus err the greatest with respect to the 1622 cm^{-1} peak (MAE = 142 cm^{-1}) and the most intense peak in the spectrum at 674 cm^{-1} (MAE = 87 cm^{-1}). This redshifting is also present for these spectra in Fig. 5 but does not impact the intense C-H out-of-plane bending feature as it does in Fig. 6. In light of this, discussion of the spectra generated by parameter sets (a) and (d) will be omitted hereafter.

The in-plane C-H wagging centered on the benzene group, predicted by B3LYP/N07D at 1439 cm^{-1}, is generally underpredicted by approximately 30-50 cm^{-1} by all parameter sets with exception of (e), where the feature is bisected into two peaks at 1466 and 1420 cm^{-1}. In contrast, the in-plane C-H wagging motion centered on the pentalene group is generally well predicted by parameter sets (b), (c), (e), and (f), with errors ranging from 11 (spectrum (b)) to 26 (spectrum (e)) cm^{-1}. The intense features centered at 849 cm^{-1}, 750 cm^{-1}, and 674 cm^{-1} all correspond to C-H out-of-plane bending motions, and are predicted by the parameter sets with varying degrees of accuracy. The 849 cm^{-1} is predicted at 866 cm^{-1} by parameter set (c) with a similar intensity. While parameter set (b) predicts this feature with similar accuracy and spectrum (f) places it within 4 cm^{-1} at 853 cm^{-1}, neither capture the intensity of the peak, and the shape of the feature predicted by parameter set (e) misses the mark entirely. Interestingly, the 750 cm^{-1} peak is the most complex of the three but is broadly computed well, with each of the four parameter sets distinguishing at least one of the two shoulder features (located at 785 and 724 cm^{-1}) from the main peak. Parameter set (b) predicts this peak center at 740 cm^{-1}, with a doubly-peaked left shoulder centered at 785 cm^{-1}. Parameter set (c) predicts well the relative peak placement within the entire feature, but redshifts the entire feature $\sim$30 cm^{-1}. Parameter set (e) blueshifts the major peak by 19 cm^{-1}, likely overtaking the left shoulder, but very accurately predicts the right shoulder at 725 cm^{-1}. Parameter set (f) predicts the existence of both shoulders but produces the largest error for the peak center (-42 cm^{-1}) and overpredicts the interpeak separation of the feature. The intense concerted C-H out-of-plane bend at 674 cm^{-1} is not predicted as accurately as it is for $(C_6H_4)_2$, with parameter sets (b) and (e) most accurately predicting this feature at 651 and 650 cm^{-1}, respectively. Overall, parameter sets (b) and (e) most closely reproduce the B3LYP/N07D spectrum of $C_{12}H_8$.

Parameter sets (b), (e), and (f) best represent the existing experimental data for $(C_6H_4)_2$ and the B3LYP/N07D spectra for $C_{12}H_8$. Parameter set (f) largely outperforms the others for $(C_6H_4)_2$, but all three parameter sets compare similarly well with the B3LYP/N07D spectrum for $C_{12}H_8$. The three corresponding parameter sets have the three lowest maximum errors in frequency prediction during parameter training as shown in Table 3 as well as three of the four lowest RMSD values for the fittings. Interestingly, the generated parameter sets that are affected by large errors in the prediction of the C-H stretching frequen-

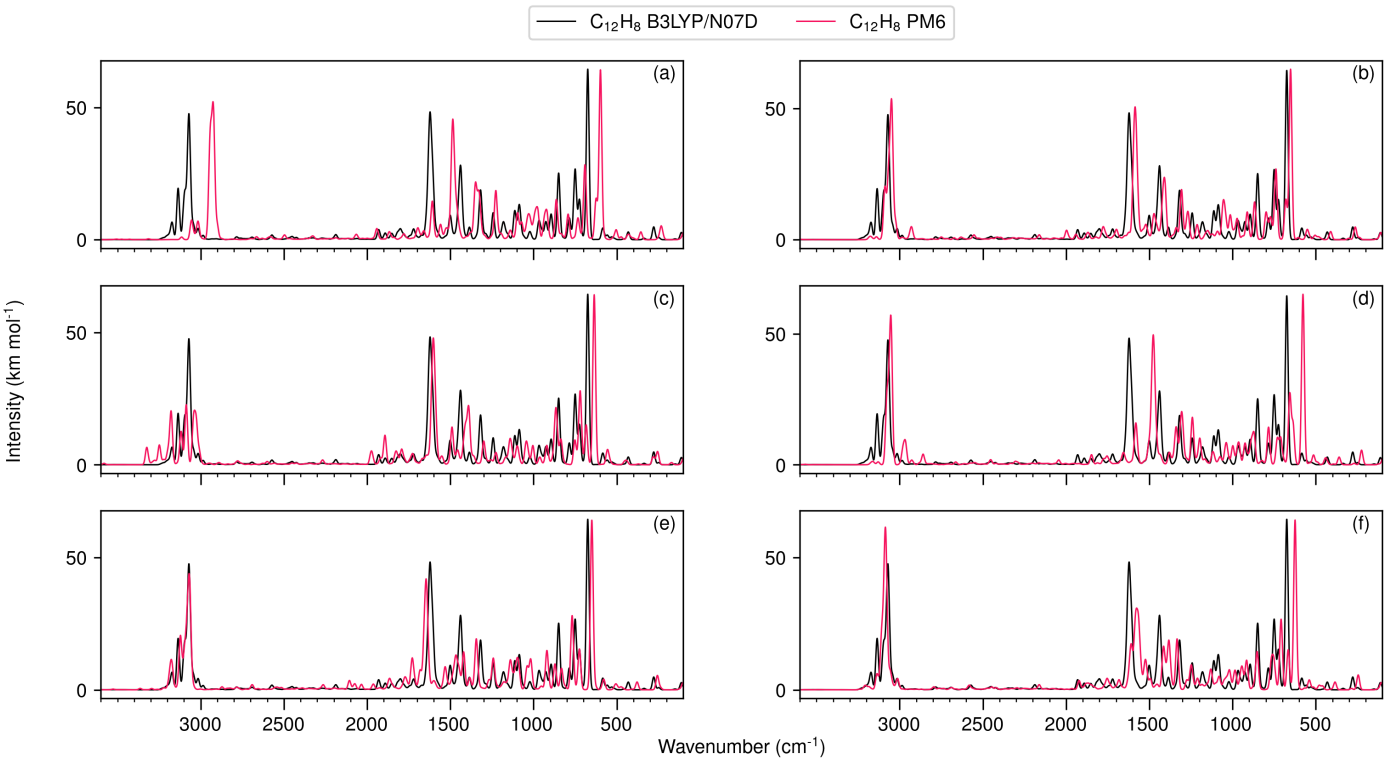

Fig. 6 Simulated spectra for benzopentalene from 3600-100 cm $^{-1}$. Each spectrum contains the B3LYP/N07D anharmonic reference spectrum. The index in the top right corner of each spectrum indicates the parameter set (see Table 3) used for the PM6 QFF calculation.

cies - (a) and (c) - result in spectra with displaced C-H stretching features. The overall poor fits by parameter sets (c) and (d) for both $(C_6H_4)_2$ and $C_{12}H_8$, despite the accuracy of the *c*-C_4H_4 training data and fitting errors similar to other reparameterizations, show that the representation of antiaromaticity alone in the parameters is not enough to predict the infrared features of larger species with some antiaromatic character. Several of the parameter sets predict the most intense peaks for $(C_6H_4)_2$ and $C_{12}H_8$, and such accurately demonstrates the potential utility of this method for generation of spectra of much larger partially-antiaromatic molecules for astronomical observation purposes.

3.2.3 Cascade Emission Spectra

The anharmonic IR cascade emission spectra of benzopentalene and biphenylene are depicted in Figure 7 (top and bottom, respectively), simulated at initial internal energies of 1 (black), 2 (dark blue), and 3 (light blue) eV. The spectra embody the same trends as the absorption spectra from Figures 5-6, and each band has the characteristic sharp rise at high energy and long tail at low energy, a product of anharmonicity. For benzopentalene (top), this involves a crowded spectrum with many bands throughout the 5 - 15 μm (2000 - 667 cm^{-1}) region. Intensity is spread across the entire region, with the bands at 13.38, 13.04, 11.79, 8.99, and 6.21 μm having a relative intensity of 0.43, 0.61, 0.41, 0.39, and 0.17, respectively, of the strongest band at 14.79 μm in the 1 eV spectrum. **As previously mentioned, these features correlate with the major features in the B3LYP/N07D absorption spectrum in Fig. 6. The prominent 14.79 and 13.88 μm peaks correspond to symmetric C-H out-of-plane bending modes while the 13.04 μm feature corresponds to an antisymmetric C-H out-of-plane bending mode. The 11.79 μm peak combines a C-H in-plane bending mode and a C-H antisymmetric out-of-plane bend, the 8.99 μm peak is another C-H in-plane bending mode, and the 6.21 μm feature is an aggregate of several C-C in-plane stretching modes.**

As is typical with a cascade emission spectrum, the relative intensity between the higher and lower frequency bands increases with increased internal energy. This phenomenon has two main sources: how the energy is distributed within the molecule and the 'cascade' effect. The internal energy is randomly distributed throughout the vibrational modes of the molecule via IVR, and it takes more energy to populate a higher frequency vibrational mode. Therefore, these modes will have a smaller probability of being excited at lower internal energy. As the internal energy increases, more energy is partitioned into the higher frequency modes, increasing their emission intensity. Separately, between each photon emission, the energy is redistributed within the molecule. The greater probability of a lower frequency mode being populated during each subsequent IVR event provides a greater probability of emission from a lower frequency mode due to the inaccessibility of the higher energy transitions. At higher internal energy, these shorter wavelength modes can be populated more readily, leading to the larger intensities. Case-in-point, the relative intensities of the bands at 13.04, 13.38, 11.79, and 8.99, and 6.21 μm in the 3 eV simulation are 0.53, 0.61, 0.51, 0.39, and 0.40, respectively. This represents a 10, 10, and 23% in-

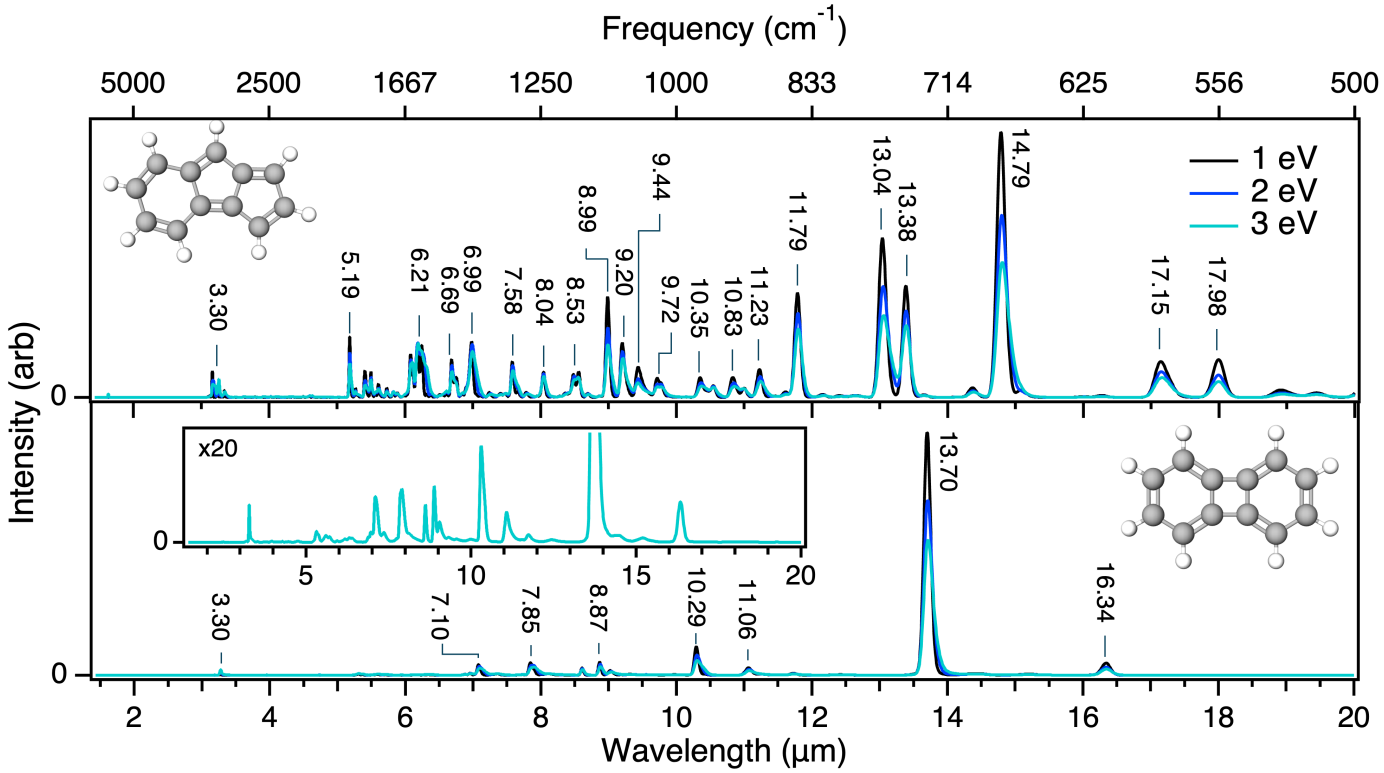

Fig. 7 Anharmonic cascade emission spectra of benzopentalene (top) and biphenylene (bottom) at starting internal energies of 1 (black), 2 (dark blue), and 3 (light blue) eV computed at the B3LYP/N07D level of theory.

crease in the intensity of the bands at 13.04, 11.79, and 6.21 μm. This physics is why the intensity ratio of the higher and lower frequency bands is used as a tracer of radiation environment. Similarly, these ratios can be used as tracers for PAH size due to the greater probability of population of higher frequency modes in small PAHs with less vibrational modes to store energy.

Moving to biphenylene in the bottom panel of Figure 7, one very intense band is computed at 13.70 μm at all three internal energies. This band originates from the symmetric out-of-plane CH bend, and does not shift discernibly from the absorption spectrum. The remaining features at shorter wavelength have intensities an order of magnitude lower. The inset shows an expanded view of the 3 eV emission spectrum in this region. Similarly to the benzopentalene spectrum, the region from 5 - 12 μm is dense with many overlapping bands. Interestingly, the intrinsic intensity of the transitions in this region are similar between biphenylene and benzopentalene making the ability to differentiate these isomers in an emission or absorption experiment difficult. Of course, the extremely strong band at 13.70 μm is the best opportunity. However, this would require observation towards a source lacking in C_2H_2, which has strong absorption[85] and emission[86] features at 13.70 μm. Due to the ubiquity of C_2H_2 in the ISM, infrared observation of $(C_6H_4)_2$ may be impractical.

There are significant differences in the CH stretch fundamental feature at 3.30 μm between the two molecules. In biphenylene, there is a single band that originates from a few overlapping CH stretch fundamentals and related combination bands as well as the symmetric nature of the molecule. This is seen in the emission spectrum as well as the experimental and computational absorption spectra in Figure 5. In contrast, the CH stretch region of benzopentalene has three identifiable bands across a 0.3 μm (950 cm^{-1}) range. This again agrees with the computed absorption spectra in Figure 6.

Conclusions

The dimerization of *o*-C_6H_4, a species observed within the **Taurus Molecular Cloud (TMC-1)**, is shown to efficiently associate into the partially-antiaromatic species biphenylene, $(C_6H_4)_2$, and its isomer benzopentalene, $C_{12}H_8$. Due to the lack of easily accessible bimolecular product sets on the dimerization pathway as well as the relatively large submerged barriers to isomerization, radiative association to form $(C_6H_4)_2$ is expected to be the dominant process upon self-interaction of *o*-C_6H_4. This pinpoints another polycyclic hydrocarbon species for potential interstellar observation and may serve to deplete *o*-C_6H_4 significantly enough to hinder its observation in other cold environments. Addition of a permanent dipole moment to biphenylene through functionalization with ·CN or ·C_2H would likely enhance its detectability, and such species may also be useful targets for astronomers.

QFFs for $(C_6H_4)_2$ computed with reparameterized versions of the semiempirical method PM6 are shown to generate spectra that, for certain parameter sets, accurately reproduce the major features of the experimental spectrum. While the B3LYP/N07D-computed spectrum is more accurate on the whole for $(C_6H_4)_2$ specifically, the reparameterized PM6 method used here can serve to generate spectra for much larger mixed aromatic-antiaromatic

molecules for which B3LYP/N07D anharmonic frequency calculations become infeasible, i.e., when the molecules of interest have dozens of atoms. While the congruence between the B3LYP/N07D-computed spectrum and the reparameterized-PM6 spectrum is lessened for $C_{12}H_8$, this is largely due to the lowered symmetry and thus the greater number and complexity of the features from 1500-750 cm^{-1} compared to the spectrum of $(C_6H_4)_2$.

The emission spectra reveal a similarly complex structure from 800 - 2000 cm^{-1} with the lower symmetry $C_{12}H_8$ having a more complex spectrum. The 3.30 μm feature embodies this, with a multi-peak structure for $C_{12}H_8$ compared to the single peak seen in $(C_6H_4)_2$. The major difference between the spectra of $(C_6H_4)_2$ and $C_{12}H_8$ is the large intensity of the 13.70 μm (730 cm^{-1}) feature for $(C_6H_4)_2$. The bands all show the characteristic tail on the low energy side due to the inclusion of anharmonicity.

The major features that could distinguish $(C_6H_4)_2$ in high-resolution IR spectra, such as the 730 cm^{-1} C-H out-of-plane bend, are still predicted accurately (relative to B3LYP/N07D) with certain parameter sets, further validating this method. However, the coincidence of this distinguishing feature in both the absorption and emission spectra with that of C_2H_2 will likely block infrared observation of $(C_6H_4)_2$. Ionization of this species could shift this peak away from the C_2H_2 feature to allow for infrared detection[87] or intensify features that are weak in the neutral spectrum.[73] Functionalization via $\cdot CN$[88] or $\cdot C_2H$ could also serve to shift the C-H out-of-plane bend, but more interestingly would likely allow observation through rotational spectroscopy. Continued comparison between spectra of mixed aromatic-antiaromatic species of interest generated using these reparameterized versions of PM6 with existing experimental data or high-accuracy computational data could be used to correlate the performance of each parameter set to specific molecular characteristics, such as the degree of antiaromaticity, and further define the utility of each PM6 reparameterization for specific subclasses of species.

Author contributions

Conceptualization, A.R.F.; Methodology, Investigation, Visualization, and Original Manuscript, A.R.F. and V.J.E.; Manuscript Revisions, A.R.F., V.J.E., and R.C.F.; Supervision and Funding, R.C.F.

Conflicts of interest

There are no conflicts to declare.

Data availability

Optimized geometries, harmonic frequencies, and T1 and D1 diagnostics for all reaction pathway structures, parameter sets from all reparameterizations, and rotational constants, harmonic, fundamental, and corrected fundamental frequencies, and Coriolis and Fermi resonances for all *o*-C_6H_4, *c*-C_4H_4, $(C_6H_4)_2$, and $C_{12}H_8$ QFFs are available in the Supplementary Information.

Acknowledgements

A.R.F. and R.C.F. acknowledge funding from NASA Grant 80NSSC24M0132 and NSF Grant AST-2103269. All authors are appreciative to the Mississippi Center for Supercomputer Research for the use of the computer resources necessary to perform this work. V.J.E. acknowledge support from the Internal Scientist Funding Model (ISFM) Laboratory Astrophysics Directed Work Package at NASA Ames. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center.

Notes and references

1 H. Henning Wenk, M. Winkler and W. Sander, *Angew. Chem. Int. Ed.*, 2003, **42**, 475–587.
2 C. Wentrup, *Austrailian Journal of Chemistry*, 2010, **63**, 979–986.
3 M. Schafer and R. S. Berry, *J. Am. Chem. Soc.*, 1965, **87**, 4497–4501.
4 G. Porter and J. I. Steinfeld, *J. Chem. Soc. A*, 1968, 877–878.
5 R. S. Tranter, S. J. Klippenstein, L. B. Harding, B. R. Giri, X. Yang and J. H. Kiefer, *J. Phys. Chem. A*, 2010, **114**, 8240–8261.
6 A. Comandini and K. Brezinsky, *J. Phys. Chem. A*, 2011, **115**, 5547–5559.
7 A. Comandini and K. Brezinsky, *J. Phys. Chem. A*, 2012, **116**, 1183–1190.
8 A. Matsugi and A. Miyoshi, *Phys. Chem. Chem. Phys.*, 2012, **14**, 9722–9728.
9 A. Comandini, S. Abid and N. Chaumeix, *J. Phys. Chem. A*, 20217, **121**, 5921–5931.
10 F. Hirsch, E. Reusch, P. Constantinidis, I. Fischer, S. Bakels, A. M. Rijs and P. Hemberger, *J. Phys. Chem. A*, 2018, **122**, 9563–9571.
11 M. N. McCabe, P. Hemberger, E. Reusch, A. Bodi and J. Bouwman, *J. Phys. Chem. Lett.*, 2020, **11**, 2859–2863.
12 L. Monluc, A. A. Nikolayev, I. A. Medvedkov, V. N. Azyazov, A. N. Morozov and A. M. Mebel, *ChemPhysChem*, 2022, **23**, e202100758.
13 M. N. McCabe, P. Hemberger, D. Campisi, J. C. Broxterman, E. Reusch, A. Bodi and J. Bouwman, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1869–1876.
14 J. Bouwman, M. N. McCabe, C. N. Shingledecker, J. Wandishin, V. Jarvis, E. Reusch, P. Hemberger and A. Bodi, *Nat. Astron.*, 2023, **7**, 423–430.
15 N. Hansen, T. Bierkandt, N. Gaiser, P. Oßwald, M. Köhler and P. Hemberger, *Proc. Combust. Inst.*, 2024, **40**, 105623.
16 D. E. Couch, M. M. San Marchi and N. Hansen, *Phys. Chem. Chem. Phys.*, 2024, **26**, 24833–24840.
17 J. Cernicharo, M. Agúndez, R. I. Kaiser, C. Cabezas, B. Tercero, N. Marcelino, J. R. Pardo and P. de Vicente, *Astron. Astrophys.*, 2021, **652**, L9.
18 F. Zhang, D. Parker, Y. S. Kim, R. I. Kaiser and A. M. Mebel, *Astrophys. J.*, 2011, **728**, 141.
19 T. A. Cole, S. R. Davis, A. R. Flint and R. C. Fortenberry, *Phys. Chem. Chem. Phys.*, 2024, **26**, 12510–12519.
20 D. B. Rap, A. Simon, K. Steenbakkers, J. G. M. Schrauwen, B. Redlich and S. Brünken, *Faraday Discuss.*, 2023, **245**, 221–244.

21 D. B. Rap, J. G. M. Schrauwen, B. Redlich and S. Brünken, *Phys. Chem. Chem. Phys.*, 2024, **26**, 7296–7307.

22 C. Puzzarini, *Front. Astron. Space Sci.*, 2022, **8**, 811342.

23 I. Novak, *J. Phys. Chem. A*, 2008, **112**, 2503–2506.

24 NIST Mass Spectrometry Data Center, William E. Wallace, director, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, National Institute of Standards and Technology, Gaithersburg, M.D., retrieved February 16, 2025.

25 P. Linstrom and W. Mallard, *J. Chem. Eng. Data*, 2001, **46**, 1059–1063.

26 M. B. Pastor, A. J. Kuhn, P. T. Nguyen, M. V. Santander, C. Castro and W. L. Karney, *J. Phys. Org. Chem.*, 2013, **26**, 750–754.

27 E. Hébrard, M. Dobrijevic, J. C. Loison, A. Bergeat, K. M. Hickson and F. Caralp, *Astron. Astrophys.*, 2013, **552**, A132.

28 C. J. Mackie, T. Chen, A. Candian, T. J. Lee and A. G. G. M. Tielens, *J. Chem. Phys.*, 2018, **149**, 134302.

29 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.

30 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

31 W. T. Yang, R. G. Parr and C. T. Lee, *Phys. Rev. A*, 1986, **34**, 4586–4590.

32 P. J. Stephens, F. J. Devlin, C. F. Chablowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

33 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

34 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

35 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmeri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang and M. Welborn, *MOLPRO, version 2024.2, a package of ab initio programs*, 2024.

36 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *WIREs Comput. Mol. Sci.*, 2012, **2**, 242–253.

37 H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Hesselmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, T. F. Miller, A. Mitruschenkov, K. A. Peterson, I. Polyak, G. Rauhut and M. Sibaev, *J. Chem. Phys.*, 2020, **152**, 144107.

38 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revision C.01*, 2016, Gaussian Inc. Wallingford CT.

39 T. B. Adler, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2007, **127**, 221106.

40 K. A. Peterson, T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2008, **128**, 084102–1.

41 G. Knizia, T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 054104.

42 W. Györffy and H.-J. Werner, *J. Chem. Phys.*, 2018, **148**, 114104.

43 H. Joachim Werner and P. J. Knowles, *J. Chem. Phys.*, 1988, **89**, 5803–5814.

44 P. J. Knowles and H. Joachim Werner, *Chem. Phys. Lett.*, 1988, **145**, 514–522.

45 T. Shiozaki, G. Knizia and H. Joachim Werner, *J. Chem. Phys.*, 2011, **134**, 034113.

46 T. Shiozaki and H. Joachim Werner, *J. Chem. Phys.*, 2011, **134**, 184104.

47 T. Shiozaki and H. Joachim Werner, *Mol. Phys.*, 2013, **111**, 607–630.

48 R. C. Fortenberry and T. J. Lee, *Annu. Rep. Comput. Chem.*, 2019, **15**, 173–202.

49 R. C. Fortenberry and T. J. Lee, *Vibrational Dynamics of Molecules*, World Scientific, 2022, pp. 235–295.

50 B. R. Westbrook and R. C. Fortenberry, *J. Chem. Theory Comput.*, 2023, **19**, 2606–2615.

51 J. F. Gaw, A. Willets, W. H. Green and N. C. Handy, 1996, *SPECTRO program*, version 3.0.

52 J. K. G. Watson, *Vibrational Spectra and Structure*, Elsevier, Amsterdam, 1977, pp. 1–89.

53 D. Papousek and M. R. Aliev, *Molecular Vibration-Rotation Spectra*, Elsevier, Amsterdam, 1982.

54 P. R. Franke, J. F. Stanton and G. E. Douberly, *J. Phys. Chem. A*, 2021, **125**, 1301–1324.

55 J. J. P. Stewart, *J. Mol. Model.*, 2007, **13**, 1173–1213.

56 J. J. P. Stewart, *Mopac2016*, Stewart Computational Chemistry, Colorado Springs, CO, USA, 2016.

57 B. R. Westbrook, J. P. Layfield, T. J. Lee and R. C. Fortenberry, *Electronic Structure*, 2022, **4**, 045003.

58 B. R. Westbrook and R. C. Fortenberry, *J. Mol. Spectrosc.*, 2023, **398**, 111846.

59 B. R. Westbrook and R. C. Fortenberry, *Annu. Rep. Comput. Chem.*, 2023, **19**, 65–85.

60 J. G. Radziszewski, B. A. H. Jr. and R. Zabradnik, *J. Am. Chem. Soc.*, 1992, **114**, 52–57.

61 G. Maier, *Angew. Chem. Int. Ed.*, 1988, **27**, 309–332.

62 A. M. Orendt, B. R. Arnold, J. G. Radziszewski, J. C. Facelli, K. D. Malsch, H. Strub, D. M. Grant and J. Michl, *J. Am. Chem.*

Soc., 1988, **110**, 2648–2650.

63 B. R. Arnold and J. Michl, *J. Phys. Chem.*, 1993, **97**, 13348–13354.

64 V. Barone, P. Cimino and E. Stendardo, *J. Chem. Theory Comput.*, 2008, **4**, 751–764.

65 V. Barone, M. Biczysko and J. Bloino, *Phys. Chem. Chem. Phys.*, 2014, **16**, 1759–1787.

66 E. Maltseva, A. Petrignani, A. Candian, C. J. Mackie, X. Huang, T. J. Lee, A. G. G. M. Tielens, J. Oomens and W. J. Buma, *Astrophys. J.*, 2015, **814**, 6.

67 I. M. Mills, *Molecular Spectroscopy: Modern Research*, Academic Press, New York, 1972, pp. 115–140.

68 R. C. Fortenberry, *J. Mol. Spectrosc.*, 2019, **357**, 4–8.

69 J. F. Gaw, A. Willets, W. H. Green and N. C. Handy, *Advances in Molecular Vibrations and Collision Dynamics*, JAI Press, Inc., Greenwich, Connecticut, 1991, pp. 170–185.

70 J. M. L. Martin, T. J. Lee, P. R. Taylor and J.-P. François, *J. Chem. Phys.*, 1995, **103**, 2589–2602.

71 J. M. L. Martin and P. R. Taylor, *Spectrochim. Acta, Part A*, 1997, **53**, 1039–1050.

72 C. J. Mackie, A. Candian, X. Huang, E. Maltseva, A. Petrignani, J. Oomens, W. J. Buma, T. J. Lee and A. G. G. M. Tielens, *J. Chem. Phys.*, 2015, **143**, 224314.

73 V. J. Esposito, L. J. Allamandola, C. Boersma, J. D. Bregman, R. C. Fortenberry, A. Maragkoudakis and P. Temi, *Mol. Phys.*, 2024, **122**, e2252936.

74 V. J. Esposito, P. Ferrari, W. J. Buma, R. C. Fortenberry, C. Boersma, A. Candian and A. G. G. M. Tielens, *J. Chem. Phys.*, 2024, **160**, 114312.

75 V. J. Esposito, P. Ferrari, W. J. Buma, C. Boersma, C. J. Mackie, A. Candian, R. C. Fortenberry and A. G. G. M. Tielens, *Mol. Phys.*, 2024, **122**, e2261570.

76 V. J. Esposito, R. C. Fortenberry, C. Boersma, A. Maragkoudakis and L. J. Allamandola, *Mon. Not. R. Astron. Soc. Lett.*, 2024, L87–L93.

77 P. M. Felker and A. H. Zewail, *Chem. Phys. Lett.*, 1984, **108**, 303–310.

78 M. Basire, P. Parneix, T. Pino, P. Bréchignac and F. Calvo, *EAS Publ. Ser.*, 2011, **46**, 95–101.

79 D. J. Cook and R. J. Saykally, *The Astrophysical Journal*, 1998, **493**, 793.

80 C. Pech, C. Joblin and P. Boissel, *Astron. Astrophys.*, 2002, **388**, 639–651.

81 C. J. Mackie, A. Candian, T. J. Lee and A. G. G. M. Tielens, *J. Phys. Chem. A*, 2022, **126**, 3198–3209.

82 V. J. Esposito, S. Bejaoui, B. E. Billinghurst, C. Boersma, R. C. Fortenberry and F. Salama, *Mon. Not. R. Astron. Soc.*, 2024, **535**, 3239–3251.

83 E. Herbst, *Front. Astron. Space Sci.*, 2021, **8**, 776942.

84 B. R. Arnold, J. G. Radziszewski, A. Campion, S. S. Perry and J. Michl, *Journal of the American Chemcial Society*, 1991, **113**, 692–694.

85 J. H. Lacy, N. J. Evans II, J. M. Achtermann, D. E. Bruce, J. F. Arens and J. S. Carr, *Astrophys. J. Lett.*, 1989, **342**, L43–L46.

86 J. S. Carr and J. R. Najita, *Science*, 2008, **319**, 1504–1506.

87 D. M. Hudgins and L. J. Allamandola, *Astrophys. J.*, 1999, **516**, L41–L44.

88 C. Cabezas, J. Janeiro, D. Pérez, W. Li, M. Agúndez, A. L. Steber, E. Guitián, J. Demaison, C. Pérez, J. Cernicharo and A. Lesarri, *J. Phys. Chem. Lett.*, 2024, **15**, 7411–7418.

Data availability

Optimized geometries, harmonic frequencies, and T1 and D1 diagnostics for all reaction pathway structures, parameter sets from all reparameterizations, and rotational constants, harmonic, fundamental, and corrected fundamental frequencies, and Coriolis and Fermi resonances for all o-C6H4, c-C4H4, (C6H4)2, and C12H8 QFFs are available in the Supplementary Information.