

Application of chemical graph theory to PAH isomer enumeration and structure in laser desorption/ionization mass spectrometry studies of particulate from an ethylene diffusion flame[☆]

J. Houston Miller*, Rachele J. Golden, Jennifer A. Giaccai,
Andrew Kamischke, Andrew Korte, Akos Vertes

Department of Chemistry, The George Washington University, 800 22nd Street, NW, Suite 4000, Washington, DC 20052, United States

Received 7 November 2019; accepted 28 June 2020

Available online 4 September 2020

Abstract

Our laboratory recently published data that showed that the PAH composition of soot can be exactly determined and spatially resolved by low-fluence laser desorption ionization, coupled with high-resolution mass spectrometry imaging [1]. This analysis revealed that PAHs of 239–838 Da, containing few oxygenated species, comprise the soot observed in an ethylene diffusion flame. In this paper, we demonstrate that the empirical formula of observed species can aid in the enumeration of isomers and places limits on their structures and thermodynamic stability. Specifically, chemical graph theory (CGT) shows that the vast majority of species observed in the sampled particulate matter may be described as benzenoid, consisting of only fused 6-membered rings. We apply CGT to determine the Dias Parameter, d_s , for observed, individual PAH peaks and demonstrate that observed PAH species cluster near low d_s , indicative of highly condensed structures, with relatively low populations of edge concavity (armchairs, bays, and fjords). Finally, we quantitatively explore the relative stability of PAH isomers using group-additivity estimates (for benzenoid structures) and those containing a single 5-membered rings using density functional theory. For the latter, we show that highly-condensed, benzenoid structures have lower free energy than those containing five-membered rings, with buried 5-membered rings showing the highest free energies.

© 2020 The Combustion Institute. Published by Elsevier Inc. All rights reserved.

Keywords: Soot; Polynuclear aromatic hydrocarbons; Graph theory

1. Introduction

A leading hypothesis for the structure of soot particles is that polynuclear aromatic hydrocarbon (PAH) molecules are their dominant components,

[☆] Colloquium: Soot, Nanomaterials, and Large Molecules.

* Corresponding author.

E-mail address: houston@gwu.edu (J.H. Miller).

with individual PAH molecules forming ordered stacks that agglomerate into primary particles. In a recent publication, we applied laser desorption and ionization mass spectrometry to the analysis of carbon particulate thermophoretically sampled from a nitrogen-diluted, ethylene/air diffusion flame previously [1]. The mass spectra and subsequent analysis revealed that the particulate is comprised predominately of medium (~ 300 – 800 Da), highly condensed structures and their average mass agreed well with our prior experimental determination of “conjugation length” from Raman scattering [2] and optical band gap measurements [3]. As shown in that contribution, an approach for developing structural insight about the PAH comprising the collected particulate is available from chemical graph theory (CGT). In this paper, we analyze these results using tools developed by the applied mathematics community and further explore isomeric variations for thermodynamic stability.

The new analysis builds on the work of Dias [4],[5] and Cyvin [6] who have made extensive contributions to the application of CGT to the enumeration of possible isomeric structures given an empirical formula. Using the mathematical tools developed in the references cited above, important limits emerge to confine the structure of molecules observed in our mass spectrometry experiments. As an example, we had considered one of the mass spectra collected in the wings of the flame. In this spectrum, there are approximately 180 identified m/z peaks (with intensities between 4200 and 86,000). The vast majority of these have empirical formulae consistent with benzenoid hydrocarbons. Specifically, for a hydrocarbon of formula C_nH_s , benzenoid species, those with only 6-membered rings, will satisfy

$$2 \left\lceil \frac{1}{2} (n + \sqrt{6n}) \right\rceil - n \leq s \leq n + 2 - 2 \left\lfloor \frac{1}{4} (n - 2) \right\rfloor. \quad (1)$$

where quantities within the curly brackets are rounded up to the nearest integer. (We consider the implications of the inclusion of five-membered ring structures in the discussion to follow).

Brunvoll and Cyvin [6] provide a number of additional limitations on and insights into structures based solely on empirical formula. For example, relationships are drawn for the number of internal carbon atoms, n_i ; the number of hexagonal rings, h ; and the numbers of isomers, both Kekuléan (closed shell) and non-Kekuléan (radicals). Finally, they provide an extensive discussion on what they refer to as the Dias Parameter, d_s , suggested by Dias to be a count of tree disconnections where “trees” are constructed from lines connecting internal carbon atoms. Brunvoll and Cyvin [6] show how the d_s can be calculated directly from the empirical formula

C_nH_s from

$$d_s = \frac{1}{2}(3s - n) - 7 \quad (2)$$

The Dias Parameter is 0 when there are no disconnections, positive when there are, and negative for branched trees which occur when there are extensive networks of interior carbons. As an example, consider several 6-ring molecules which were observed in our mass spectra: $C_{24}H_{12}$ (coronene), $C_{25}H_{13}$, $C_{26}H_{14}$, $C_{27}H_{15}$. All the isomers of the odd-carbon species are radicals (as well as a small number of the isomers of the even-carbon species). The number of possible isomers (1, 3, 10, and 5, respectively [6]) grows with d_s ($-1, 0, 1,$ and 2) while it is observed that the MS intensity is greater for the low d_s molecules (10,365, 26,997, 5226, and 8558, respectively) despite the statistically smaller number of structures that satisfy the empirical formula for low d_s . The same trends are seen for groups containing higher numbers of rings observed in our spectra, as shown in Fig. 2.

2. Description of experimental procedures and data

A report of the experimental procedure has appeared previously [1], [7],[8], and is summarized here. Additional experimental details are provided in the Supplemental Materials. The flame studied was a nitrogen-diluted, ethylene-air flames (60% ethylene by volume) supported on a burner developed at Yale University. It has been extensively studied and modelled by our laboratory and our collaborators at Yale [2],[9–17]. Additionally, it is a target, non-premixed flame of the International Sooting Flame workshop because of its established dilution levels that span a range of sooting levels and residency times. The burner consists of a central fuel tube with an inner diameter of 0.4 cm and a concentric air co-flow tube with an inner diameter of 5.1 cm. A stable flame is maintained using co-flowing air emitted from a honeycomb [10],[17].

All measurements were made on a Thermo LTQ MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, San Jose, CA). This instrument is equipped with an intermediate-pressure (75 mTorr) commercial MALDI source, which includes a nitrogen laser emitting 337 nm light with 4 ns pulse width and a repetition rate of 60 Hz. Silicon and aluminum sample substrates were affixed to a commercial sample plate using double-sided carbon tape (Ted Pella, Inc., Redding, CA) before insertion into the source. The carbon tape has not been shown to result in spectral features unless directly exposed to laser radiation. All spectra were acquired in the Orbitrap analyzer with nominal mass resolution set to 60,000.

Particulate from the flame was thermophoretically sampled onto low resistivity p-type silicon wafers (Silicon Valley Microelectronics, Inc., Santa

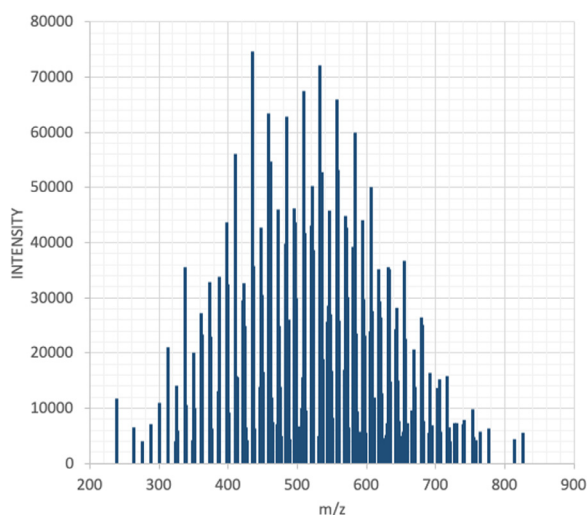


Fig. 1. Representative LDI-MS data from 30 mm above the burner in the flame wing.

Clara, CA). The substrates were inserted into the center of the axially symmetric flame, and the resulting deposition pattern provided a map of soot composition that may be compared with radial distribution of molecular and condensed-phase species measured by other techniques. After sampling, the wafers were analyzed by laser desorption ionization (LDI)-mass spectrometry (MS) imaging with a mass spectrometer capable of high mass resolution and accurate mass measurements.

When samples were introduced into this instrument, a camera takes a visible image of the surface so that the area of analysis can be selected by the user. In these imaging experiments, the laser power was adjusted for each measurement. One image pixel was an average of four scans, taken at four adjacent areas; each scan was acquired with 10 laser shots and separated by 100 μm in physical space on the sample surface, with the next pixel centered 200 μm over. Pixels were obtained in a raster pattern. All measurements reported were the average of 49 total scans.

Initial data processing is done using Xcalibur QualBrowser software (Thermo Scientific, San Jose, CA). Spectra were averaged, and the resulting averaged spectrum lists were exported to mMass version 5.5, an open source software tool [18]. A homebuilt library was constructed in Python (v. 2.7) consisting of all possible molecular combinations weighing less than 2000 Da of 150 carbons, 150 hydrogens, and 4 oxygens, without accounting for the stability of the resulting molecule. It is well known that the availability and number of various elements has a significant impact on peak identification; a specific m/z value could correspond very closely with a PAH of 56 carbons, but if the library is limited to 50 carbons, a similar m/z may be

achieved with 38 carbons and 13 oxygens, regardless of the practical likelihood of oxygenation.

There is an extensive literature available on the analysis of PAH, including their determination in environmental samples, including water and soil, or food products. (Many PAHs are tightly regulated by the Environmental Protection Agency due to their known carcinogenic effects [19].) Mass spectrometric techniques are used extensively to analyze more complex mixtures of larger PAHs from raw fuel materials, a field known as petroleomics. The main limitation of LDI is that there is an inherent requirement that the sample efficiently absorb the incident wavelength. PAH molecules are broadly absorbing and are ionized efficiently by UV radiation. As a consequence, molecular complexes (e.g. dimers, trimers, etc.) or fragmentation is often observed upon ionization. To explore the possibility for artifacts introduced in the experiments, a series of experiments were performed on pure coronene or a PAH standard mixture. (See Supplemental Materials). Finally, it is not uncommon to observe the formation of carbon clusters, or more specifically fullerenes, in mass spectra of carbon rich material upon laser absorption [20],[21]. In our studies, under the lowest fluence conditions, there were no clusters of carbon observed in the mass spectrum. Importantly, when the laser fluence was increased, there were many carbon cluster ions observed even when coronene was the only primary analyte.

A representative mass spectrum for a region 30 mm above the burner and in the soot “wing” area is shown in Fig. 1. Our PAH distribution is centered at higher mass values than previously reported MBMS or early LMMS results. In addition to the sampling method, the two most significant

differences between the current study and others are the low laser fluence (3.2 mJ/cm^2) combined with a more sensitive, and higher-resolution imaging mass spectrometer. The high mass resolving power of the Orbitrap analyzer allows for the resolution of nearly-isobaric species. This, combined with accurate m/z measurements improved by internal calibration to previously identified PAH ions, allows for the confident assignment of molecular formulae.

In the observed data, the flame areas of maximum mass spectrometric intensity was positioned between the region rich in hydrocarbon intermediates formed directly from fuel ethylene decomposition [22] and the high-temperature, radical rich region height [2].

3. Isomer enumeration and thermodynamic analysis

3.1. Generation of molecular structures from graph theory

The *Chemical and abstract Graph environment* (CaGe) program is an open-source software tool that applies graph theory to molecular structures [23]. CaGe's molecular generator enumerates *graphs* that serve as models for molecules. *Graphs* consist of *vertices* (or nodes) and *edges* (or connections) with adjacent vertices connected by an edge. The number of edges starting at a vertex (or the number of adjacent vertices) is the *valency* (or *degree*) of that vertex.

The *hydrocarbon* generator in CaGe package creates graphs consisting of planar (or near planar) structures with pentagonal and hexagonal faces. In the language of CGT, for 2-connected plane graphs with a distinguished outer face where all bounded faces are pentagons or hexagons, all vertices not in the boundary of the outer face have degree 3 and all vertices in the boundary of the outer face have degree 2 or 3. In the more familiar language of carbon-bonding, the working assumption is that all carbons in the structure are sp^2 hybridized and the degree can be used to distinguish edge from interior carbons. Further, by generating the entire suite of isomers characteristic of a given empirical formula, an inventory of edge-carbon morphology can be developed that will distinguish “zig zag”, “armchair”, “bay”, and “fjord” sites, as shown below.

3.2. Thermodynamic stability and equilibrium calculations

The enumeration of isomeric structures using CaGe software produces a mapping of atomic connectivity, but does not produce optimized chemical structures and thus does not provide quantitative insight into thermodynamic stability. After the structures of the possible isomers were

generated by CaGe software, adjacency lists were created using our own python program. The initial focus was on hexagonal, completely benzenoid structure. To determine the thermodynamic stability of different isomers, the group additivity approach developed by Yu was used [24], which can be accessed using the thermodynamic part of the Reaction Mechanism Generator (RMG) Python package or the RMG website produced by the Green group at MIT and West group at Northeastern University [25].

The thermodynamic libraries generated by RMG contained the NASA polynomials for each isomer CaGe constructed. For a family of PAH structures with a fixed number of aromatic rings, a gas mixture was constructed by adding these to the GRI 3.0 mechanism [26] and was solved for their equilibrium concentrations at a given temperature and pressure using the chemical equilibria solver of Cantera [27]. These calculations were done for 7-, 8-, 9-, and 10-ring PAH families. The starting elemental concentrations for carbon and hydrogen were set to match PAH with the greatest H/C ratio. These calculations were run every 100 K from 1500 to 2000 K and 1 atm. for each set.

When molecules containing 5-membered rings were evaluated, all bonds needed to be explicitly determined as the CaGe program would often produce non-Kekulean isomeric structures. For these isomers as well as ovalene, thermodynamic properties were evaluated by using density functional theory (DFT) in the NWChem computational chemistry program version 6.5 [28],[29]. Molecules were first geometry optimized using the 6–31G* basis set and the B3LYP exchange-correlation functional before frequency calculations were carried out using the 6–311G* basis set every 100 K between 300 and 2200 K. This basis set and functional have been shown to provide acceptable resolution for geometry optimization for PAH [30],[31]. Although this level of theory will not produce quantitatively accurate thermodynamics values, our goal here was only to compare relative magnitudes between structurally similar molecules. Specifically, we were interested in the equilibrium between single 5-membered-ring-containing isomers of $C_{32}H_{14}$ (abbreviated below as 1_5_XX, where XX is structural index generated by CaGe code) and ovalene.



$$\Delta G = (\varepsilon_0 + G_{corr})_{\text{product}} - (\varepsilon_0 + G_{corr})_{\text{reactant}} \quad (3)$$

Where ε_0 is the zero-point energy of the DFT calculation and G_{corr} is given by quantities generated in the NWCHEM vibrational analysis through

$$G_{corr} = H_{corr} - T \cdot S_{tot} \quad (4)$$

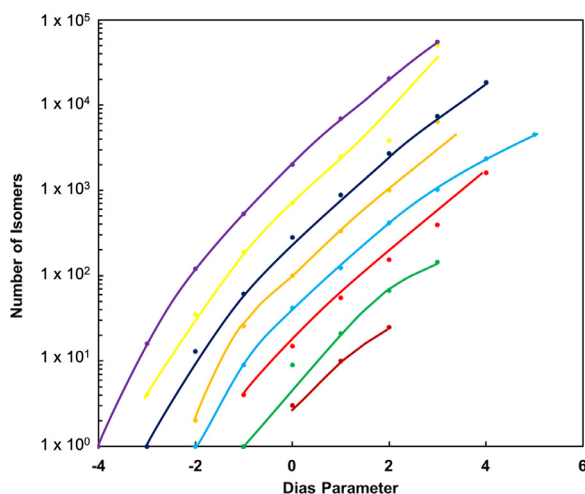


Fig. 2. *CaGe* predictions of the number of isomers as a function of ring count and d_s for observed LDI-MS peaks. Color Key: orange-7 rings, aqua-8 rings, red-9 rings, blue-10 rings, dark yellow-11 rings, dark blue-12 rings, yellow-13 rings, purple-14 rings. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

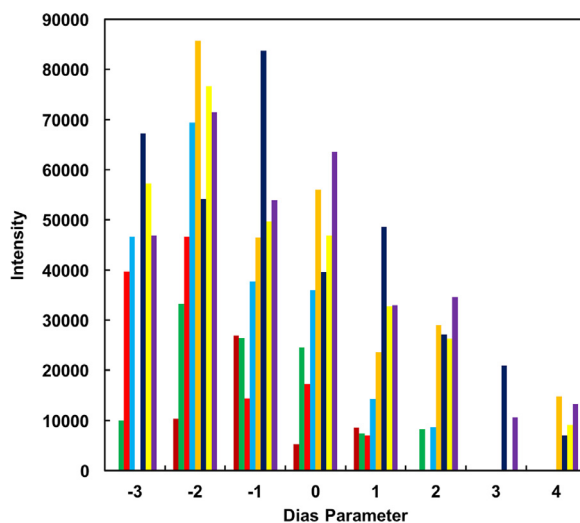


Fig. 3. Observed peak intensities for a range of ring counts sorted by d_s . See Fig. 2 for color scale.

4. Computational results and discussion

4.1. Isomer enumeration and observed structural trends

As noted in the introduction, the number of possible isomers increases rapidly with d_s and identifying molecular structural trends within this data is a daunting task, one that is only enabled through the use of *CaGe*. Shown in Fig. 2 is the exponential increase in isomer count for the observed 7–14 ring compounds. Some observations are immediately obvious in this data. First, the range in d_s ex-

pands with ring count, h . Second, for either a given d_s or a given ring count, the number of isomers grows exponentially. Third, for the lowest possible d_s in a given ring count family the number of isomers of even-numbered ring counts is one, for odd-numbered ring counts the number of isomers is in the single digits.

Fig. 3 provides insight into the observed distribution of these isomeric species as a function of d_s . As seen in Fig. 2 the number of possible isomers increases with d_s while it is observed that the LDI-MS intensity is greater for the low d_s molecules despite the statistically smaller number of structures

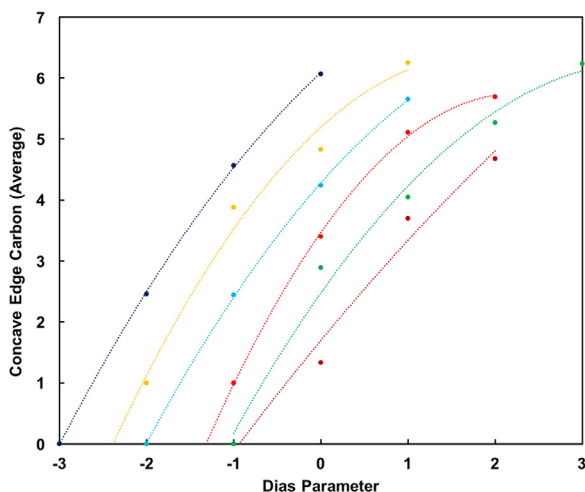


Fig. 4. Number of concave edge carbons for a range of ring counts sorted by d_s . See Fig. 2 for color scale.

that satisfy the empirical formula for low d_s . As discussed below, this is an important observation with implications for thermodynamic stability of the chemical components of incipient particles.

For likely PAH structures observed in flames, it is expected that as the d_s is increased, the number of interior carbons and the extent of aromatic conjugation must both decrease. To explore this hypothesis, *CaGe* was used to identify structures for (in some cases thousands of) isomers for a given ring count and d_s . This was accomplished by counting the number of non-zigzag, concave edge carbons found in the armchair, bay, and fjord regions in a given structure. These counts were then averaged to determine the number of concave edge carbons noting that an armchair represents two edge carbons, a bay three, and a fjord four. The result of this procedure is shown in Fig. 4 for each observed peak in the mass spectrum shown in Fig. 1. As anticipated, increasing d_s leads to an increase in the count of concave edge carbons. Further, this effect is more pronounced with increasing ring count.

4.2. Implications for thermodynamic stability

In a recent publication from our group, we showed how the Clar model, a method of studying the aromaticity of a structure, could be applied to an understanding of thermodynamic stability for different PAH isomers [32]. In Clar structures, the π electrons in a polycyclic ring structure are assigned so that the maximum number of rings has 6 π electrons, known as an aromatic sextet [33]. The Clar structure approach to understanding molecular structure has been validated by both atomic force microscopy and scanning tunneling microscopy [34]. The kinetic and thermodynamic stability of isomers increases with an increase in

number of aromatic sextets because of additional resonance energy [33].

The evaluation of resonant energy in large PAH has found increased focus as these molecules are thought to be ideal model systems for the understanding of bonding and electronic structure in nanographene compounds [35]. It has been found that the nature of PAH edges has a profound effect on the number of Clar sextets that may exist for a PAH isomer and thus electronic structure and thermodynamic stability. It is worth noting that counting sextets for irregularly shaped, large PAH can be challenging [36], and new measures of stability have been introduced, including bond resonance energy (BRE) and superaromatic stabilization energy (SSE). SSE is defined as the extra stabilization energy due to macrocyclic conjugation in these large π -systems, and thus is manifested in interior regions of large PAH. Specifically, Dias notes that for benzocoronenes, in the peripheral hexagons $BRE = SSE$, but for inner hexagons $BRE > SSE$ [37]. In this work, we note that additional resonance stabilization for interior carbons will be more relevant for structures with the most extensive network of interior conjugation, structures with minimal d_s .

Clearly, the presence of non-zigzag carbon features on the periphery of aromatic structures decreases the size of these superaromatic regimes and thus the thermodynamic stability of a particular isomer. Further, the presence of fjords in a molecule further decrease stability due to the steric interference of hydrogen atoms within the edge cavity leading to non-planar structures. The fact that the LDI-MS data clusters near low d_s suggests that the PAH components of soot are these most stable structures and are dominated by planar, hexagonal networks. This conjecture is explored quantitatively in the next section.

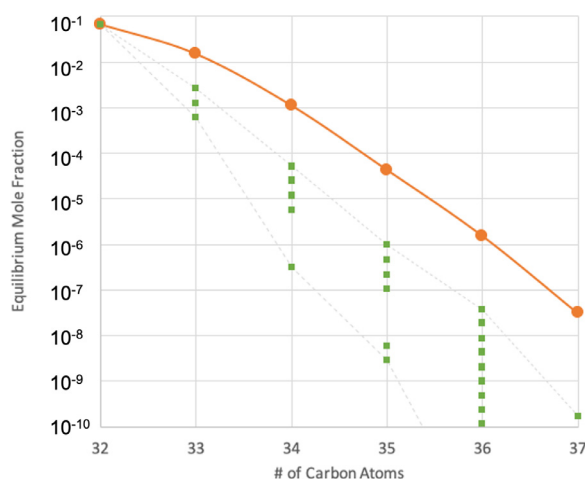


Fig. 5. Equilibrium mole fractions of 10-ring PAH at 1500 K and 1 atm. Total concentration are the orange circles and individual isomers are the green squares. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2.1. Thermodynamic stability of benzenoid PAH

In the mid 1980s Stein and Fahr [38] showed that certain isomers of PAH were the most thermodynamically stable isomeric form for given inventories of carbon and hydrogen atoms and these were labeled as stabilomers. It is important to note that although the original stabilomer study included only PAH with even numbers of carbon atoms, benzenoid structures can be comprised of an odd number of carbons, e.g., $C_{13}H_9$, $C_{47}H_{17}$ [39]. The initial Stein and Fahr work, and many follow-on studies used group additivity methods to estimate thermodynamics properties for these species. The Green group of MIT has greatly improved on the older group additivity estimates for PAH by adding an algorithm for bicyclic decomposition for large (>3 rings) PAH species [40].

Fig. 5 shows equilibrium mole fractions for 10-ring, benzenoid PAH at 1500 K and 1 atm as described in Section 3.2. Similar trends were observed for temperatures between 1500 and 2000 K. A total of 8467 unique isomers of these 10-ring PAH were included in this equilibrium mixture ranging from 32 to 39 carbon atoms. (For this family of PAH, $\#H = \#C - 18$. Further, via Eq. (2) we see that $d_s = \#C - 34$). As shown in Fig. 2, the number of isomers increases rapidly with d_s . However, RMG predicted a range of thermodynamic stability among molecules with the same d_s . As the complexity of PAH structures increased (towards larger d_s), degeneracies within these subsets also increased. For example, for $C_{33}H_{15}$, there are 9 unique isomeric structures that are predicted to have one of three values for free energy at 1500 K and thus three equilibrium mole fractions, shown as green squares in Fig. 5. The orange circle at each $\#C$ gives the sum of all isomeric concentrations for that empirical

formula. In agreement with the observed LDI-MS peak intensities (Fig. 3), for these 10-ring PAH the lower d_s species are predicted to be the most thermodynamically stable. The same trend is observed for PAH with higher and lower ring counts.

4.2.2. Comparative stability for $C_{32}H_{14}$ PAH containing one 5-membered ring

As reviewed in the earlier report of the measurements described here [1], recent literature suggests the importance of PAH structures which include five-membered rings [41–43]. Addition of even one pentagon into a PAH structure (fluoranthenoids) dramatically increases the number of possible isomers and, fortunately, these may also be enumerated using CaGe. Here, we concentrate on a subset of 10-ring PAH isomers of $C_{32}H_{14}$ including ovalene, shown above to be the most stable benzenoid structure, and six of the 75 isomers that contain one 5-membered rings of the same empirical formula. Two of the structures have buried pentagons leading to bowl shaped geometries, two have edge pentagons that include one C–H bond, and two have edge pentagons that contain two C–H bonds. (Structures for these species are included in the Supplementary Material). In our previous paper [1], we showed that inclusion of 5-membered rings increases sigma character and leads to greater reactivity of interior carbons.

For the examples studied at 1500 K we found that ovalene remained the most stable of the $C_{32}H_{14}$ isomers, with structures with edge 5-membered rings with two C–H groups next (ΔG of R1 = -30 ± 0 kcal/mol), followed by edge 5-membered rings with one C–H groups (-41 ± 0 kcal/mol), and finally structures with buried 5-membered rings (-65 ± 7 kcal/mol). These

results highlight the fact that the bowl-shaped molecules with pentagons occurring in the interior of the PAH molecule are the least preferred isomers, when compared to ovalene or the isomers with pentagons occurring on the edge of the molecule.

5. Conclusions

In this paper, we extend the analysis of recently published results of low-fluence, high-resolution, imaging LDI-MS in which we suggested that benzenoid PAH molecules are the building blocks of soot particles. From observed empirical formulae, we show that chemical graph theory can be used to determine important structural parameters and lead to both isomer enumeration and to inventory edge carbon topology. Further, through the application of both group additivity techniques and thermodynamic analysis vis DFT computational chemistry, we show that the PAH observed in soot tend towards highly condensed and thermodynamically stable structures.

These new results illuminate an ongoing debate in the chemical growth community: are there sufficiently high concentrations of large monomer PAH in flames to explain observed primary particle concentrations? Although we raised this concern in an early calculation of the equilibrium constant for the dimerization constant for PAH [44], in subsequent publications, we qualified this conclusion by noting that sufficiently large PAH will physically condense at temperatures characteristic of soot formation in diffusion flames [45], and this conclusion has been supported by recent molecular dynamics simulations of both physical and chemical condensation [46],[47]. The new LDI-MS results and the analysis reported here provide evidence that species of sufficient size are present within carbon particulates in flames and that their size is consistent with the conjugation length inferred by both Raman scattering and optical band gap measurements observed in the same burner/flame system [2],[3]. What remains to be determined is how these PAH came to be there, either through more complex chemistry in the gas phase [48–51],[52],[53], or graphitization within the growing particles.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This material is based upon work supported by the U.S. National Science Foundation under grant number CBET-1706757 with Drs. Song-Charn

Kong and Harsha Chelliah serving as technical monitors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.proci.2020.06.080.

References

- [1] R.S. Jacobson, A.R. Korte, A. Vertes, J.H. Miller, *Angew. Chem. Int. Ed.* 59 (2020) 4484–4490.
- [2] J.D. Herdman, B.C. Connelly, M.D. Smooke, M.B. Long, J.H. Miller, *Carbon N.Y.* 49 (2011) 5298–5311.
- [3] E.M. Adkins, J.H. Miller, *PCCP* 17 (2015) 2686–2695.
- [4] J.R. Dias, *J. Mol. Struct. THEOCHEM* 137 (1986) 9–29.
- [5] J. Dias, *Cheminform* (2003) 34.
- [6] J. Brunvoll, S. Cyvin, *Z. Naturforschung A* 45 (1990).
- [7] R.S. Jacobson, A.R. Korte, A. Vertes, J.H. Miller, *Angew. Chem. Int. Ed.* (2019) submitted for publication.
- [8] R.S.J. Golden, *The Development of Laser-Based Analytical Tools for Molecular Composition Characterization Across Combustion and Biological Systems, Chemistry*, George Washington University, Washington, DC, 2019.
- [9] M.D. Smooke, R.J. Hall, M.B. Colket, J. Fielding, M.B. Long, C.S. McEnally, L.D. Pfeifferle, *Combust. Theory Model.* 8 (2004) 593–606.
- [10] M.D. Smooke, M.B. Long, B.C. Connelly, M.B. Colket, R.J. Hall, *Combust. Flame* 143 (2005) 613–628.
- [11] S.A.K. Blair, C. Connelly, M.D. Smooke, M.B. Long, *Proc. US Sect. Combust. Inst.* 4 (2005).
- [12] S.B. Dworkin, M.D. Smooke, V. Giovangigli, *Proc. Combust. Inst.* 32 (2009) 1165–1172.
- [13] M.A. Puccio, J.D. Herdman, J.H. Miller, B.C. Connelly, M.D. Smooke, M.B. Long, *Chem. Phys. Process. Combust.* (2005) 217–220.
- [14] B.C. Connelly, B.A. Bennett, S. Dworkin, M.D. Smooke, M.B. Long, M.A. Puccio, J.D. Herdman, J.H. Miller, *Chem. Phys. Process. Combust.* (2005) 259–262.
- [15] S.B. Dworkin, B.C. Connelly, A.M. Schaffer, B.A.V. Bennett, M.B. Long, M.D. Smooke, M.P. Puccio, B. McAndrews, J.H. Miller, *Proc. Combust. Inst.* 31 (2007) 971–978.
- [16] J.H. Miller, J.D. Herdman, C.D.O. Green, E.M. Webster, *Proc. Combust. Inst.* 34 (2013) 3669–3675.
- [17] E.M. Adkins, J.H. Miller, *Phys. Chem. Chem. Phys.* 17 (2015) 2686–2695.
- [18] M. Strohal, M. Hassman, B. Kořata, M. Kodíček, *Rapid Commun. Mass Spectrom.* 22 (2008) 905–908.
- [19] S.-S. Cai, J.A. Syage, K.A. Hanold, M.P. Balogh, *Anal. Chem.* 81 (2009) 2123–2128.
- [20] H.W. Kroto, A.W. Allaf, S.P. Balm, *Chem. Rev.* 91 (1991) 1213–1235.
- [21] H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, R.E. Smalley, *Nature* 318 (1985) 162.

- [22] K.C. Smyth, J.H. Miller, R.C. Dorfman, W.G. Mallard, R.J. Santoro, *Combust. Flame* 62 (1985) 157–181.
- [23] G. Brinkmann, O. Friedrichs, S. Liskén, A. Peeters, N. Van Cleemput, *MATCH Commun. Math. Comput. Chem.* 63 (2010) 533–552.
- [24] J. Yu, R. Sumathi, W.H. Green, *J. Phys. Chem. A* 110 (2006) 6971–6977.
- [25] C.W. Gao, J.W. Allen, W.H. Green, R.H. West, *Comput. Phys. Commun.* 203 (2016) 212–225.
- [26] G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriarty, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, J. William C. Gardiner, V.V. Lissianski, Z. Qin, GRI-Mech 3.0. <http://combustion.berkeley.edu/gri-mech/releases.html> (2020).
- [27] D. Goodwin, H. Moffat, R. Speth, Cantera: An Object-oriented Software Toolkit for Chemical Kinetics, Thermodynamics, and Transport Processes. Version 2.4.0. <https://www.cantera.org>, doi:10.5281/zenodo.1174508.
- [28] M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T.L. Windus, W.A. de Jong, *NWChem, Comput. Phys. Commun.* 181 (2010) 1477–1489.
- [29] E. Aprà, E.J. Bylaska, W.A. de Jong, N. Govind, K. Kowalski, T.P. Straatsma, M. Valiev, H.J.J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F.W. Aquino, R. Atta-Fynn, J. Autschbach, N.P. Bauman, J.C. Becca, D.E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G.N. Chuev, C.J. Cramer, J. Daily, M.J.O. Deegan, T.H. Dunning, M. Dupuis, K.G. Dyall, G.I. Fann, S.A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A.W. Götz, J. Hammond, V. Helms, E.D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B.G. Johnson, H. Jónsson, R.A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R.D. Lins, R.J. Littlefield, A.J. Logsdail, K. Lopata, W. Ma, A.V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J.E. Moore, J.M. Mullin, T. Nakajima, D.R. Nascimento, J.A. Nichols, P.J. Nichols, J. Nieplocha, A. Otero-de-la-Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati, J. Pittner, L. Pollack, R.M. Richard, P. Sadayappan, G.C. Schatz, W.A. Shelton, D.W. Silverstein, D.M.A. Smith, T.A. Soares, D. Song, M. Swart, H.L. Taylor, G.S. Thomas, V. Tipparaju, D.G. Truhlar, K. Tsemekhman, T. Van Voorhis, Á. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K.D. Vogiatzis, D. Wang, J.H. Weare, M.J. Williamson, T.L. Windus, K. Woliński, A.T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, R.J. Harrison, *J. Chem. Phys.* 152 (2020) 184102.
- [30] E.M. Adkins, J.A. Giaccai, J.H. Miller, *Proc. Combust. Inst.* 36 (2017) 957–964.
- [31] T.C. Allison, D.R. Burgess, *J. Phys. Chem. A* 119 (2015) 11329–11365.
- [32] E.M. Adkins, J.H. Miller, *Phys. Chem. Chem. Phys.* 19 (2017) 28458–28469.
- [33] E. Clar, R. Schoental, *Polycyclic Hydrocarbons*, Springer, 1964.
- [34] M. Solà, *Front. Chem.* 1 (2013) 1–22.
- [35] K. Sakamoto, N. Nishina, T. Enoki, J.-I. Aihara, *J. Phys. Chem. A* 118 (2014) 3014–3025.
- [36] N. Nishina, M. Makino, J.-i. Aihara, *J. Phys. Chem. A* 120 (2016) 2431–2442.
- [37] J.R. Dias, *J. Phys. Chem. A* 123 (2019) 3229–3238.
- [38] S.E. Stein, A. Fahr, *J. Phys. Chem.* 89 (1985) 3714–3725.
- [39] K.H. Homann, *Angew. Chem. Int. Ed.* 37 (1998) 2434–2451.
- [40] K. Han, A. Jamal, C.A. Grambow, Z.J. Buras, W.H. Green, *Int. J. Chem. Kinet.* 50 (2018) 294–303.
- [41] N. Hansen, M. Schenk, K. Moshhammer, K. Kohse-Höinghaus, *Combust. Flame* 180 (2017) 250–261.
- [42] P.J.F. Harris, Z. Liu, K. Suenaga, *J. Phys. Condens. Matter* 20 (2008) 362201.
- [43] J.W. Martin, R.I. Slavchov, E.K.Y. Yapp, J. Akroyd, S. Mosbach, M. Kraft, *J. Phys. Chem. C* 121 (2017) 27154–27163.
- [44] J.H. Miller, K.C. Smyth, W.G. Mallard, *Symp. Int. Combust.* 20 (1985) 1139–1147.
- [45] J.H. Miller, *Symp. Int. Combust.* 23 (1991) 91–98.
- [46] P. Elvati, A. Violi, *Fuel* 222 (2018) 307–311.
- [47] J.S. Lowe, J.Y.W. Lai, P. Elvati, A. Violi, *Proc. Combust. Inst.* 35 (2015) 1827–1832.
- [48] J.H. Miller, *Proc. Combust. Inst.* 30 (2005) 1381–1388.
- [49] M. Frenklach, A.M. Mebel, *Phys. Chem. Chem. Phys.* 22 (2020) 5314–5331.
- [50] R.A. Krueger, G. Blanquart, *Phys. Chem. Chem. Phys.* 21 (2019) 10325–10335.
- [51] G. Vitiello, G. De Falco, F. Picca, M. Commodo, G. D’Errico, P. Minutolo, A. D’Anna, *Combust. Flame* 205 (2019) 286–294.
- [52] H. Wang, *Proc. Combust. Inst.* 33 (2011) 41–67.
- [53] J.W. Martin, D. Hou, A. Menon, L. Pascazio, J. Akroyd, X. You, M. Kraft, *J. Phys. Chem. C* 123 (2019) 26673–26682.

Supplemental Material

Title: Application of Chemical Graph Theory to PAH Isomer Enumeration and Structure in Laser Desorption/Ionization Mass Spectrometry Studies of Particulate from an Ethylene Diffusion Flame

Authors: J. Houston Miller*, Rachele J. Golden, Jennifer A. Giaccai, Andrew Kamischke, Andrew Korte, Akos Vertes

Department of Chemistry
The George Washington University
Washington, DC 20052

*Corresponding Author: J. Houston Miller
800 22nd Street, NW
Suite 4000
George Washington University
Washington, DC 20052, USA
+1 202-994-7474
houston@gwu.edu

Table of Contents

<i>S1. Additional Details on Experimental Procedures</i>	1
S1.1 Background on Mass Spectrometric Analysis of PAH	2
S1.2 Studies of PAH standards.....	2
S1.3 Spectra of Isolated and Mixed PAH Standards on Aluminum Foil Surfaces	3
S1.4 Source-Generated Ions on Aluminum Foil Substrates	8
S1.5 Spectra of Isolated and Mixed PAH Standards on Silicon Wafer Surfaces	12
S1.6 Source-Generated Ions on Silicon Wafer Substrates	16
S1.7 Fullerene Formation	19
<i>S.2 Structures of 5-membered ring-containing PAH considered in DFT analysis of thermochemistry</i>	22
<i>Supplemental References</i>	22

S1. Additional Details on Experimental Procedures

The burner consists of a central fuel tube with an inner diameter of 0.4 cm and a concentric air co-flow tube with an inner diameter of 5.1 cm. The fuel flow at the nozzle exit has a parabolic profile with an average velocity of 35 cm/s with the air velocity profiles at the burner surface being plug flow of 35 cm/s. A stable flame is maintained using co-flowing air emitted from a honeycomb [1, 2]. The flow of ethylene and nitrogen were controlled through a flow meter, Teledyne Hastings HFM-200 CH₄O-600scm and Teledyne Hastings HFM-200 N₂O-1000scm, while the air was controlled through a flow controller Omega FMA5542 N₂O-100L respectively.

All measurements were made on a Thermo LTQ MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, San Jose, CA). This instrument is equipped with an intermediate-pressure (75 mTorr) commercial MALDI source, which includes a nitrogen laser emitting 337 nm light with 4 ns pulse width and a repetition rate of 60 Hz. The laser beam interacts with the sample at a 32° angle, in the reflection geometry configuration. The laser power is precisely controlled via software and is continuously monitored, with one 50% neutral density filter in the beam path. Silicon and aluminum sample substrates were affixed to a commercial sample plate using double-sided carbon tape (Ted Pella, Inc., Redding, CA) before insertion into the source. The carbon tape has not been shown to result in spectral features unless directly exposed to laser radiation. All spectra were acquired in the Orbitrap analyzer with nominal mass resolution set to 60,000, though the resolution is m/z dependent. The mass range is user selected on this instrument, but it loses sensitivity when an overly wide range is selected, most experiments were conducted between 75-1000 m/z unless otherwise specified.

Particulate from the flame was thermophoretically sampled onto silicon wafer substrates. The substrates were inserted into the center of the axially symmetric flame, and the resulting deposition pattern provided a map of soot composition that may be compared with radial distribution of molecular and condensed-phase species measured by other techniques. After sampling, the wafers were analyzed by laser desorption ionization (LDI)-mass spectrometry (MS) imaging with a mass spectrometer capable of high mass resolution and accurate mass measurements.

When samples were introduced into this instrument, a camera takes a visible image of the surface so that the area of analysis can be selected by the user. While the design of the MALDI source is for matrix assisted ionization, in the absence of a matrix applied to the sample it functions as an LDI source. In these imaging experiments, the laser power was adjusted for each measurement. One image pixel was an average of four scans, taken at four adjacent areas; each scan was acquired with 10 laser shots and separated by 100 μm in physical space on the sample surface, with the next pixel centered 200 μm over. Pixels were obtained in a raster pattern. Due to homogeneity in standard solutions, the imaging aspect of these measurements was to ensure that only areas with sample were interrogated and averaged and standardize the number of scans for analysis. All measurements reported were the average of 49 total scans.

Initial data processing is done using Xcalibur QualBrowser software (Thermo Scientific, San Jose, CA). Spectra were averaged, and the resulting averaged spectrum lists were exported to mMass version 5.5. mMass is an open source software tool developed primarily by Martin Strohmalm.[3] A homebuilt library was constructed in Python (v. 2.7) consisting of all possible molecular combinations weighing less than 2000 Da of 150 carbons, 150 hydrogens, and 4 oxygens, without accounting for the stability of the resulting molecule. It is well known that the availability and number of various elements has a significant impact on the identification; a specific m/z value could correspond very closely with a PAH of 56 carbons, but if the library is limited to 50 carbons, a similar m/z may be achieved with 38 carbons and 13 oxygens, regardless of the practical likelihood of oxygenation. The library was imported into mMass to use for spectral feature matching.

S1.1 Background on Mass Spectrometric Analysis of PAH

There is an extensive literature available on the analysis of PAH, including their determination in environmental samples, including water and soil, or food products. (Many PAHs are tightly regulated by the Environmental Protection Agency due to their known carcinogenic effects [4].) Mass spectrometric techniques are used extensively to analyze more complex mixtures of larger PAHs from raw fuel materials, a field known as petroleomics. One key, defining feature of most petroleomics research is the use of extremely high resolution mass spectrometers, typically using Fourier transform - ion cyclotron resonance (FT-ICR) detectors. Acquisition of an FT-ICR instrument is incredibly expensive, due largely to the need for a superconducting magnet. There are available facilities in Florida and Washington where users can bring samples to analyze using FT-ICR instruments with >20 T magnetic fields.

The main limitation of LDI is that there is an inherent requirement that the sample efficiently absorb the incident wavelength. PAH molecules are broadly absorbing and are ionized efficiently by UV radiation. As a consequence, molecular complexes (e.g. dimers, trimers, etc) or fragmentation is often observed upon ionization. The absorbance of UV light by PAH molecules is so substantial that they have been used in various forms as MALDI matrices, to enhance sensitivity of biological analytes.[5-9]

S1.2 Studies of PAH standards

Standard PAH compounds available for purchase were limited. In the current work, 98.4% pure solid coronene and an EPA certified reference mixture of standard PAH molecules were dissolved in acetone (EPA 525 PAH mix B, certified purity $\geq 98.8\%$ pure for each component) were purchased from Sigma-Aldrich (St. Louis, MO). All samples were used without further purification. The standard mix contained each of the compounds listed in the supplementary materials). Aliquots of the standard mix solution were dried in a vacuum centrifuge for storage. Solutions prepared for analysis were all dissolved in HPLC grade toluene from Sigma-Aldrich (St. Louis, MO).

Samples were deposited on either low resistivity p-type silicon wafers (Silicon Valley Microelectronics, Inc., Santa Clara, CA) or Ultra Clean Supremium aluminum foil from

VWR, after gently cleaning with methanol (foil only). Typical experiments were performed on 500 nL droplet depositions of sample onto the substrate. The entire sample was not interrogated in any one measurement, so these values should be used only as a guide to relative amounts. In the PAH standard mix, though 0.9-1.6 nmol of material deposited for each molecule, several indistinguishable isomers were present. Typical coronene stock solution concentrations were 2 mM, or 500 nmol in a 500 nL sample. All samples were air dried prior to MS analysis. Each combination of standard solution and substrate was analyzed in triplicate and the results were averaged with standard deviations calculated.

The simplest case studied was a standard coronene solution deposited on cleaned aluminum foil, a substrate used in other combustion sampling measurements.[10] The laser fluence is a very significant aspect of these measurements and was carefully controlled. For example, in the analysis of coronene there is a marked difference in the appearance of the spectrum at the lowest (3.18 mJ/cm²), and higher fluence (11.94 mJ·cm⁻²). Specifically, coronene (C₂₄H₁₂, *m/z* 300.0934) is not the only detectable ion in the spectrum at higher fluence; notably the loss of 2C or 2H is observed. (See Supplement for complete details.)

In addition to coronene, a mixture of several PAH components was measured in the same fashion. As with coronene, there are many peaks present in all spectra that do not correspond to ions specified in the standard material. The concentrations of all species are lower than that of coronene (~150×), therefore it follows that the signal overall would be lower. Likely because of this concentration difference, there are fewer ions unaccounted for in the PAH mix, though this may also be related to purity of the original solutions. Many trends observed in the data of the PAH standard mix are similar to what was observed in coronene; chiefly, the loss of 2C or 2H is observed. (See Supplement.)

Carbon clusters, or more specifically fullerenes, are observed in mass spectra of carbon rich material upon laser excitation at high fluence. The 1996 Nobel Prize in chemistry was awarded to Professors Curl, Kroto, and Smalley for discovering C₆₀ with a similar experiment as performed here. In those experiments, laser light was focused onto graphitic surfaces, and extensive follow-up research was performed [11, 12]. In our studies, under the lowest fluence conditions, there were no clusters of carbon observed in the mass spectrum. Importantly, when the laser fluence was increased sufficiently, there were many carbon cluster ions observed even when coronene was the only primary analyte.

S1.3 Spectra of Isolated and Mixed PAH Standards on Aluminum Foil Surfaces

The simplest case studied was a standard coronene solution deposited on cleaned aluminum foil, a substrate used in other combustion sampling measurements.[10] The laser fluence is a very significant aspect of these measurements and were carefully controlled. **Figure S.1** shows a typical spectrum of coronene analysis with the lowest laser fluence, 3.18 mJ/cm², and **Figure S.2** shows a similar spectrum acquired with 11.94 mJ·cm⁻² of laser fluence. Coronene, or C₂₄H₁₂, (*m/z* 300.0934) is not the only detectable ion in the spectrum, and is the dominant ion observed.

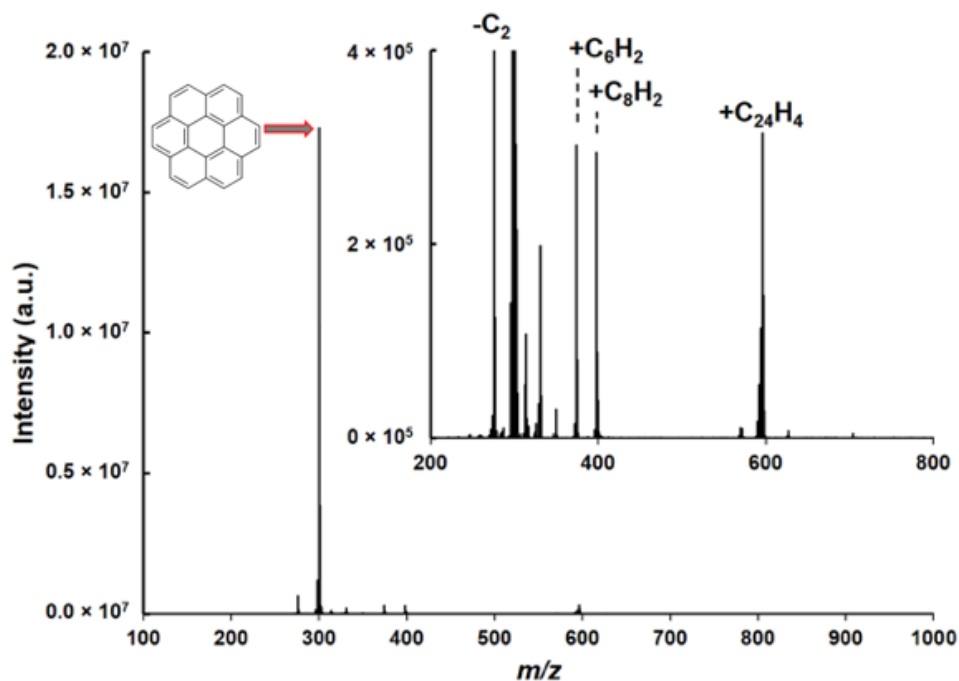


Figure S.1. Coronene spectrum after sample deposition on aluminum foil surface. The fluence used in this acquisition was 3.18 mJ/cm². The arrow indicates the coronene base peak.

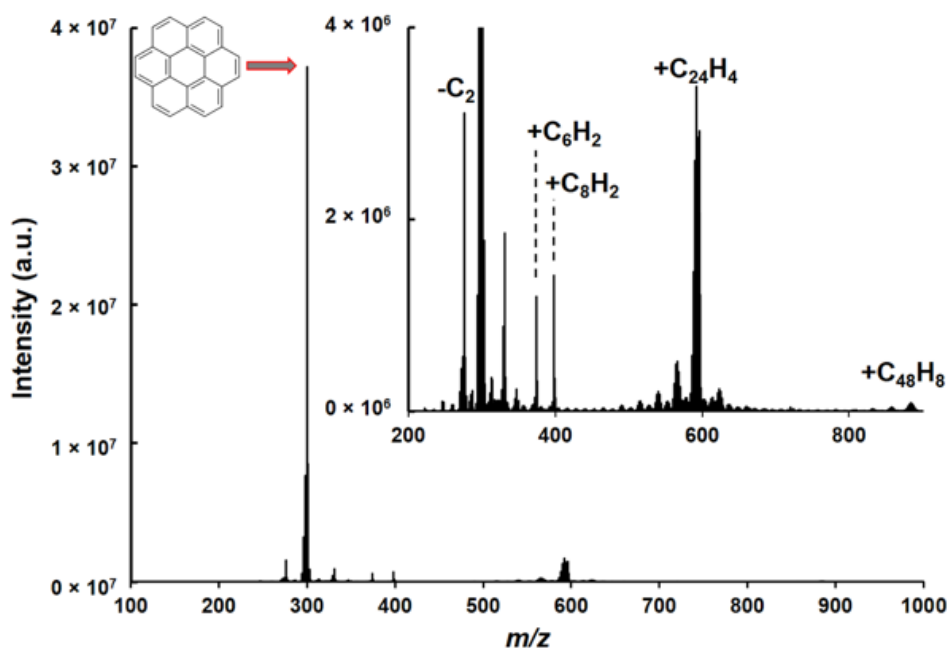


Figure S.2. Coronene spectrum after sample deposition on aluminum foil surface. The fluence used in this acquisition was 11.94 mJ/cm². The arrow indicates the coronene base peak.

In addition to coronene, a mix of several PAH components was measured in the same fashion. For a list of the components in the PAH mix, see **Table S.1**. A typical spectrum from the standard mix at $3.18 \text{ mJ}\cdot\text{cm}^{-2}$ and $11. \text{ mJ}\cdot\text{cm}^{-2}$ are presented in **Figure S.3** and **5.5**, respectively. In all spectra, the smallest PAH from the mix that was observed experimentally was $\text{C}_{20}\text{H}_{12}$. As with coronene, there are many peaks present in all spectra that do not correspond to ions specified in the standard material. The concentrations of all species are lower than that of coronene ($\sim 150\times$), therefore it follows that the signal overall would be lower. Likely because of this concentration difference, there are fewer ions unaccounted for in the PAH mix, though this may also be related to purity of the original solutions. An important aspect of **Figures S.3** and **S.4** is that there are two groups of isomers in the standard mix that are similar in m/z values: $\text{C}_{22}\text{H}_{14}$ and $\text{C}_{22}\text{H}_{12}$. Though the amount of $\text{C}_{22}\text{H}_{12}$ present in solution was $2\times$ greater than $\text{C}_{22}\text{H}_{14}$, $\text{C}_{22}\text{H}_{14}$ was the base peak in all spectra from aluminum foil, suggesting a greater ionization efficiency of $\text{C}_{22}\text{H}_{14}$ than $\text{C}_{22}\text{H}_{12}$.



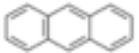





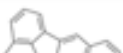


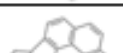

Name	Formula	Structure	Monoisotopic Mass (Da)	Amount Deposited (nmol)
Acenaphthylene	C ₁₂ H ₈		152.0626	1.6
Fluorene	C ₁₃ H ₁₀		166.0783	1.5
Anthracene	C ₁₄ H ₁₀		178.0783	2.8
Phenanthrene	C ₁₄ H ₁₀		178.0783	
Pyrene	C ₁₆ H ₁₀		202.0783	1.2
Benz[a]anthracene	C ₁₈ H ₁₂		228.0939	2.2
Chrysene	C ₁₈ H ₁₂		228.0939	
Benzo[b]fluoranthene	C ₂₀ H ₁₂		252.0939	3.0
Benzo[k]fluoranthene	C ₂₀ H ₁₂		252.0939	
Benzo[a]pyrene	C ₂₀ H ₁₂		252.0939	
Benzo[ghi]perylene	C ₂₂ H ₁₂		276.0939	1.8
Indeno[1.2.3-cd]pyrene	C ₂₂ H ₁₂		276.0939	
Dibenz[a,h]anthracene	C ₂₂ H ₁₄		278.1096	0.9

Table S.1. Molecules included in the PAH standard mix.

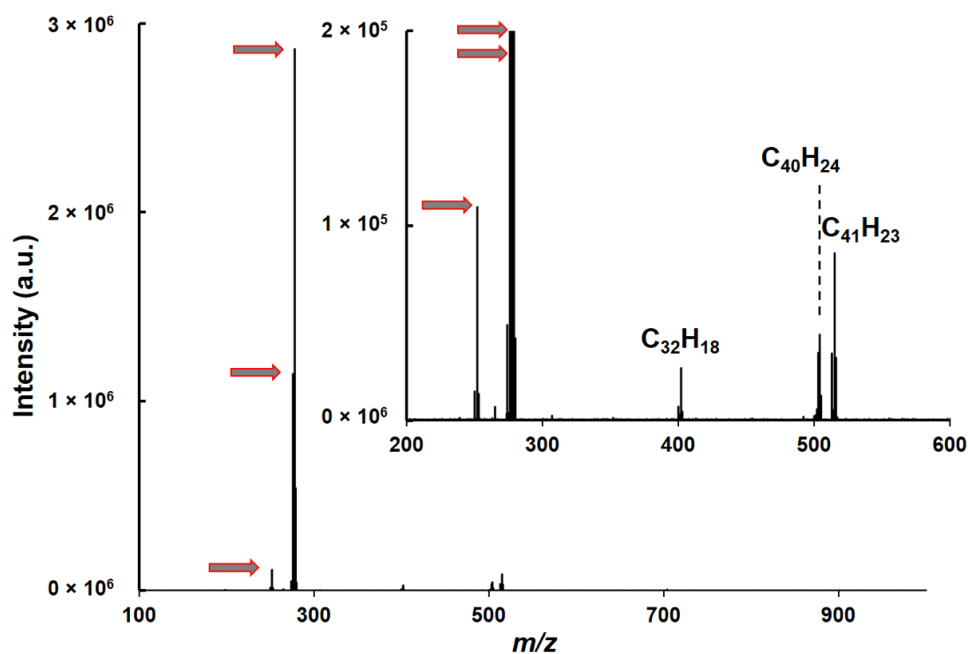


Figure S.3. Spectrum of PAH standard mix after sample deposition on aluminum foil surface. The fluence used in this acquisition was 3.18 mJ/cm^2 . The arrows indicate the parent peaks of the PAH components included in the original solution.

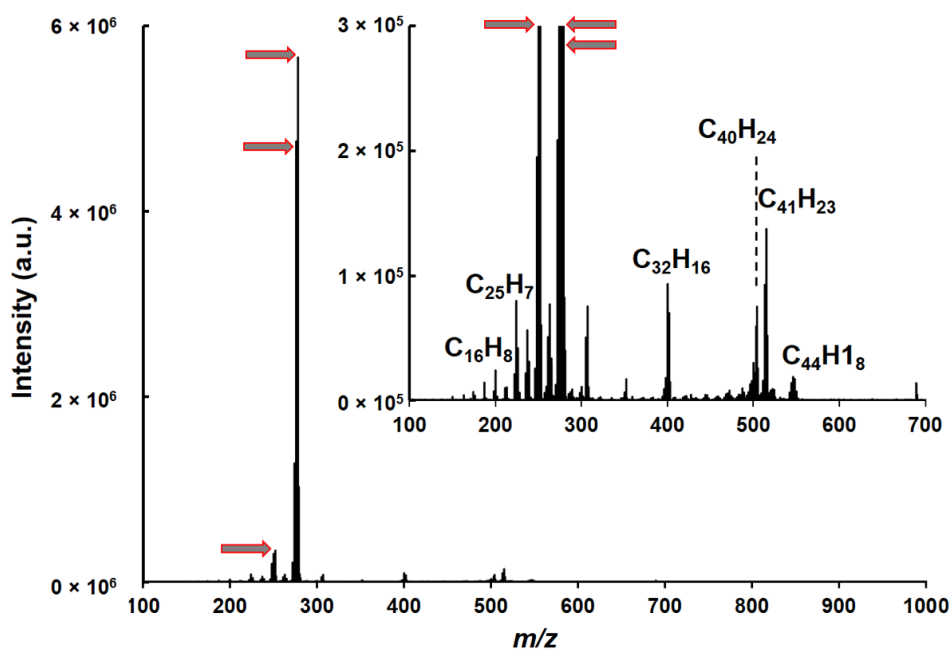


Figure S.4. Spectrum of PAH standard mix after sample deposition on aluminum foil surface. The fluence used in this acquisition was 11.94 mJ/cm^2 . The arrows indicate the parent peaks of the PAH components included in the original solution.

S1.4 Source-Generated Ions on Aluminum Foil Substrates

The fluence dependence of the coronene peak is shown in **Figure S.5**. On the secondary axis, the number of peaks that were observed within a 0.1% relative intensity threshold in all three of the triplicate spectra is also plotted for each fluence studied. There is a linear trend from 3.18 mJ·cm⁻² until 9.55 mJ/cm², but the trend drops off at 11.94 mJ·cm⁻², likely because the ion is saturated. The equation for the linear region of the plot is $y = 2.98 \times 10^6 x + 9.75 \times 10^6$, and $R^2 = 0.9978$. Although in other types of photoionization measurements, the slope of the log-log plot of laser fluence vs ion intensity has suggested the number of photons required for ionization, the slope in this case is 0.614, suggesting that this relationship is not valid for the current method of ionization as the number of 337 nm photons required to ionize coronene is greater than one. This trend may be the result of differences in the phase of the material during ionization, the fluence, or the relatively long wavelength as compared with other measurements discussed previously.

The phase of the material is defined in the gas phase for ionization energy measurements, and though the mechanism of ionization in LDI-MS analysis isn't well understood, this data suggests that the ionization is not happening in the gas phase. It is possible this relationship suggests one photon is necessary for ionization, however coronene has a known ionization energy of 7.21 ± 0.02 eV in the gas phase, and the energy of a photon of 337 nm is 3.68 eV. Meanwhile, the number of peaks that were detectable with the established parameters ranged from 15-69, which suggests a lot of potential interference peaks.

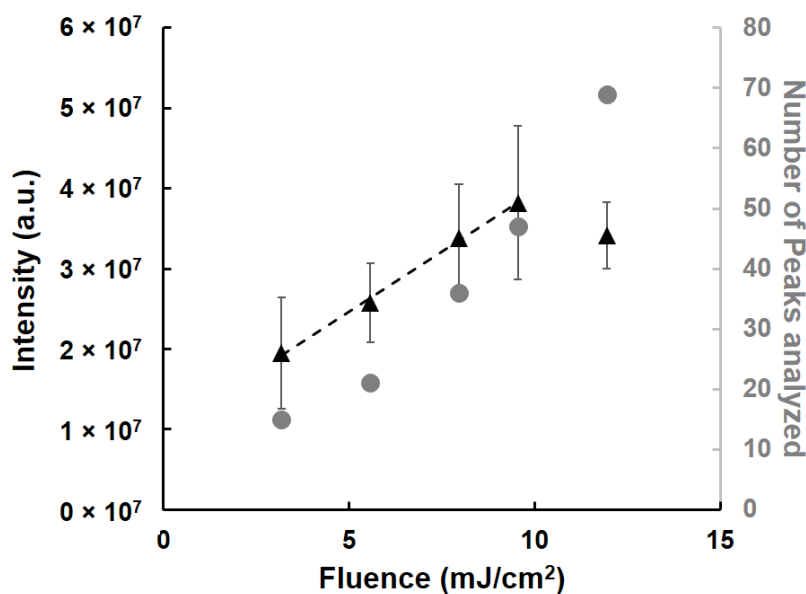


Figure S.5. Fluence dependence of the coronene molecular ion (m/z 300.0934) as a function of laser fluence is shown in black triangles. Plotted values are the average of three replicates and error bars are standard deviation. Number of peaks analyzed is plotted in gray circles.

Looking deeply into the non-coronene peaks present in both spectra, there are numerous overlapping peaks observed in all cases. Right around the coronene peak, there are also -2H and -4H peaks observed, as well as a peak corresponding to a loss of two carbons. In the higher mass range, the group of peaks around $C_{48}H_{20}$ (m/z 596.1560) are likely dimers of coronene, however to obtain that C/H ratio, these dimers are likely formed in-plane, as opposed to a stack confirmation, which would correspond to $C_{48}H_{24}$. In fact, the peak corresponding to $C_{48}H_{16}$ in **Figures S.1** and **S.2** are somewhat higher in intensity, which corresponds to the loss of four hydrogens from a $C_{48}H_{20}$ dimer. Interestingly, the intensity ratios of parent peaks, -2H peaks, and -4H peaks are different for the monomer and dimer. There are also additional peaks, including $C_{30}H_{14}$ and $C_{32}H_{14}$, which are of unclear origin. By examining the ratio between coronene and each of these other peaks as a function of fluence, the origin of the peaks may be discerned.

Figure 5.7 shows such comparison for several major non-coronene peaks. Peaks assigned as $C_{22}H_{12}$, $C_{28}H_{14}$, $C_{30}H_{14}$, and $C_{32}H_{14}$ all show approximately constant ratios with coronene regardless of the laser fluence used, suggesting they may coexist with coronene in the standard acquired from Millipore Sigma. On the contrary, $C_{24}H_{10}$, $C_{48}H_{16}$, $C_{48}H_{18}$, and $C_{48}H_{20}$ all show an increase in relative intensity as a function of fluence, which suggests that they are being formed in greater percentages as the laser fluence is increased. In other words, the former four ions are likely contaminants, whereas the latter are source-generated ions. The most common source-generated ions are losses of hydrogen, coronene dimers, or a combination of the two.

It is worth discussing that though there is a significant contribution of all the ions presented in **Figure S.6**, in addition to other ions observed, if fluence is kept to a minimum (i.e., <7.96 mJ/cm²), all the source generated ions are $<10\%$. Given the wealth of information that can be contributed to the soot composition literature by focusing on peaks with $>10\%$ relative intensity, there is still a lot to be learned from the remaining data in a more complicated soot spectrum, discussed in the text.

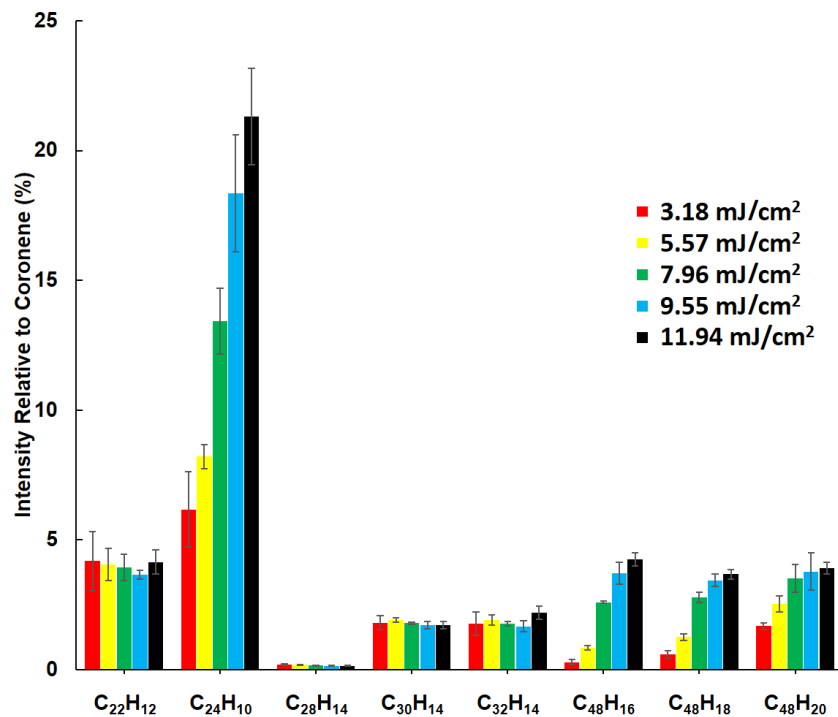


Figure S.6. Relative ion intensity, compared to coronene, from an aluminum foil substrate. Values are averaged across three replicates with error bars showing the standard deviation.

The PAH mix was analyzed similarly to the coronene standard and is presented in **Figures S.7** and **S.8**. Figure 5.8 shows the intensity of the PAH C₂₂H₁₄ with increasing fluence in the black triangles; that ion was the base peak in all spectra obtained in this group of measurements. In the linear region of **Figure S.7**, the points follow the line $y = 5.35 \times 10^5 x - 4.49 \times 10^5$, with $R^2 = 0.9622$. As the amount of material is lower in the PAH standard mix, it follows that the linear region doesn't show the same saturation effect that was observed with coronene analysis. In fact, this data suggests below 5.57 mJ·cm⁻² laser fluence, there is no effect on the ion intensity because the signal is relatively low. As shown in **Figures S.5** and **S.7**, the number of peaks identified across all replicates, shown in the gray circles in both figures are remarkably similar, which is a surprising result given the numerous additional species present in the PAH standard mix. We expected that there would be many additional peaks observed in the PAH standard mix case. As discussed above, concentration may play a role in the number of source-generated ions, as well as sensitivity of very low-abundance contaminants.

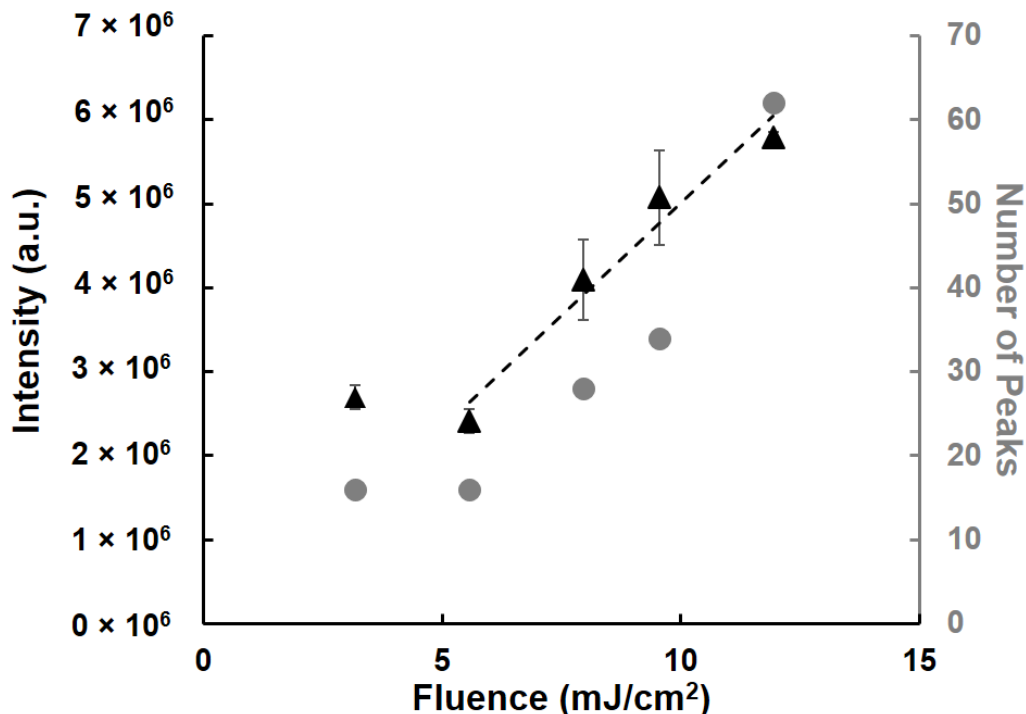


Figure S.7. Fluence dependence of the $C_{22}H_{14}$ molecular ion (m/z 278.1090) intensity from aluminum foil substrates is shown in black triangles. Values are averaged across three replicates and error bars show the standard deviation. The total number of peaks analyzed in each fluence condition across all replicates is shown by the gray circles.

The spectral features in the PAH standard mix measurements were analyzed with respect to their intensity ratio with the base peak, which in all cases was m/z 278.1090, coming from $C_{22}H_{14}$, shown in **Figure S.8**. The molecular formulae with italic labels indicate the molecular species that were known to be present in the PAH standard mix. As discussed above, the PAH standard mix contained both $C_{22}H_{12}$ and $C_{22}H_{14}$, with the former having 2× the amount of material because there were equimolar concentrations of two isomers.

Many trends observed in the data of the PAH standard mix are similar to what was observed in coronene; chiefly, the loss of 2C or 2H is observed. Though initially it appeared that the planar dimerization was again observed, these “dimer” ions, e.g. $C_{41}H_{21}$ and $C_{40}H_{24}$, did not have a fluence dependence. Possibly, these molecules are either contaminants or molecules that formed in the reference solution, however they appear to be independent of the laser power. For the ion corresponding to $C_{22}H_{12}$, which is itself present in the original solution, the lowest fluence case is the nearest representation of the intensity coming from the original $C_{22}H_{12}$ ions. As fluence is increased the intensity of the $C_{22}H_{12}$ peak becomes the sum of native $C_{22}H_{12}$ molecules with the additional intensity coming from $C_{22}H_{14}$ losing 2H. Though the data is not presented **Figure S.8**, at fluences greater than 5.57 $mJ\cdot cm^{-2}$, there are also losses of -4H observed from several PAH standard mix species. If fluence values remain sufficiently low, the contribution of source-generated hydrogen losses was low.

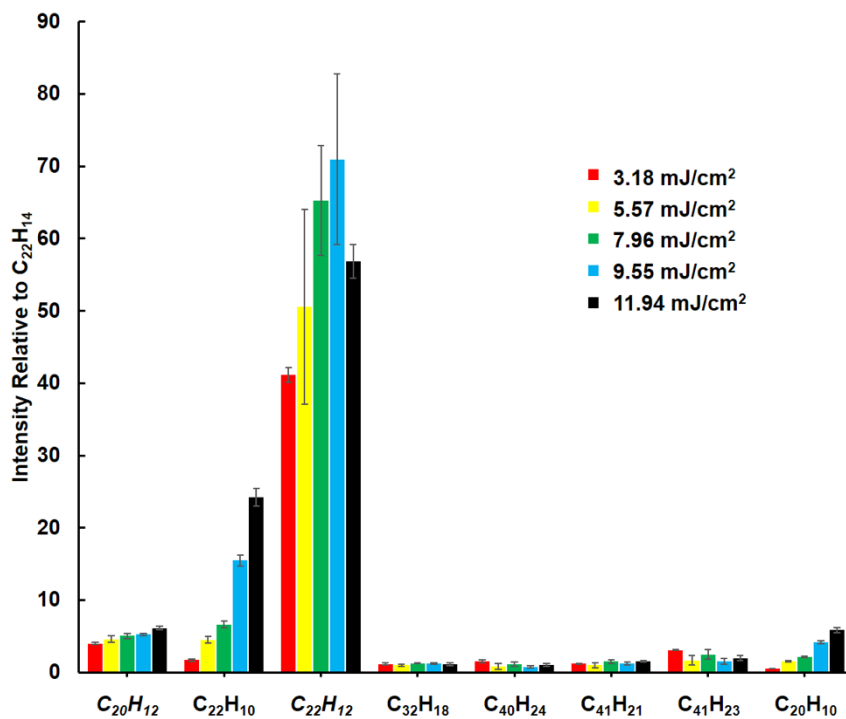


Figure S.8. Selected ion intensity relative to $C_{22}H_{14}$, averaged across three replicates with error bars showing the standard deviation, at different fluences analyzed on aluminum foil substrates.

S1.5 Spectra of Isolated and Mixed PAH Standards on Silicon Wafer Surfaces

An analytical method nearly identical to that which was described for samples on aluminum foil was used for samples analyzed on silicon wafer substrates. These wafers were not subjected to any surface treatment prior to sample deposition. The same fluence range, 3.18–11.94 $\text{mJ}\cdot\text{cm}^{-2}$ was also used in this set of experiments, and a representative spectrum for the lowest and highest coronene experiments are shown in **Figures S.9** and **5.11**, respectively. In most spectra, the overall intensity of the molecular ions is lower on silicon surfaces. The surface tension of the deposited toluene solution was lower on the silicon wafers, therefore the solution spread out more. The area interrogated during analysis was the same size regardless of the material used as a substrate, thus there would have been somewhat less material per area on the silicon wafers. The ionization efficiency of PAH on silicon wafers may be lower than on aluminum foil.

In the lowest fluence case, several similar ions are observed as were on aluminum foil, except the only dimer-related peak is the completely hydrogenated form, $C_{48}H_{20}$. Though the ion intensity does have a $S/N > 3$, it is below the threshold intensity of 0.1% used in mMass peak filtering. Similarly, there is a lower abundance of $-2H$ peaks at low fluence on silicon wafer as compared with aluminum foil. Additionally, the presence of $C_{22}H_{12}$, $C_{30}H_{12}$, and $C_{32}H_{12}$ on silicon as well does agree with the assertion that they are contaminants of the solution as opposed to source-generated ions, given the lower abundance of source-generated ions like $C_{48}H_{20}$, $C_{48}H_{18}$, $C_{24}H_{10}$, etc. Though many similar peaks are also observed in the highest fluence cases across substrates, there are fewer peaks that appear only in the respective high fluence spectra, suggesting that there are

more fluence-dependent source-generated ions when samples are analyzed on aluminum foil compared with silicon wafer.

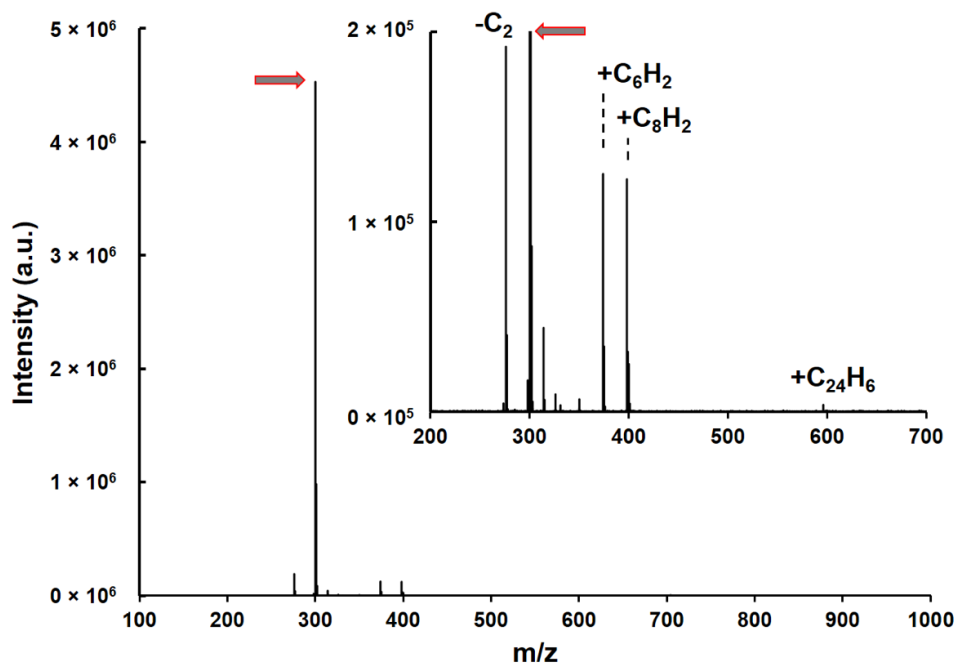


Figure S.9. Coronene spectrum after sample deposition on a silicon wafer surface. The fluence used in this acquisition was $3.18 \text{ mJ} \cdot \text{cm}^{-2}$. The arrow indicates the coronene base peak.

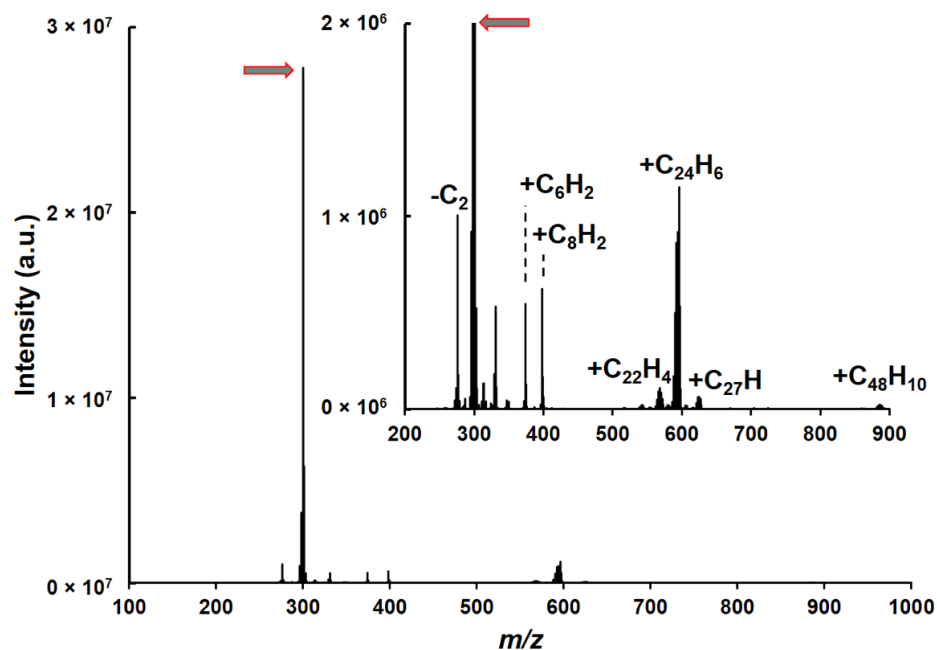


Figure S.10. Coronene spectrum after sample deposition on a silicon wafer surface. The fluence used in this acquisition was 11.94 mJ cm^{-2} . The arrow indicates the coronene base peak.

The standard mix of PAH molecules was analyzed on silicon wafer substrates, and the resulting spectra of the lowest and highest fluence used are shown in **Figures S.11** and **S.12**, respectively. The most striking significance between the two substrates is that on aluminum foil, $\text{C}_{22}\text{H}_{14}$ remained the most intense ion in all spectra regardless of fluence, whereas on the silicon wafer, the intensity of $\text{C}_{22}\text{H}_{12}$ eclipsed it substantially. This reveals that although with coronene the amount of H and C loss was reduced on silicon wafer substrates, $\text{C}_{22}\text{H}_{14}$ lost -2H at a greater rate in the silicon wafer case as fluence was increased.

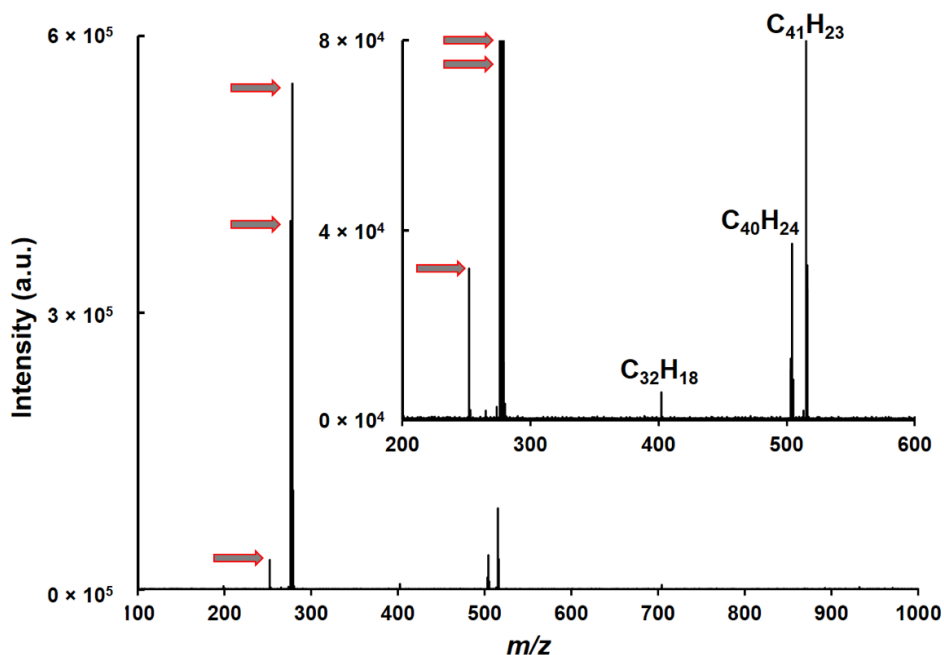


Figure S.11. Spectrum of PAH standard mix after sample deposition on a silicon wafer surface. The fluence used in this acquisition was $3.18 \text{ mJ}\cdot\text{cm}^{-2}$. The arrows indicate the PAH components known to be included in the original solution.

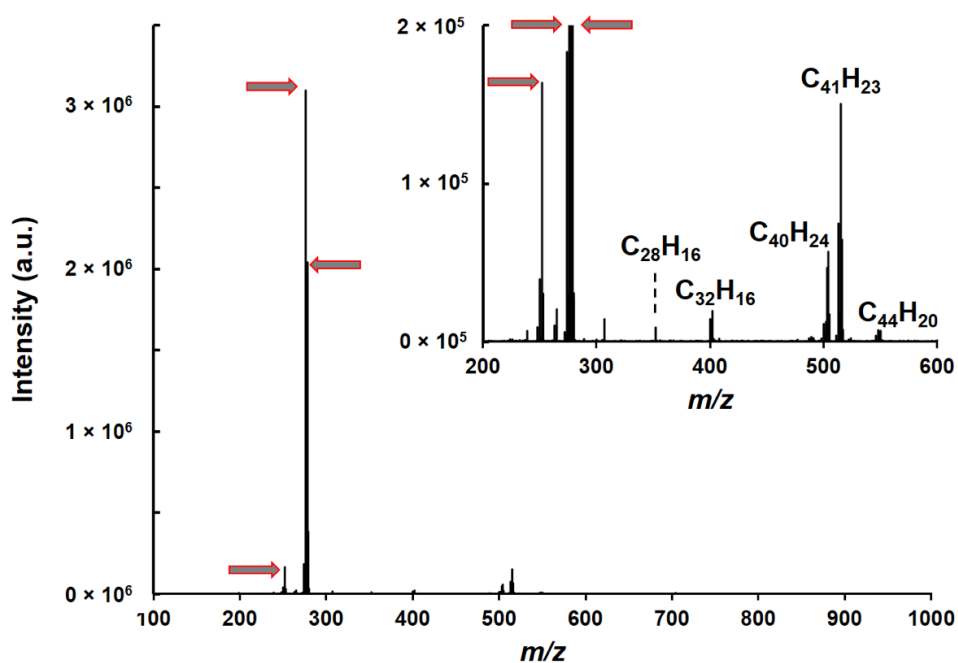


Figure S.12. Spectrum of PAH standard mix after sample deposition on a silicon wafer surface. The fluence used in this acquisition was $11.94 \text{ mJ}\cdot\text{cm}^{-2}$. The arrows indicate the PAH components known to be included in the original solution.

S1.6 Source-Generated Ions on Silicon Wafer Substrates

To more thoroughly examine peaks aside from the coronene molecular ion, the intensity of that peak was plotted as a function of laser fluence. The number of peaks identified was also monitored as a function of fluence alongside the intensity of coronene ion signal in **Figure 5.14**. **Figure 5.14** also shows that the linear range of fluences used on silicon wafer substrates spans the entire fluence range studied, and the data follows the equation $y = 2.81 \times 10^6 x - 1.51 \times 10^6$ although the R^2 value is only 0.958 as compared with **Figure 5.6**. The increased dynamic range is likely due to the reduced sample concentration. Converting the values shown in Figure 5.14 to a log-log scale as was performed and resulted in a slope of 1.063. Although the slope is greater than one, the ionization energy of coronene should require at least 2 photons.

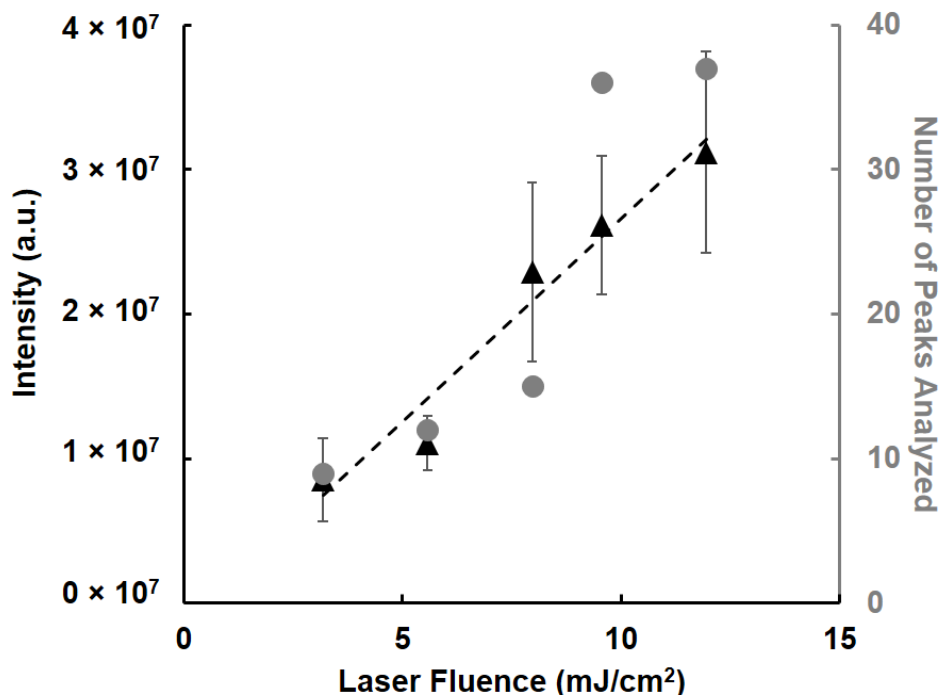


Figure S.13. Fluence dependence of the coronene molecular ion (m/z 300.0934) signal intensity is shown in black triangles. Values are averaged across three replicates and error bars are standard deviation. The total number of peaks analyzed is shown by the gray circles.

In comparison with the analysis on aluminum foil, there are nearly half as many peaks detected from the silicon wafer substrate even in the highest fluence case: 37 on Si, 69 on Al. This translates to a cleaner spectrum with more identifiable peaks. Looking more deeply into the identities of these peaks, it is clear that on aluminum foil, we observed both more H and C loss than on silicon wafer substrates. The results of evaluating similar ions to **Figure S.6** are presented in **Figure S.14**. In the coronene spectra from the silicon wafer substrate, all of the fluence cases show very minimal H loss or dimerization, yet the ions thought to be contaminants, e.g. $C_{22}H_{12}$, $C_{30}H_{14}$, and $C_{32}H_{14}$ all remain in similar ratio with coronene regardless of the laser fluence.

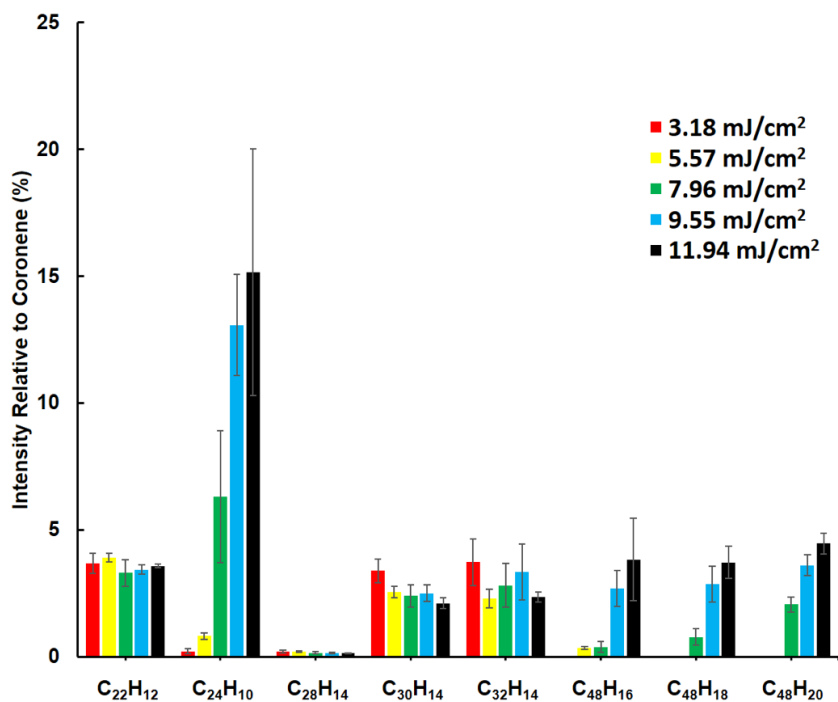


Figure S.14. Selected ion intensity relative to coronene from a silicon wafer surface, averaged across three replicates with error bars showing the standard deviation, at different fluences.

A plot of C₂₂H₁₄ ion (m/z 278.1090) was also tracked across all fluence cases and is presented in **Figure S.15**. Similar to **Figure S.7**, the linear range of this plot does not hold through to the lowest fluence studied. The linear range follows the equation $y = 1.81 \times 10^5 x - 3.76 \times 10^5$, with an $R^2 = 0.9940$. The number of peaks identified was incorporated into **Figure S.15**, and similar to coronene spectra on aluminum foil compared with silicon wafer, there were less than half as many peaks observed with the introduction of silicon.

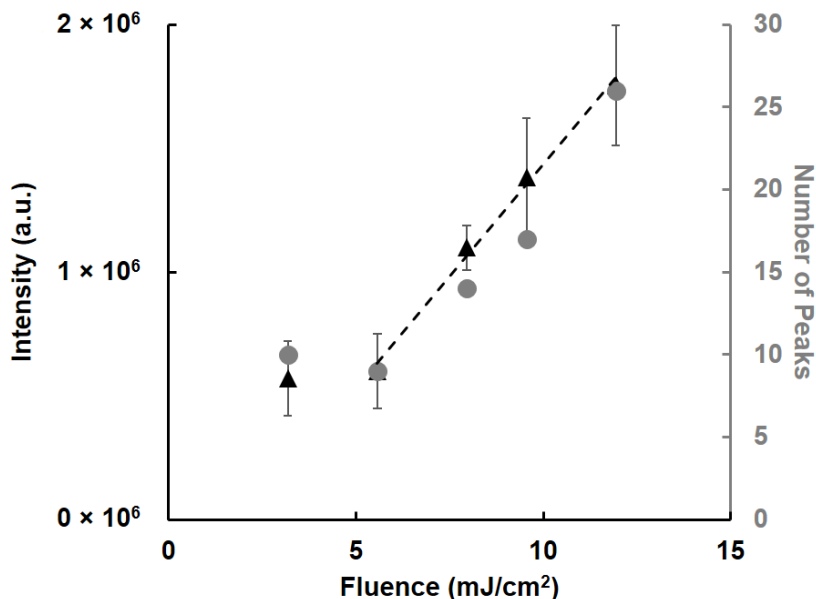


Figure S.15. Fluence dependence of the $C_{22}H_{14}$ molecular ion (m/z 278.1090) intensity from silicon wafer substrates is shown in black triangles. Values are averaged across three replicates and error bars show the standard deviation. The total number of peaks analyzed in each fluence condition across all replicates is shown by the grey circles.

Other peaks in the PAH mix spectrum obtained from the silicon wafer substrates are highlighted in **Figure S.16** for comparison with **Figure S.8**. The intensity of $C_{22}H_{12}$ at the lowest fluence was also higher on silicon wafers, therefore there is a more equal ionization rate of $C_{22}H_{12}$ and $C_{22}H_{14}$ on the silicon wafer than on aluminum foil. This is further exemplified by the average relative intensity $C_{22}H_{12}/C_{22}H_{14}$ was $69.8 \pm 5.4\%$ on silicon wafer, whereas the same ratio was $41.2 \pm 1\%$ on aluminum foil.

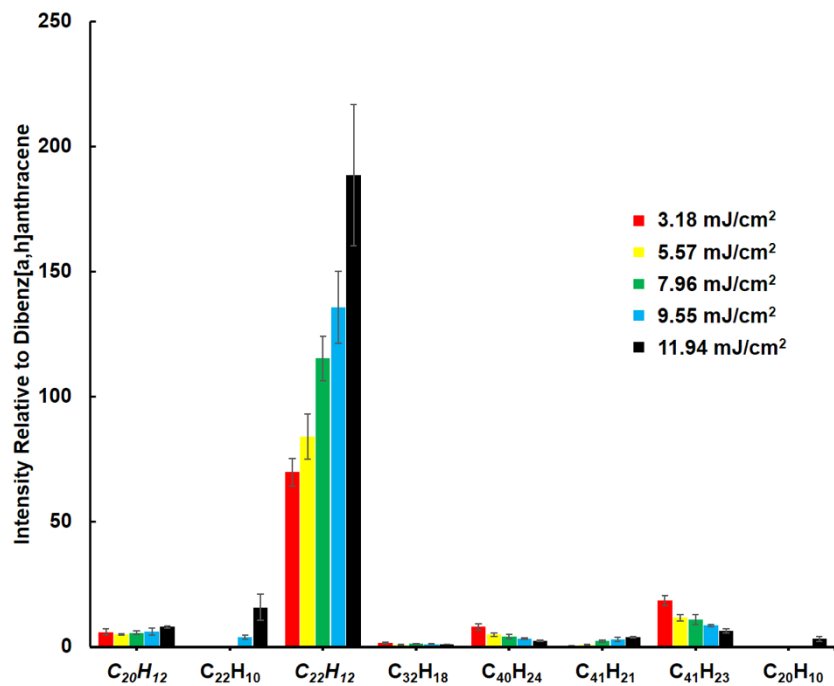


Figure S.16. Selected ion intensity relative to $C_{22}H_{14}$, averaged across three replicates with error bars showing the standard deviation, at different fluences analyzed on silicon wafer substrates.

S1.7 Fullerene Formation

Carbon clusters, or more specifically fullerenes, are observed in mass spectra of carbon rich material. The 1996 Nobel Prize in chemistry was awarded to Professors Curl, Kroto, and Smalley for discovering C_{60} with a similar experiment as performed here. In those experiments, laser light was focused onto graphitic surfaces, and extensive follow-up research was performed.[11, 12] With the fluence conditions discussed above, there were no clusters of carbon observed in the mass spectrum. Importantly, when the laser fluence was increased sufficiently, there were many carbon cluster ions observed even when coronene was the only primary analyte. **Figure S.17** shows a spectrum of the same coronene solution discussed above on aluminum foil, with $15.92 \text{ mJ}\cdot\text{cm}^{-2}$ laser fluence, where many carbon cluster ions are clearly visible. These carbon clusters are separated by 24 m/z , or even numbered carbon clusters observed, between C_{58} and the maximum mass range studied in this case, which was raised to 2000 m/z . Although C_{60} and C_{70} are not the observed with the highest intensity compared with other clusters, they are greater than adjacent peaks, which is also like the results observed by Kroto and coworkers. On aluminum foil, the intensity of these carbon clusters is approximately 5-10% of the coronene peak.

On a silicon wafer substrate, carbon clusters are again observed under the same fluence conditions, though they have a lower relative abundance. **Figure S.18** shows a spectrum of coronene on a silicon wafer also analyzed by $15.92 \text{ mJ}\cdot\text{cm}^{-2}$ laser fluence. Though the same carbon cluster ions are observed, the intensity of those clusters is approximately

0.1-0.7% of the coronene peak, which is a substantial improvement over aluminum. The coronene trimer peaks are observed in **Figure S.18** more clearly in the inset than in **Figure S.17** only because the carbon clusters in the latter are more intense. Likewise, on the silicon wafer, the C_{70} peak is comingled with the trimer ions and is unlabeled for clarity.

This analysis clearly shows that carbon clusters, including C_{60} and C_{70} , are capable of being preferentially formed using excess laser fluence in this regime. The fluence must be closely monitored when more complex PAH mixtures (e.g. soot) are analyzed by this technique to prevent these source-generated carbon clusters from interfering with spectra. As these clusters are formed with only one small PAH present in the original analyte, a rich, complex mixture of larger PAH will result in more carbon cluster formation with excessive laser fluence.

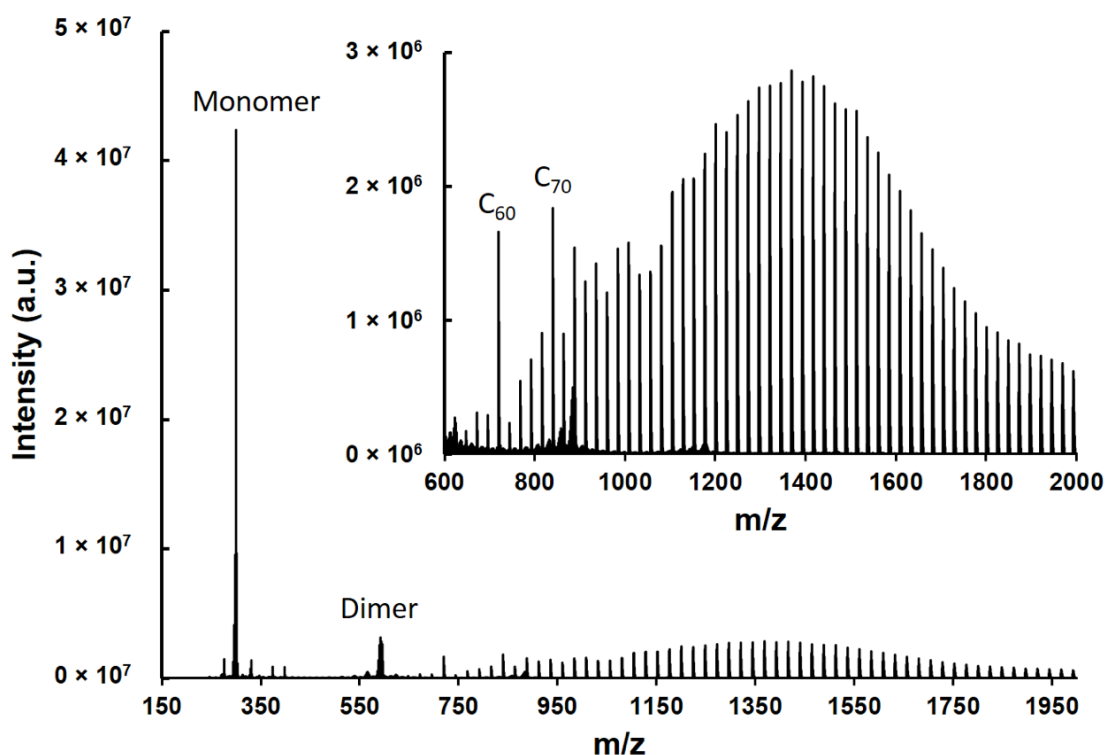


Figure S.17. Spectrum of coronene solution on an aluminum foil substrate analyzed with $15.92 \text{ mJ}\cdot\text{cm}^{-2}$ laser fluence. Inset shows clear distribution of carbon clusters. With higher laser fluence, there is an increased presence of H-loss, peaks associated with coronene monomers and dimers.

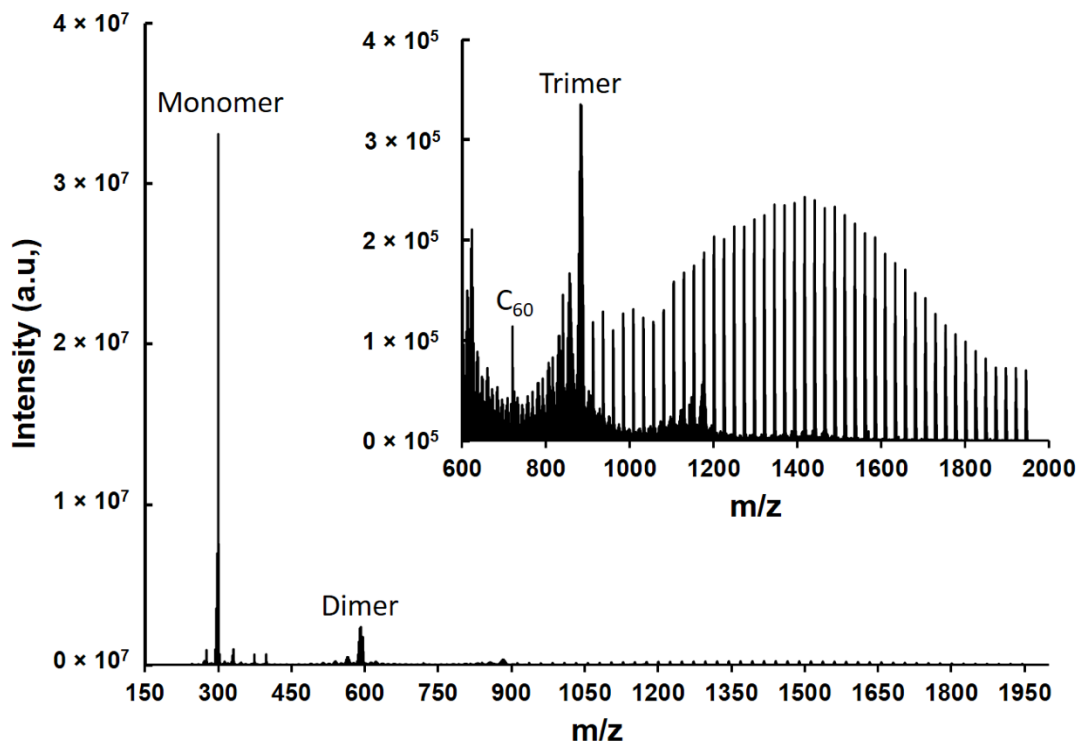
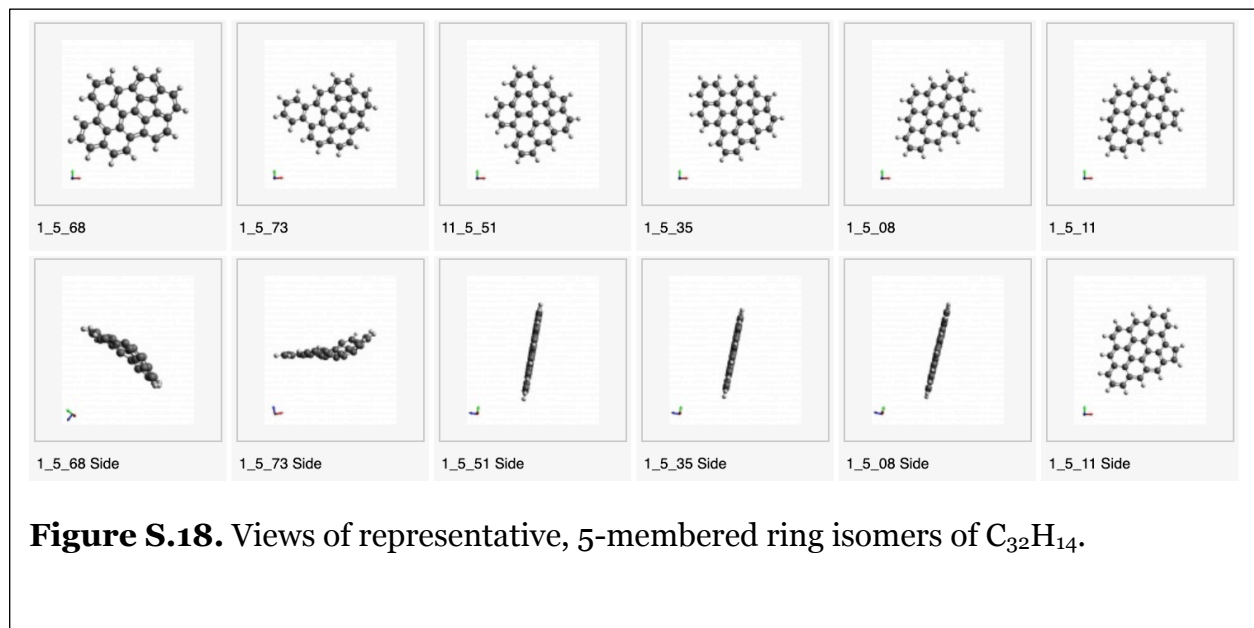


Figure S.18. Spectrum of coronene solution on a silicon wafer substrate analyzed with $15.92 \text{ mJ}\cdot\text{cm}^{-2}$ laser fluence. Inset shows clear distribution of carbon clusters. With higher laser fluence, there is an increased presence of H-loss peaks associated with coronene monomers and dimers.

S.2 Structures of 5-membered ring-containing PAH considered in DFT analysis of thermochemistry



Supplemental References

- [1] M.D. Smooke, M.B. Long, B.C. Connelly, M.B. Colket, R.J. Hall, *Combustion and Flame*, 143 (2005) 613-628.
- [2] E.M. Adkins, J.H. Miller, *Phys Chem Chem Phys*, 17 (2015) 2686-2695.
- [3] M. Strohalm, M. Hassman, B. Košata, M. Kodíček, *Rapid Commun. Mass Spectrom.*, 22 (2008) 905-908.
- [4] S.-S. Cai, J.A. Syage, K.A. Hanold, M.P. Balogh, *Anal. Chem.*, 81 (2009) 2123-2128.
- [5] E. Berger-Nicoletti, F. Wurm, A.F.M. Kilbinger, H. Frey, *Macromolecules*, 40 (2007) 746-751.
- [6] S.M. Chen, H.Z. Zheng, J.N. Wang, J. Hou, Q. He, H.H. Liu, C.Q. Xiong, X.L. Kong, Z.X. Nie, *Anal. Chem.*, 85 (2013) 6646-6652.
- [7] S.F. Macha, P.A. Limbach, P.J. Savickas, *J. Am. Soc. Mass. Spectrom.*, 11 (2000) 731-737.
- [8] A.D. Warren, U. Conway, C.J. Arthur, P.J. Gates, *J. Mass Spectrom.*, 51 (2016) 491-503.
- [9] S.Y. Xu, Y.F. Li, H.F. Zou, J.S. Qiu, Z. Guo, B.C. Guo, *Anal. Chem.*, 75 (2003) 6191-6195.
- [10] J. Cain, A. Laskin, M.R. Kholghy, M.J. Thomson, H. Wang, *PCCP*, 16 (2014) 25862-25875.
- [11] H.W. Kroto, A.W. Allaf, S.P. Balm, *Chem. Rev.*, 91 (1991) 1213-1235.
- [12] H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, R.E. Smalley, *Nature*, 318 (1985) 162.

