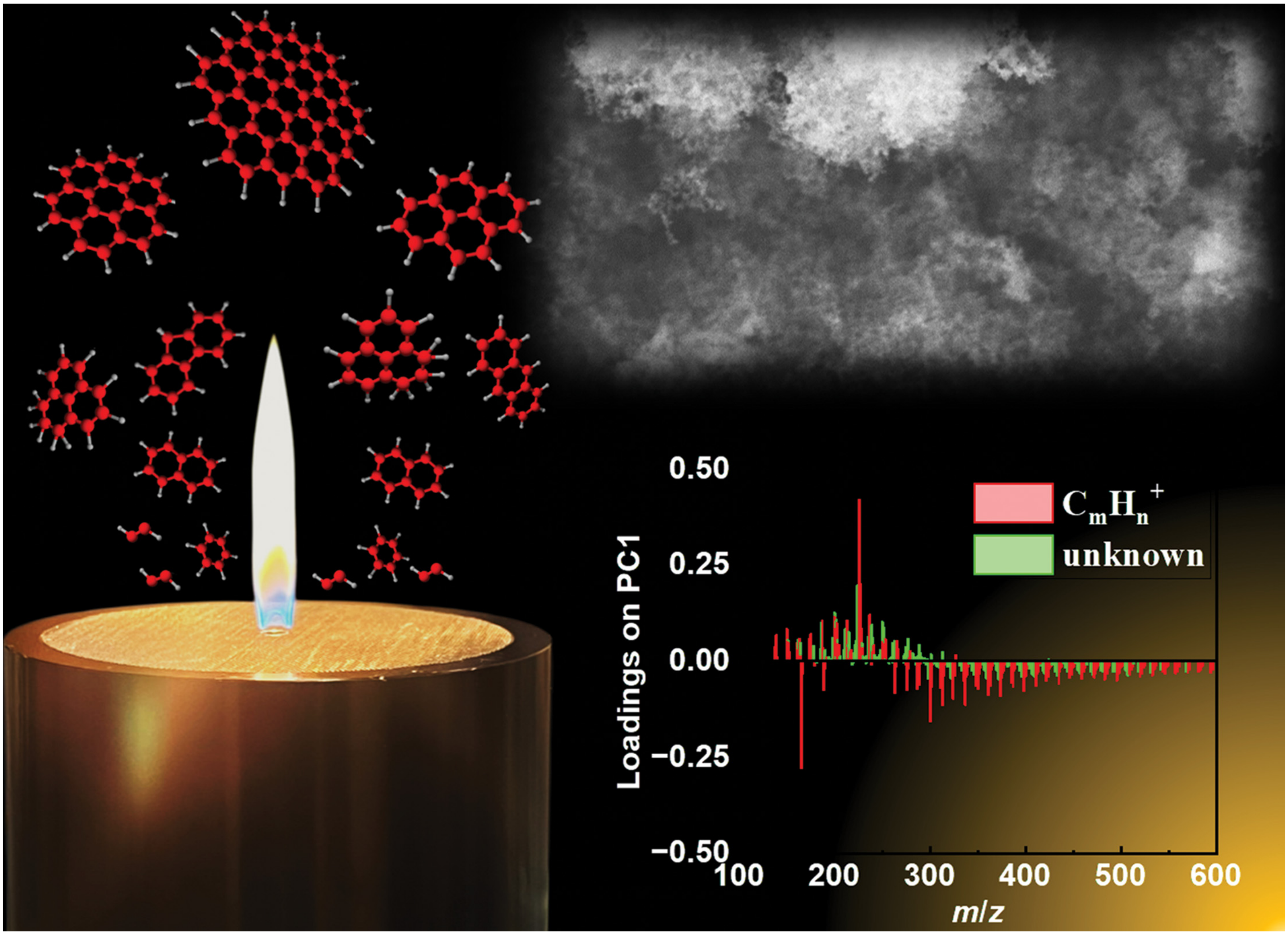


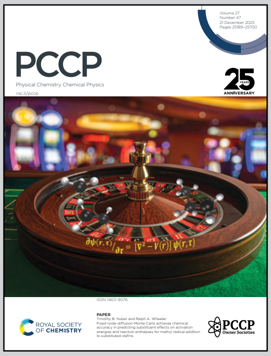
Showcasing research from the University of Lille and George Washington University.

Ex situ characterization of the precursors of incipient nanoparticles in a laminar diffusion flame of ethylene

The light from a candle masks the light source's labyrinthine journey from small molecular fragments of carbon and hydrogen to particles, each with tens of thousands of carbon atoms, and within a fraction of a second. This work presents a multi-diagnostic approach to identify molecular precursors involved in the inception process of carbon nanoparticles in flame combustion. New experimental evidence suggests that small molecular clusters are initially bound by physical forces then rapidly stabilized by the formation of C–C covalent bonds.

Image reproduced by permission of Harsh Chaliyawala from *Phys. Chem. Chem. Phys.*, 2025, **27**, 25300.

As featured in:



See Harsh Chaliyawala *et al.*, *Phys. Chem. Chem. Phys.*, 2025, **27**, 25300.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/pccp

Registered charity number: 207890

PCCP Owner Societies

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 25300

Ex situ characterization of the precursors of incipient nanoparticles in a laminar diffusion flame of ethylene

Harsh Chaliyawala,[a] Erin McCaughey,[b] Xavier Mercier,[a] J. Houston Miller,[b] Myriam Moreau,[c] Nicolas Nuns,[d] Pardis Simon[e] and Alessandro Faccinetto*[a]

Detailed information on the chemical and physical properties of the precursors of incipient carbon nanoparticles (CNPs) in flame combustion provides clues on the reaction pathways for CNP formation and growth. Therefore, this data is needed by the modeling community who seeks to understand inception at a fundamental level. However, identifying and isolating the precursors in a reactive environment at high temperature remains a challenging task. The present work reports a multi-diagnostic approach to identify the molecular species involved in the inception of CNPs based on the comparative analysis of surface morphology (scanning electron microscopy), chemical composition (time of flight secondary ions mass spectrometry), chemical state (X-ray photoelectron spectroscopy), and structure (Raman spectroscopy) of samples extracted from a nitrogen-diluted ethylene laminar diffusion flame stabilized on a Yale burner. Statistical analysis enables the reduction of the pool of species to be considered by showing, for instance, that large polyaromatic molecules are not required for the CNP inception to occur. Several low m/z species are identified as likely candidates that are consistent with polycyclic aromatic hydrocarbons (PAHs) and their derivatives. Among them, this work stresses the importance of species slightly above the curve representing the maximally condensed aromatics and of non-benzenoid PAHs (containing 5-member aromatic rings for instance). This new experimental evidence reveals trends consistent with the "combined physical and chemical inception" group of hypotheses, according to which small clusters (typically dimers) of PAHs initially bound by physical forces are rapidly stabilized by the formation of C–C covalent bonds according to various postulated mechanisms (extended HACA, spin localization).

Received 7th July 2025,
Accepted 6th October 2025

DOI: 10.1039/d5cp02594j

rsc.li/pccp

1. Introduction

The molecular-level understanding of carbon nanoparticle inception in combustion environments represents a fundamental challenge at the intersection of physical chemistry, materials science, and combustion physics. Despite decades of research into the chemical kinetics and thermodynamics governing soot formation, the precise molecular mechanisms underlying the transition from gas-phase polycyclic aromatic hydrocarbons (PAHs) to incipient carbon nanoparticles (CNPs) in the condensed phase remains incompletely understood.[1,2]

The physical chemistry of CNP inception involves the complex interactions of molecular clustering, chemical reactivity, and phase transitions under extreme conditions of temperature and chemical environment. Recent hypotheses have proposed several pathways including physical clustering followed by chemical stabilization,[3,4] the formation of fullerene-like structures from large PAHs *via* chemical growth mechanisms,[5] and chemical coalescence of mid-sized PAHs into cross-linked three-dimensional structures through the addition of PAHs with PAH radicals.[6,7] Extensive kinetic modeling on the formation of PAHs in various ethylene and methane flames provides a quantitative measurement of precursors and possible reaction pathways besides well-known mechanisms like the HACA.[8–12] However, experimental validation of these hypotheses requires molecular-level characterization of the precursor species and their transformation products, which presents significant analytical challenges

---

[a] *Univ. Lille, CNRS, UMR 8522 - PC2A - F-59000 Lille, France. E-mail: alessandro.faccinetto@univ-lille.fr*

[b] *Department of Chemistry, George Washington University, Washington DC 20052, USA*

[c] *Univ. Lille, CNRS, UMR 8516 - LASIRe - F-59000 Lille, France*

[d] *Univ. Lille, CNRS, INRAE, Centrale Lille, Univ. Artois, - IMEC - F-59000 Lille, France*

[e] *Univ. Lille, CNRS, UMR 8181 - UCCS -, Centrale Lille, Univ. Artois, F-59000 Lille, France*

due to the reactive nature of the flame environment and the transient existence of intermediate species.

Many of the recent measurements to study CNP inception represent either ensemble-averaged properties of particles samples[13] or are biased toward specific molecular subsets within the complex mixture.[14] Experimental data providing species specific molecular identification, critical for detailed chemical mechanistic modeling and validation, remains rarer. Techniques such as laser induced incandescence (LII),[13] laser induced fluorescence (LIF),[15] and transmission electron microscopy (TEM)[16] excel at characterizing mature particles, but offer limited insight into the earliest stages of particle formation. Conversely, gas-phase *ex situ* methods like laser desorption mass spectrometry (LDI-MS)[17] and time of flight secondary ions mass spectrometry (ToF-SIMS)[18] can identify individual molecular species but often struggle to capture heterogeneous processes at the gas–solid interface during inception. Additionally, while radical-mediated inception mechanisms have gained theoretical and computational support, direct experimental evidence of persistent radicals in flames has been limited.

The work presented here addresses these fundamental questions through a multi-technique approach that combines structural characterization (Raman spectroscopy), surface chemical analysis (ToF-SIMS), surface chemical state (X-ray photoelectron spectroscopy, XPS) and morphological examination (scanning electron microscopy, SEM) of CNP precursors extracted from a well-characterized laminar diffusion flame. By applying statistical analysis to correlate observations across these complementary techniques, we aim to identify the molecular signatures associated with CNP inception.

Our investigation focuses on an axially symmetric, non-premixed nitrogen-diluted ethylene flame stabilized on a "Yale" burner, a system that has been extensively characterized by both experimental and computational modeling efforts.[19–25] The choice of this flame system allows more direct comparisons with existing data, while providing a well-controlled environment for isolating the signals of interest. Through *ex situ* analysis of samples extracted at different heights above the burner (HABs), we track the evolution of chemical composition and molecular structure as the system progresses from gas-phase PAH precursors towards mature CNPs.

The physical chemistry insights gained from this work have broader implications for understanding the inception and growth processes. Furthermore, the analytical methodologies used here demonstrate the power of combining multiple characterization techniques with statistical analysis to extract molecular level information from complex, heterogeneous systems.

2. Materials and methods

2.1. Flame and sampling

A 50 mm visible flame high, 60% ethylene-in-nitrogen co-flow laminar diffusion flame is stabilized on a "Yale" burner at atmospheric pressure. The ethylene/nitrogen mixture flows through a 0.4 cm inner diameter central injector surrounded by a 7.5 cm inner diameter honeycomb for the air co-flow (see SI). The fuel and air velocity at the burner surface are stabilized at an average of 35 cm s^{-1}. This flame was the object of several past investigations.[13,19–23] In this work, CNPs and condensable gas are extracted from four different HABs (20, 25, 30, and 35 mm) from the central axis of the flame using a dilutive extraction microprobe[26] and deposited on 10 × 10 × 1 mm Ti (99.5%, American Elements) substrates for *ex situ* analyses (pictures of the samples are available in the SI). These HABs are chosen based on previous LII measurements showing the soot inception region.[13] Impaction at high velocity ($>$30 m s^{-1}) results in the aggregation of the carbonaceous particles on the center of the substrate, while condensable gases scatter over the surface of the substrate. Particles that bounce off the impaction site create an additional ring-like structure.

2.2. SEM

Scanning electron microscopy with energy dispersive X-ray spectroscopy (SEM/EDX) is used to study the evolution of soot morphology at various HABs. Images are taken with a JEOL JSM-7800F LV scanning electron microscope equipped with an FEG source at 15k magnification at 5 kV with a 7–8 mm working distance from the source gun. EDX analysis is also performed on some of the samples. However, since H cannot be detected with EDX, the usefulness of this diagnostic is limited to assisting the interpretation of the mass spectra by determining the level of impurities in the samples.

2.3. Raman spectroscopy

Raman spectroscopy analyses are performed with a Horiba HR800 microscope equipped with a 40× objective (NA = 0.7) at λ_{ex} = 325 nm excitation wavelength. The acquisition time is fixed to 15 min per spectrum, and the laser power at the sample is limited to around 0.1 mW to avoid structural modifications. The laser spot measured at the sample surface is 1.7 μm. The measurements discussed herein are performed only on the impaction spot regions: 3–4 spectra are recorded for each sample, which are fitted using the peak analyzer tool included in Origin Pro 2024. The best results ($R^2 \geq 0.95$) are obtained with six peak fits: D4, D5, D1, D3, G and D2. As detailed in our previous work, a mix of Gaussian (D3) and Lorentzian (D4, D5, D3, G and D2) waveforms is used for the fitting.[18]

The baseline is fitted together with the Raman peaks (no baseline subtraction is performed) using a 3rd-order polynomial function for all spectra.

2.4. ToF-SIMS

Time of flight secondary ions mass spectrometry (ToF-SIMS) is an analytical technique widely used for the characterization of the chemical composition of the surface of solid materials and can probe surface layers as thin as 2–3 nm. In this work, a Bi_3^+ primary ion beam (25 keV, 0.3 pA) is used to sputter and ionize the surface of the deposited samples. These settings, resulting in the total ion dose ($\sim 10^{11}$ ions cm^{-2}) being well below the limit of static mode, are suited for the analysis of high mass organic species ($m/z > 200$). The generated secondary ions are accelerated and analyzed in a time-of-flight mass spectrometer with $m/\Delta m \approx 10^4$ nominal resolving power. Mass spectra in positive polarity are recorded at 50 scans/acquisition on a 500 × 500 μm^2 surface with an image resolution of 128 × 128 pixels. Acquisitions are performed on the impaction spot region. The collected mass spectra are first aligned, calibrated and normalized by the total useful ion count. A detailed description of the post-treatment[27] is available. Briefly, the full peak list contains 1115 peaks having SNR ≥ 3, of which 687 are identified by mass defect analysis and assigned a molecular formula. After removal of background and fragment ions, 393 peaks are retained.

2.5. XPS

X-ray photoelectron spectroscopy (XPS) is used to investigate the sp^2/sp^3 ratio of soot at various HABs. XPS spectra are recorded with a Kratos Axis Ultra DLD spectrometer equipped with a monochromatized Al Kα X-ray source working at 225 W. The base pressure of the instrument is 10^{-9} mbar and the analysis area on the sample 300 × 700 μm^2. The instrument charge neutralization system is used during all acquisitions. C 1s core level spectra and C KLL Auger peak are recorded at 20 eV and 80 eV fixed Pass Energy with 0.1 eV and 0.5 eV energy step. Peak fitting is done using CasaXPS software.[28] Shirley type backgrounds are subtracted to the C 1s photopeak and 70% Gaussian and 30% Lorentzian product function are used except for the sp^2 peak for which an asymmetrical Lorentzian line shape is used. Binding energy scales are adjusted on the sp^2 carbon peak positioned at 284.4 eV.

2.6. Multivariate analysis

Principal component analysis (PCA) is used to reduce the dimensionality of the Raman and ToF-SIMS databases and as a support for the interpretation of the variance and the peak assignment and selection. The Raman peaks normalized by the total area of the Raman spectrum, and the mass spectra ion counts normalized by the total useful ion count are used as variables, while acquisitions on different zones of the substrates are used as observables.[29] The results are discussed in terms of scores and loadings to find correlations between the HAB and the structure of the samples (Raman), and the changes in the chemical composition of the sample directly associated with CNP formation (ToF-SIMS).

3. Results and discussion

3.1. Surface morphology evolution

SEM images, shown in Fig. 1, display the soot morphological evolution along the flame central axis with increasing the flame height from 20 to 35 mm HAB. The analysis of the surface morphology shows no soot deposits below 20 mm HAB (not shown in the figure). Smooth droplet-like deposits can be seen from 20 mm HAB, and fractal-like mature soot particles from 30 mm HAB.[30] Likewise, the substrate surface surrounding the impaction location is analyzed as shown in the SI: almost no traces of fractal-like soot particles are observed on the substrate surface outside of the central impaction zone. In addition, the elemental composition is determined by EDX in order to detect the presence of impurities on different regions of the samples at various HABs and assist the creation of the ToF-SIMS peak list. Si impurities, later identified as siloxanes, are found in four samples on the spot region.

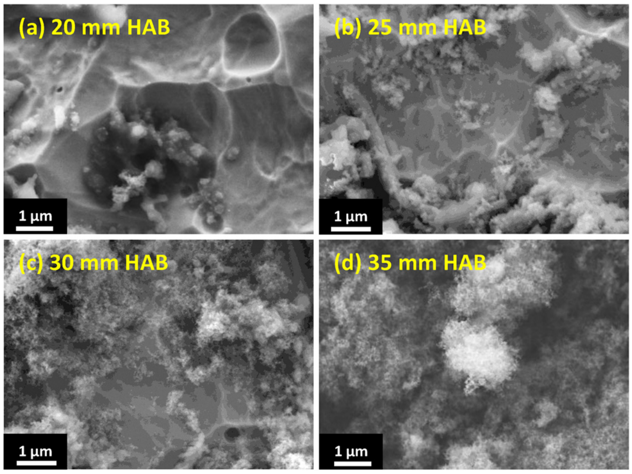


Fig. 1 SEM morphology images at (a) 20, (b) 25, (c) 30 and (d) 35 mm HAB. Notice the presence of fractal-like soot aggregates visible from 30 mm HAB that are absent at 20 mm HAB.

3.2. On the structure

Fig. 2 shows the Raman spectra of samples at λ_{ex} = 325 nm. The first-order feature of the Raman spectra, consisting of the D and G bands in the 900–2000 cm^{-1} spectral range, is analyzed by peak fitting to extract information on the carbonaceous materials from these partially overlapped bands. At this λ_{ex}, no significant broad fluorescence emission is observed. The two typical broad bands at ~1350 cm^{-1} (D or defect band) and ~1600 cm^{-1} (G or graphitic band) can be observed in all the spectra. As mentioned above, the approach proposed herein is based on a six-peak fitting (D4, D5, D1, D3, G, D2) and focuses on finding correlations between spectral parameters as a function of the HAB to evaluate the structural behavior of the CNPs. The physical significance and the detailed interpretation of each band can be found elsewhere.[31,32] Briefly, peaks D1, G, and D2 have been assigned, respectively, to the A$_{1g}$ mode of graphene only active for lattice defects, to the E$_{2g}$ mode of the ideal graphitic lattice, and to graphene layers not sandwiched between other layers. Peaks D4, D5 and D3 have not yet been assigned to specific vibration modes: proposed interpretations in the literature include contributions from ionic impurities, peripheric sp^2 or sp^3 carbon atoms, finite crystal size, and poorly linked polyaromatic units. Furthermore, the I_{D1}/I_G ratio is estimated at every HAB. It is found that the I_{D1}/I_G ratio slightly increases from 0.80 at 20 mm HAB up to 0.93 at 35 mm HAB. This may be due to the creation of lattice defects, progressing towards more amorphization carbon structure, consistent with the reported literature.[31,32]

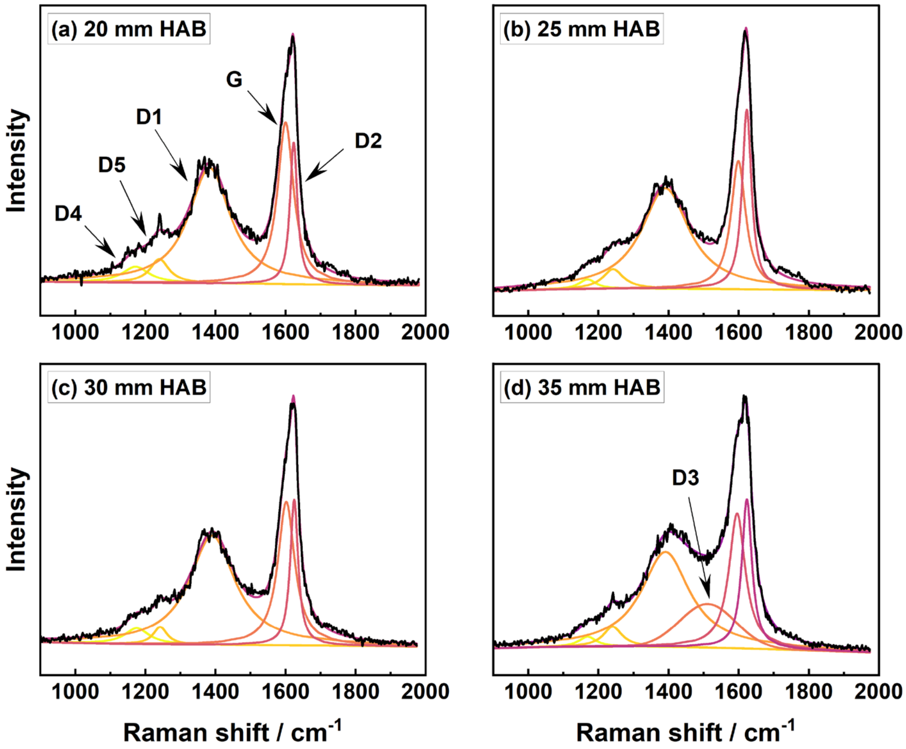


Fig. 2 Fitted Raman spectra obtained from samples collected at (a) 20, (b) 25, (c) 30 and (d) 35 mm HAB at λ_{ex} = 325 nm. The spectra include the raw data (black line), and the fitted peaks (D4, D5, D1, D3, G and D2, in color).

PCA is performed on the Raman normalized areas A_n, including all available data at λ_{ex} = 325 nm and various HABs. The result is plotted in Fig. 3. In the score plot in Fig. 3(a), the datapoints represents the Raman spectra: as shown in the figure, two main clusters of datapoints based on the sampling HAB are found: datapoints at 20–25–30 mm HAB (with one exception) have negative scores on PC1, while datapoints at 35 mm HAB have positive scores on PC1. These two data clusters represent the tendency of A_n to assume constant but different values at lower and higher HAB. According to the PCA, 35 mm HAB are well separated from the remaining HABs, to indicate that there is a net dependency of the structure of the samples on the HAB.

The loading plot of PC1 in Fig. 3(b) identifies the peak D3, which has the highest positive loadings on PC1, as the main responsible for the observed dependency on the HAB. Peak D3 is centered around 1530 cm^{-1} between peaks D1 and G, and only occurs high in the flame at 35 mm HAB by forming a wider overlapping region than at lower HAB. The appearance of peak D3 is therefore a strong indicator for the appearance of mature and probably partially graphitic soot particles.[18,33–35] Unfortunately, its origin is not yet clarified. Possibly a convolution of several other modes,[36] peak D3 is proposed to be associated to vibrations of amorphous carbon,[33,34,37] finite-size graphite crystals[35] or defects outside the graphene layers.[38] In particular, similar trends have been observed in a laminar diffusion flame

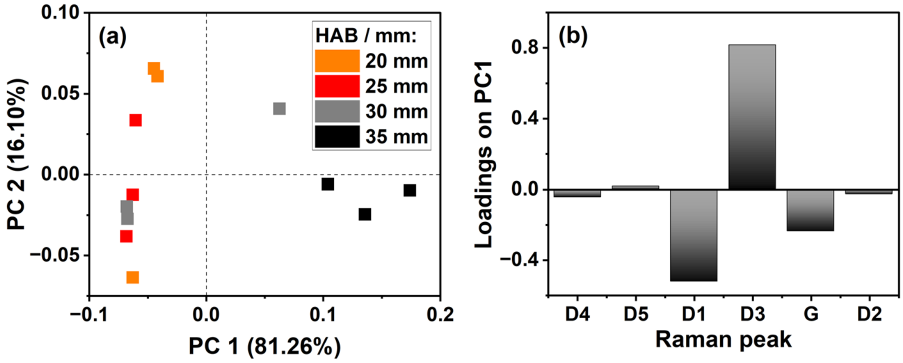


Fig. 3 Results of the PCA performed on the normalized peak areas (A_n) extracted from the Raman spectra at λ_{ex} = 325 nm. The figure shows the (a) PC2 against PC1 score plot and the (b) loading plot of PC1.

in our previous reports.[18] Peaks D1 and G have lower loadings than D3 and therefore contribute less to the phenomenon represented by PC1. In contrast, peaks D4, D5, and D2 have loadings close to zero, and therefore contribute marginally to the phenomenon represented by PC1.[33,37]

3.3. On the chemical composition

The discussion of the common patterns of signals obtained by the *ex situ* mass spectrometry analysis of samples extracted from flames is not the object of this work, which focuses instead on a chemometric-based approach to extract hidden information from the database. The mass spectra and the list of identified peaks are shown in the SI. Briefly, the mass spectra are rich of signals in *m*/*z* 150–800 that are identified by mass defect analysis mainly as $C_mH_n^+$ ions. The signals recur in groups having the same *m* and variable *n*. The most intense signals within a group characterized by the same *m* depend on the parity of *m*. If *m* is even, the strongest peaks exhibit even nominal mass. If *m* is odd, the strongest peaks exhibit odd nominal mass. This pattern is well-known and consistent with the identification of the signals as PAH.[39]

The information extracted from the mass spectra includes *m*/*z* and normalized ion counts of 393 peaks. The results of the PCA on the ensemble of the samples are shown in Fig. 4. The interpretation provided here is based on the analysis of PC1 and PC2 (62.02 and 27.58% of the total variance explained, respectively), while the lower-variance principal components are not discussed. Fig. 4(a) shows the PC2 against PC1 score plot, in which the datapoint represents the mass spectra. Three different data clusters can be identified: 20 and 25 mm HAB (orange-red) have high negative scores on PC1: they form a compact cluster, indicating that the PCA cannot find statistically significant differences among them. 30 mm HAB (gray) have high negative scores on PC2. Finally, 35 mm HAB (black) have high positive scores on PC1. Samples in different data clusters have sufficiently distinct chemical compositions to allow a classification along PC1 or PC2, and the corresponding loadings identify the variables responsible for the data clustering observed in the score plot.

In this database, PC1 clearly describes an HAB-dependent phenomenon: in the score plot in Fig. 4(a), the HAB monotonically increases from the left to the right. The corresponding loading plot in Fig. 4(b) shows that the responsible variables are high-*m*/*z* ions (negative loadings) that are more prominent at 20 and 25 mm HAB, against low-*m*/*z* ions (positive loadings) that are more prominent at 30 and 35 mm HAB. The transition from low- to high-*m*/*z* ions occurs at the sign inversion of PC1, around *m*/*z* 250.

By contrast, the phenomenon described by PC2 does not appear to be HAB-dependent. The scores on PC2 are approximately constant at low positive values for all samples except 30 mm HAB, for which PC2 assumes high negative values. As shown in Fig. 4(c), the distribution of positive loadings spans almost the entire investigated *m*/*z* range, and includes ions identified as $C_mH_n^+$ and other unidentified ions. The latter are defined by a lower mass defect than the identified PAHs, and hypotheses on their elemental composition include oxygen-containing hydrocarbons or hydrogen-poor carbon clusters:[40] the available resolving power is insufficient for certain identification. Conversely, the distribution of negative loadings is remarkable for several reasons. First, this distribution is bimodal as indicated by the arrows in Fig. 4(c). Then, almost all ions having negative loadings on PC2 are identified by mass defect analysis as $C_mH_n^+$, and their molecular formulae are consistent with PAHs or their derivatives. Finally, mode 2 appears at about twice the *m*/*z* of mode 1: mode 1 spans *m*/*z* 140–230 and includes C11-C18 hydrocarbons, while mode 2 spans *m*/*z* 340–450 and includes C27-C36 hydrocarbons.

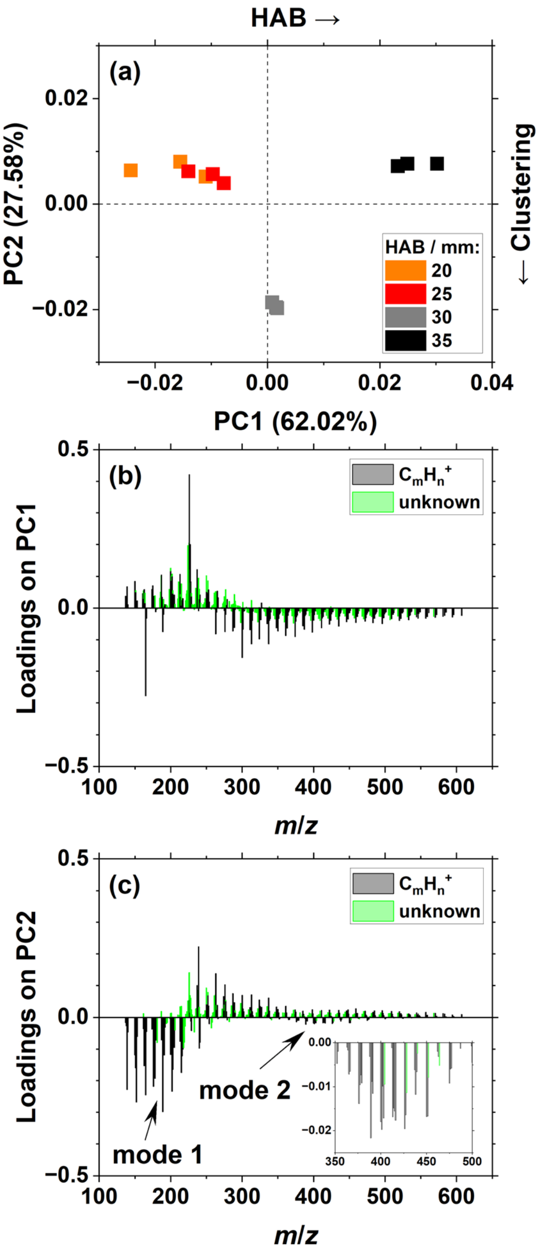


Fig. 4 Results of the PCA on the ensemble of the mass spectra. (a) PC2 *vs.* PC1 score plot. (b) PC1 and (c) PC2 loading plots. The black arrows point to the two modes representing correlated groups of peaks.

3.4. Discriminating carbon hybridization

The sp^3 to sp^2 ratio of the samples is estimated by XPS using either the so-called C KLL Auger peak *D*-parameter or the C 1s photopeak fitting. The *D*-parameter is calculated from the separation in energy between the maximum and the minimum in the derivative Auger C KLL spectra,[41,42] as shown in Fig. 5. As the signal-to-noise ratio in the first derivative C KLL spectra of carbon materials is low, a Savitzky–Golay algorithm, 3rd order polynomial with 20 energy window points smoothing is applied to determine the *D*-parameter. The sp^2 percentage is then estimated using a linear interpolation between reported values for pure sp^3 and pure sp^2 carbon, found at 13.2 eV and 23.1 eV respectively.[41] The C 1s spectra in Fig. 6 can be decomposed into six contributions with the lowest binding energy contribution corresponding to the graphitic C sp^2. C sp^3 and oxygen-containing groups like C sp^3 C–OH, C sp^2 C=O and O–C=O appear at higher binding energy and are separated from graphitic C sp^2 by 0.5 eV, 2 eV, 3.5 eV and 4.5 eV, respectively. An additional satellite peak separated from graphitic C sp^2 by 6.4 eV is found that corresponds to a $\pi-\pi^*$ transition found in graphitic carbon. The full width at half maximum (FWHM) is taken for the whole C 1s photopeak whereas the estimation of sp^2 content is calculated by summing all the sp^2 component areas (graphitic C, C=O and O–C=O peaks).

The calculated sp^2 percentages and all the parameters used for the calculations with using both methods are shown in Table 1. The calculations based on C KLL Auger peak and on C 1s photopeak show a monotonically increasing percentage of C sp^2 with the sampling HAB that is consistent with the progressive maturation of the samples. To explain the discrepancy between the numerical values, it is interesting to note that the two methods have different analysis depth, ~3.4 nm for C KLL and ~10 nm for C 1s. This discrepancy, which is consistent across all HABs, hints to a rapid changing chemical composition with the analysis depth and suggests that C sp^2 is located a few nm below the surface of the samples. The most likely explanation is the surface oxidation and consequent formation of C–OH bonds that occurs at the outermost layers of all samples exposed to air.

3.5. Evidence of emerging reactivity

The bimodal distribution of loadings on PC2 highlights a series of correlated *m*/*z* at 30 mm HAB, and aligning with transitions seen in LII measurements of this flame[13] marking the beginning of the soot inception region on the flame axis. Following the approach proposed in ref. 40, the hypothesis is made that mode 1 represents small reactive hydrocarbons, while mode 2 the resulting adducts. To verify this hypothesis, correspondences between pairs of ions in mode 1 ($C_{m1}H_{n1}^+$ and $C_{m2}H_{n2}^+$) and individual ions in mode 2 ($C_{m1+m2}H_{n1+n2-k}^+$) are therefore searched as follows. The sums of the *m*/*z* of each pair of ions in mode 1 are calculated and compared to individual ions in mode 2 (only the ions having the highest loadings of each mode are considered). If a match is found within a variable number of H atoms *k*, the value is recorded (the full table of correspondences is shown in the SI).

The number of correspondences against *k* is shown in Fig. 7. Several matches are found for $k = [0,8]$ with the highest frequency for $k = [2,4]$. On the one hand, the detected ions are all assigned to PAHs and their derivatives and therefore matches within an arbitrary number of H atoms are expected to occur as a result of the increasing molecular size. On the other hand, these matches exist between two specific sets of correlated ions, and they only occur for even values of *k* despite many identified ions having odd *m* and *n*. C15–C18 ions have the highest probability of generating a match. Furthermore, *k* is highest for C16–C17, then decreases for heavier species: if the matches were only due to the increasing molecular size, *k* would keep increasing with *m*/*z* as the disparity between the number of H and C atoms increases with the size of the PAHs.

Linking the reactivity of the (unknown) species in the flame to the species detected in the *ex situ* samples is not trivial. It is interesting to notice however that a hypothetically increasing rate of cross-linking reactions of low-*m*/*z* PAHs, possibly because of the flame temperature reaching a critical threshold, provides a consistent explanation for all the experimental evidence. First, the cross-linking of low-*m*/*z* PAHs would explain the two sets of correlated *m*/*z*, *i.e.* the bimodal distribution of loadings emerging locally at 30 mm HAB. Second, an increasing rate of formation of C–C bonds in the flame between 25 and 30 mm HAB would explain the maximum of *k* also emerging locally at 30 mm HAB, as the elimination of two H atoms can be interpreted as evidence of the formation of one C-C covalent bond.

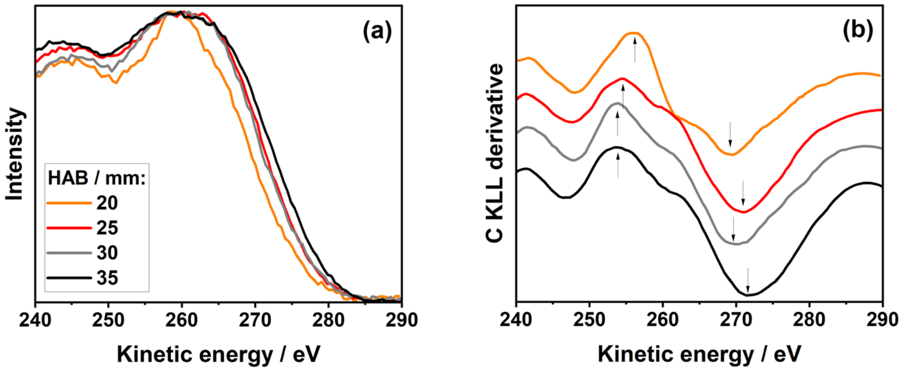


Fig. 5 (a) C KLL Auger spectra and (b) C KLL first derivative recorded from samples collected at 20, 25, 30, and 35 mm HAB.

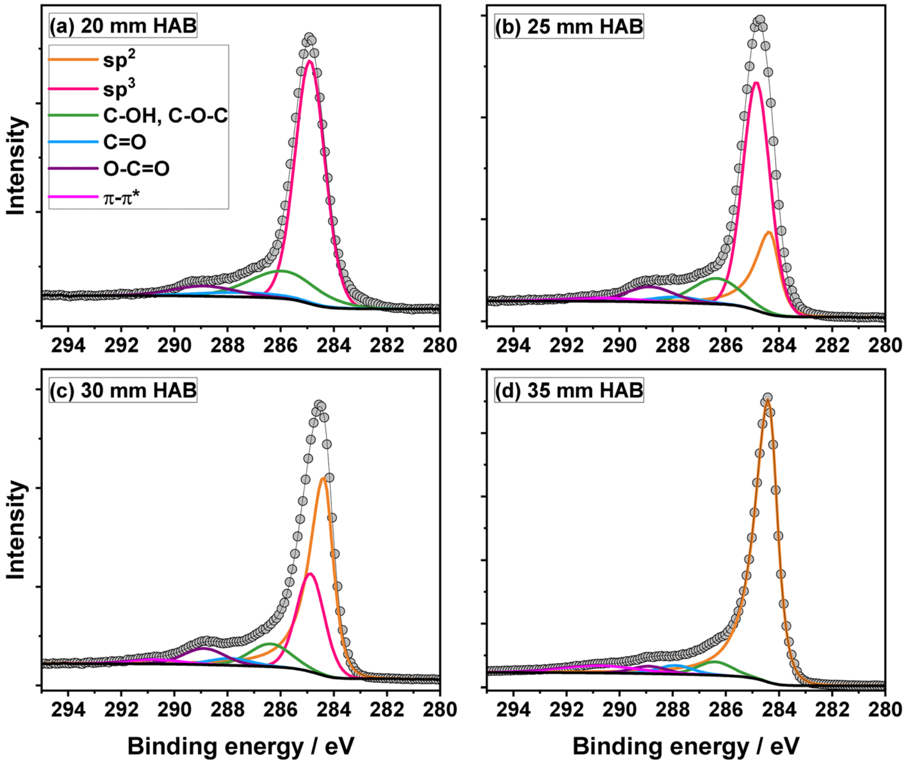


Fig. 6 XPS C 1s core level spectra of samples collected at (a) 20, (b) 25, (c) 30, and (d) 35 mm HAB.

Fig. 8 shows the atomic fraction of hydrogen [H] against the number of C atoms of the identified ions at 30 mm HAB. The two panels compare the ion count of all identified ions (a) and their variability across the CNPs inception zone represented by the loadings on PC2 (b), and clearly shows that only a subset of the detected *m*/*z* are involved in the reactivity the analysis of the mass spectra called attention to. The assignment of molecular structures to the detected ions remains a challenging task. The majority of the ions identified as $C_mH_n^+$ are found dispersed around the limit of *peri*-condensed PAHs (red dashes line in the figure). Structures consistent with molecular formulae above this limit include acenes, cata-condensed PAHs, bridged PAHs or PAHs substituted by short (possibly aliphatic) side chains. Structures consistent with molecular formulae below this limit include PAHs containing 5-member rings or substituted with highly unsaturated side chains ($-C_2H$ for instance).

It is noteworthy that ions with odd *m* values tend to have both high ion counts and variability. For instance, many cases are found for which the ion count of M + 1 is higher than M, which cannot obviously be explained by the isotopic distribution of C. Instead, this observation suggests the occurrence to some degree of post-ionization dissociation reactions like H-elimination from closed shell structures (for instance, non-aromatic CH_2 groups bridging a bay site).

Recent hypotheses on soot inception *via* E-bridges[43] or multicenter bonds[44] postulate small PAHs and their derivatives as key reactive species for dimerization and cross-linking reactions. The predictions of these hypotheses are in good agreement with the analysis proposed in this work. In the

Table 1 C KLL Auger peak *D*-parameter and XPS C 1s total photopeak FWHM associated with the calculated sp^2 percentage

	C KLL Auger peak		C 1s photopeak	
HAB/mm	*D*-parameter/eV	sp^2%	C 1s FWHM/eV	sp^2%[a]
20	13.5	3	1.40	10.3
25	17.0	38	1.35	33.1
30	17.0	38	1.36	65.6
35	18.5	54	1.01	95.1

[a] sp^2 content determined using the C 1s fitting from Fig. 6.

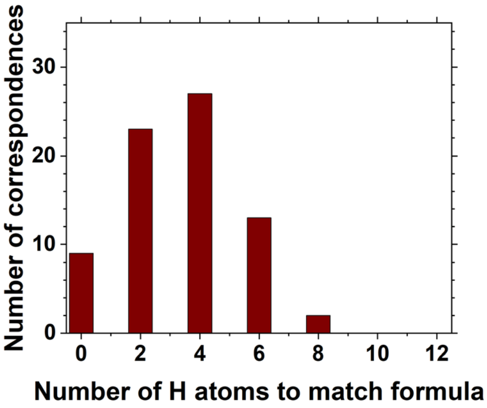


Fig. 7 Number of matches between pairs of ions of mode 1 ($C_{m1}H_{n1}^+$ and $C_{m2}H_{n2}^+$) and individual ions of mode 2 ($C_{m1+m2}H_{n1+n2-k}^+$) against *k* (within 15 ppm from the expected position of the ion in mode 2).

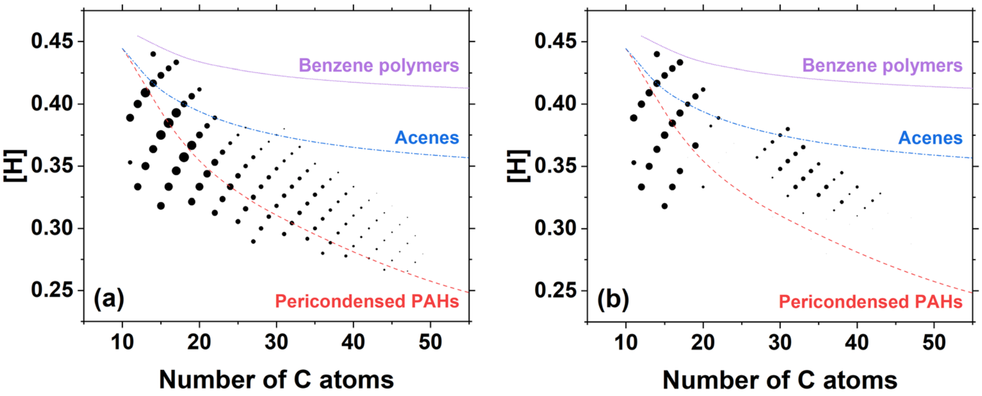


Fig. 8 [H] against number of C atoms of all identified ions at 30 mm HAB. (a) The size of the datapoints is proportional to the logarithm of the ion count. (b) The size of the datapoints is proportional to the logarithm of the (absolute value) of the loadings on PC2.

investigated flame, the identified reactive species are C12–C20 hydrocarbons with [H] = 0.30–0.45, while their adducts are C27–C40 hydrocarbons with [H] = 0.32–0.38.

3.6. Correlating structure and composition

As shown in Fig. 9, the identified descriptors of the structure and of the chemical composition of the samples are not independent. The scores on PC1 (Raman spectroscopy at 325 nm) are positively correlated to the scores on PC1 (ToF-SIMS), thus the increase of intensity of the Raman peak D3 is correlated to the disappearance of high-*m*/*z* hydrocarbons. As mentioned above, this change in the behavior of both Raman spectra and mass spectra becomes evident between 25 and 30 mm HAB, and it is consistent with the beginning of the soot inception region as revealed by *in situ* LII measurements.[13] This evidence is further corroborated by the XPS analyses (Fig. 5 and 6) that show the progressive transition from C sp^3 to C sp^2 against HAB, where C sp^2 is detected as early as 25 mm HAB and becomes dominant over C sp^3 at 30 mm HAB. Although peak D3 has been suggested to be related to the formation of amorphous carbon or finite-size graphitic crystals, the vibration mode(s) directly responsible have not yet been identified. Therefore, understanding its detailed role on the inception of CNPs remains a challenging task.

No correlation is found between the scores on PC1 (Raman) and the scores on PC2 (ToF-SIMS), thus the emergent reactivity highlighted by the analysis of the mass spectra might not be mirrored by changes in the spectral region where the D and G Raman bands are located. A more detailed exploration of this aspect will be the object of a future work.

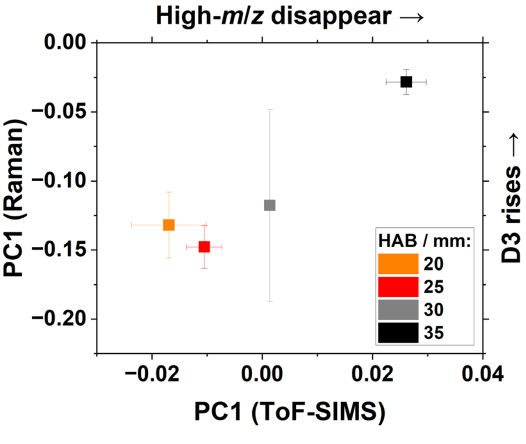


Fig. 9 Scores on PC1 from the analysis of the Raman spectra against scores on PC1 from the analysis of the mass spectra. Since it is not possible to compare analyses performed at the same location at the surface of the samples, the averages over all the available datapoints are used instead for the comparison (the error bars represent the 95% confidence interval).

4. Conclusions

This multi-diagnostic investigation of incipient CNP formation and their precursors in a nitrogen-diluted ethylene laminar diffusion flame has provided new molecular level evidence for the fundamental physical chemistry governing gas-to-particle transitions. By combining structural, chemical and morphological characterization with statistical analysis, we have identified signatures that illuminate the mechanistic pathways of CNP inception.

Our results provide compelling experimental evidence for a combined physical and chemical inception mechanism that occurs close to 30 mm HAB on this flame system. The bimodal distribution observed in ToF-SIMS data characterized by loss of 2–4 H atoms is consistent with the formation of covalent C–C bonds between initially physically associated clusters. This even-numbered hydrogen elimination pattern may indicate a concerted mechanism involving the stabilization of weakly bound PAH dimers through radical-mediated bonding reactions consistently with hypotheses recently proposed in the literature. The XPS results provide independent confirmation of this transformation revealing an increase in C sp^2 content with HAB and highlighting depth-dependent compositional changes consistent with surface oxidation, which strengthens the evidence for progressive graphitization during inception.

Raman spectra reveal a fundamental transformation in molecular structures during CNP maturation. The emergence of the D3 peak at 35 mm HAB indicates a transition from the extended aromatic domains towards more defective, finite-size graphitic structures. This structural evolution suggests that CNP inception involves the formation of curved or multi-layered structures rather than simple planar PAH aggregation.

The disappearance of high-*m*/*z* species from mass spectra challenges models that assumes preservation of large, flat PAHs during inception and supports mechanisms that involve three-dimension cluster formation. The spatial localization of important "transitions" at 30 mm HAB, coincides with the onset of particle formation observed in previous LII studies,[13] and establishes this region as a critical "inception zone" where thermodynamic and kinetic factors converge to enable particle inception.

Neither Raman spectroscopy, mass spectrometry, nor XPS alone can provide definitive structural characterization of CNPs, which underscores the need for the integration of both surface-sensitive and structure-sensitive methods with appropriate statistical frameworks. On-going studies of this flame system will incorporate multi-wavelength Raman spectroscopy to further explore the correlation of excitation wavelength and CNP electronic structure as revealed by key spectral features, refining our understanding of the soot formation process.

Conflicts of 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.

Data availability

Data for this article, including ToF-SIMS mass spectra, Raman spectra and SEM images, are available at the Open Science Framework at **https://doi.org/10.17605/OSF.IO/FJA2D**.

Supplementary information (SI) is available. See DOI: **https://doi.org/10.1039/d5cp02594j**.

Acknowledgements

The authors would like to thank Ing. Bénédicte Calimet for her invaluable support in the preparation of the measurement campaign. The authors would also like to acknowledge the contribution of the CERLA platform for the materials and equipment. Financial support for this work was provided by the Labex CaPPA under contract ANR-11-LABX-0005-01. Partial support for some of the laboratory equipment used in this collaboration and salary support for EEMc and JHM was provided by the US National Science Foundation through grants CBET-0828950, CBET-1142284, CBET-1706757, and CBET-2412623.

References

1 P. R. Buseck, K. Adachi, A. Gelencsér, É. Tompa and M. Pósfai, Are black carbon and soot the same? Aerosols/Field Measurements/Troposphere/Chemistry (chemical composition and reactions), *Atmos. Chem. Phys.*, 2012, **12**(9), 24821–24846.
2 Ü. Ö. Köylü, G. M. Faeth, T. L. Farias and M. G. Carvalho, Fractal and projected structure properties of soot aggregates, *Combust. Flame*, 1995, **100**, 621–633.
3 T. S. Totton, A. J. Misquitta and M. Kraft, A quantitative study of the clustering of polycyclic aromatic hydrocarbons at high temperatures, *Phys. Chem. Chem. Phys.*, 2012, **14**, 4081–4094.
4 C. A. Schuetz and M. Frenklach, Nucleation of soot: Molecular dynamics simulations of pyrene dimerization, *Proc. Combust. Inst.*, 2002, **29**, 2307–2314.
5 K.-H. Homann, Fullerenes and Soot Formation— New Pathways to Large Particles in Flames, *Angew. Chem., Int. Ed.*, 1998, **37**, 2434–2451.
6 A. Violi, A. F. Sarofim and G. A. Voth, Kinetic Monte Carlo–Molecular Dynamics Approach to Model Soot Inception, *Combust. Sci. Technol.*, 2004, **176**, 991–1005.
7 A. Violi, A. Kubota, T. N. Truong, W. J. Pitz, C. K. Westbrook and A. F. Sarofim, A fully integrated kinetic monte carlo/molecular dynamics approach for the simulation of soot precursor growth, *Proc. Combust. Inst.*, 2002, **29**, 2343–2349.
8 H. Jin, J. Guo, T. Li, Z. Zhou, H. G. Im and A. Farooq, Experimental and numerical study of polycyclic aromatic hydrocarbon formation in ethylene laminar co-flow diffusion flames, *Fuel*, 2021, **289**, 119931.
9 H. Jin, A. Cuoci, A. Frassoldati, T. Faravelli, Y. Wang, Y. Li and F. Qi, Experimental and kinetic modeling study of PAH formation in methane coflow diffusion flames doped with n-butanol, *Combust. Flame*, 2014, **161**, 657–670.
10 H. Jin, A. Frassoldati, Y. Wang, X. Zhang, M. Zeng, Y. Li, F. Qi, A. Cuoci and T. Faravelli, Kinetic modeling study of benzene and PAH formation in laminar methane flames, *Combust. Flame*, 2015, **162**, 1692–1711.
11 H. Jin, W. Yuan, Y. Wang, Y. Li, F. Qi, A. Cuoci, A. Frassoldati and T. Faravelli, Experimental and kinetic modeling study of laminar coflow diffusion methane flames doped with 2-butanol, *Proc. Combust. Inst.*, 2015, **35**, 863–871.
12 H. Jin, G. Wang, Y. Wang, X. Zhang, Y. Li, Z. Zhou, J. Yang and F. Qi, Experimental and kinetic modeling study of laminar coflow diffusion methane flames doped with iso-butanol, *Proc. Combust. Inst.*, 2017, **36**, 1259–1267.
13 J. D. Herdman, B. C. Connelly, M. D. Smooke, M. B. Long and J. H. Miller, A comparison of Raman signatures and laser-induced incandescence with direct numerical simulation of soot growth in non-premixed ethylene/air flames, *Carbon*, 2011, **49**, 5298–5311.
14 E. M. Adkins, J. A. Giaccai and J. H. Miller, Computed electronic structure of polynuclear aromatic hydrocarbon agglomerates, *Proc. Combust. Inst.*, 2017, **36**, 957–964.
15 J. Elias, L. Labarrière, A. Faccinetto, A. Moncomble, J.-P. Cornard and X. Mercier, Advanced characterization of soot precursors via excitation emission matrices fluorescence spectroscopy and molecular modeling, *Carbon*, 2024, **228**, 119355.
16 M. L. Botero, E. M. Adkins, S. González-Calera, H. Miller and M. Kraft, PAH structure analysis of soot in a non-premixed flame

using high-resolution transmission electron microscopy and optical band gap analysis, *Combust. Flame*, 2016, **164**, 250–258.

17 R. S. Jacobson, A. R. Korte, A. Vertes and J. H. Miller, The Molecular Composition of Soot, *Angew. Chem., Int. Ed.*, 2020, **59**, 4484–4490.

18 J. Elias, A. Faccinetto, C. Irimiea, N. Nuns, C. Pirim, C. Focsa, H. Vezin and X. Mercier, On the chemical composition and structure of incipient soot in a laminar diffusion flame, *Fuel*, 2024, **373**, 132056.

19 E. M. Adkins and J. H. Miller, Extinction measurements for optical band gap determination of soot in a series of nitrogen-diluted ethylene/air non-premixed flames, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2686–2695.

20 B. C. Connelly, B. A. V. Bennett, M. D. Smooke and M. B. Long, A paradigm shift in the interaction of experiments and computations in combustion research, *Proc. Combust. Inst.*, 2009, **32**, 879–886.

21 N. J. Kempema and M. B. Long, Combined optical and TEM investigations for a detailed characterization of soot aggregate properties in a laminar coflow diffusion flame, *Combust. Flame*, 2016, **164**, 373–385.

22 N. J. Kempema, B. Ma and M. B. Long, Investigation of in-flame soot optical properties in laminar coflow diffusion flames using thermophoretic particle sampling and spectral light extinction, *Appl. Phys. B*, 2016, **122**, 232.

23 B. Ma and M. B. Long, Combined soot optical characterization using 2-D multi-angle light scattering and spectrally resolved line-of-sight attenuation and its implication on soot color-ratio pyrometry, *Appl. Phys. B*, 2014, **117**, 287–303.

24 D. Bartos, M. Sirignano, M. J. Dunn, A. D'Anna and A. R. Masri, Soot inception in laminar coflow diffusion flames, *Combust. Flame*, 2019, **205**, 180–192.

25 M. Smooke, M. Long, B. Connelly, M. Colket and R. Hall, Soot formation in laminar diffusion flames, *Combust. Flame*, 2005, **143**, 613–628.

26 C. Irimiea, A. Faccinetto, Y. Carpentier, I. Ortega, N. Nuns, E. Therssen, P. Desgroux and C. Focsa, A comprehensive protocol for chemical analysis of flame combustion emissions by secondary ion mass spectrometry, *Rapid Commun. Mass Spectrom.*, 2018, **32**, 1015–1025.

27 M. Daoudi, N. Nuns, P. Schiffmann, A. Frobert, B. Hanoune, P. Desgroux and A. Faccinetto, A mass defect–based approach for the automatic construction of peak lists for databases of mass spectra with limited resolution: Application to time-of-flight secondary ion mass spectrometry data, *Rapid Commun. Mass Spectrom.*, 2024, **38**, e9777.

28 N. Fairley, V. Fernandez, M. Richard-Plouet, C. Guillot-Deudon, J. Walton, E. Smith, D. Flahaut, M. Greiner, M. Biesinger, S. Tougaard, D. Morgan and J. Baltrusaitis, Systematic and collaborative approach to problem solving using X-ray photoelectron spectroscopy, *Appl. Surf. Sci.*, 2021, **5**, 100112.

29 C. Irimiea, A. Faccinetto, X. Mercier, I.-K. Ortega, N. Nuns, E. Therssen, P. Desgroux and C. Focsa, Unveiling trends in soot nucleation and growth: When secondary ion mass spectrometry meets statistical analysis, *Carbon*, 2019, **144**, 815–830.

30 H. Chu, W. Han, W. Cao, C. Tao, M. Raza and L. Chen, Experimental investigation of soot morphology and primary particle size along axial and radial direction of an ethylene diffusion flame via electron microscopy, *J. Energy Inst.*, 2019, **92**, 1294–1302.

31 A. Sadezky, H. Muckenhuber, H. Grothe, R. Niessner and U. Pöschl, Raman microspectroscopy of soot and related carbonaceous materials: Spectral analysis and structural information, *Carbon*, 2005, **43**, 1731–1742.

32 C. Russo and A. Ciajolo, Effect of the flame environment on soot nanostructure inferred by Raman spectroscopy at different excitation wavelengths, *Combust. Flame*, 2015, **162**, 2431–2441.

33 B. Dippel, H. Jander and J. Heintzenberg, NIR FT Raman spectroscopic study of flame soot, *Phys. Chem. Chem. Phys.*, 1999, **1**, 4707–4712.

34 T. Jawhari, A. Roid and J. Casado, Raman spectroscopic characterization of some commercially available carbon black materials, *Carbon*, 1995, **33**, 1561–1565.

35 R. J. Nemanich and S. A. Solin, First- and second-order Raman scattering from finite-size crystals of graphite, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1979, **20**, 392–401.

36 P. Parent, C. Laffon, I. Marhaba, D. Ferry, T. Z. Regier, I.-K. Ortega, B. Chazallon, Y. Carpentier and C. Focsa, Nanoscale characterization of aircraft soot: a high-resolution transmission electron microscopy, Raman spectroscopy, X-ray photoelectron and near-edge X-ray absorption spectroscopy study, *Carbon*, 2016, **101**, 86–100.

37 A. Cuesta, P. Dhamelincourt, J. Laureyns, A. Martínez-Alonso and J. M. D. Tascón, Raman microprobe studies on carbon materials, *Carbon*, 1994, **32**, 1523–1532.

38 O. Beyssac, B. Goffé, J.-P. Petitet, E. Froigneux, M. Moreau and J.-N. Rouzaud, On the characterization of disordered and heterogeneous carbonaceous materials by Raman spectroscopy, *Spectrochim. Acta A*, 2003, **59**, 2267–2276.

39 J. W. Martin, M. Salamanca and M. Kraft, Soot inception: Carbonaceous nanoparticle formation in flames, *Prog. Energy Combust. Sci.*, 2022, **88**, 100956.

40 A. Faccinetto, C. Irimiea, P. Minutolo, M. Commodo, A. D'Anna, N. Nuns, Y. Carpentier, C. Pirim, P. Desgroux, C. Focsa and X. Mercier, Evidence on the formation of dimers of polycyclic aromatic hydrocarbons in a laminar diffusion flame, *Commun. Chem.*, 2020, **3**, 112.

41 B. Lesiak, L. Kövér, J. Tóth, J. Zemek, P. Jiricek, A. Kromka and N. Rangam, C sp2/sp3 hybridisations in carbon nanomaterials – XPS and (X)AES study, *Appl. Surf. Sci.*, 2018, **452**, 223–231.

42 J. C. Lascovich and S. Scaglione, Comparison among XAES, PELS and XPS techniques for evaluation of sp2 percentage in a-C:H, *Appl. Surf. Sci.*, 1994, **78**, 17–23.

43 M. Frenklach and A. M. Mebel, On the Mechanism of Soot Nucleation. IV. Molecular Growth of the Flattened E-Bridge, *J. Phys. Chem. A*, 2022, **126**, 9259–9267.

44 J. W. Martin, D. Hou, A. Menon, L. Pascazio, J. Akroyd, X. You and M. Kraft, Reactivity of Polycyclic Aromatic Hydrocarbon Soot Precursors: Implications of Localized π-Radicals on Rim-Based Pentagonal Rings, *J. Phys. Chem. C*, 2019, **123**, 26673–26682.