Supplementary material

Figure S1 illustrates the difference between PAH binding energy as a function of the

dimer reduced mass proposed by Miller [16] and Herdman and Miller [49].



Figure S1. Comparison of two intermolecular potentials used by [16, 49].

The following set of differential equations are solved to obtain pyrene [A4] and pyrene dimer [A4-A4] concentration as a function of time at different temperatures using the forward and reverse rate constants calculated with equations 1 and 3 of the manuscript.

$$\frac{d[A4]}{dt} = -2k_f[A4]^2 + 2k_r[A4 - A4]$$
(S1)

$$\frac{d[A4 - A4]}{dt} = k_f [A4]^2 - k_r [A4 - A4]$$
(S2)

 $[A4]_0 = [A4]_{@\ t = 0\ for\ 60\ K,\ or\ 235\ K}$ (S3)

$$[A4 - A4]_0 = 0$$

Figure S2 compares the predicted pyrene concentrations by RNC and IN models with the measured values [40]. Because in the RNC model all PAHs contribute to nucleation, pyrene concentrations are not affected by nucleation. However, the IN model only uses pyrene for nucleation and the predicted results are much lower than the RNC model and outside of the experimental error bars.



Figure S2. Comparison of measured pyrene concentrations with the predicted values when all PAHs contribute to soot nucleation (⁻⁻) and when only pyrene is considered for nucleation (⁻⁻). By only nucleating from a single species, its concentrations cannot be predicted well.

Figure S3 compares the rate constants for chemical bond formation reactions used in the proposed nucleation mechanism (reactions 6, 7 and 8 listed in table one) with an empirical rate constant for soot carbonization [44]. The empirical rate constant is based on the decrease in soot H/C ratio along the centerline of an ethylene coflow diffusion flame.



Figure S3. Comparison of an empirical rate constants for soot dehyrogenation [44, 60] with the first step of PAH radical-molecule reactions obtained from density functional theory [45: Table 2: reactions 1, 5 and 7] and reactions 6, 7 and 8.