Atomistic simulations of the aggregation of small aromatic molecules in homogenous and heterogenous mixtures

Michael Thomas, Irene Suarez-Martinez, Li-Juan Yu, Amir Karton, Graham Chandler, Marc Robinson, Isabelle Cherchneff, Dahbia Talbi, Dino Spagnoli

Atomistic simulations of the aggregation of small aromatic molecules in homogenous and heterogenous mixtures

Michael Thomas [1], Irene Suarez Martinez [2], Li-Juan Yu [3,4], Amir Karton [3], Graham Chandler [3], Marc Robinson[2], Isabelle Cherchneff [5], Dahbia Talbi [6*], and Dino Spagnoli[3*]

[1] College of Engineering and Computer Science, Australian National University, Canberra, ACT 2601, Australia
[2] Department of Physics and Astronomy, Curtin University, Perth, WA, 6845, Australia
[3] School of Molecular Sciences, University of Western Australia, Perth, WA, 6009, Australia
[4]ARC Centre of Excellence for Electromaterials Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia
[5] Instituto de Fisica Fundamental, Consejo Superior de Investigaciones Cientificas (CSIC), Serrano 113 bis 28006, Madrid, Spain
[6] Laboratoire Univers et Particules Montpellier, UMR5299-CNRS-Université de Montpellier, France

*corresponding authors:
Dino Spagnoli: dino.spagnoli@uwa.edu.au,
Dahbia Talbi: dahbia.talbi@umontpellier.fr

Abstract

The relatively weak London dispersion forces are the only interactions that could cause aggregation between simple aromatic molecules. The use of molecular dynamics and high-level *ab initio* computer simulations has been used to describe the aggregation and interactions between molecular systems containing benzene, naphthalene and anthracene. Mixtures containing one type of molecule (homogenous) and more than one type of molecule (heterogenous) were considered. Our results indicate that as molecular weight increases so does the temperature at which aggregation will occur. In all simulations, the mechanism of aggregation is through small clusters coalescing into larger clusters. The structural analysis of the molecules within the clusters reveals that benzene will orient itself in T-shaped and parallel display configurations. Molecules of anthracene prefer to orient themselves in a similar manner to a bulk crystal with no T-shaped configuration observed. The aggregation of these aromatic molecules is discussed in the context of astrochemistry with particular reference to the dust formation region around stars.

Introduction

Weak intermolecular forces control the physical properties of aromatic molecules and play a crucial role in processes such as protein-ligand recognition,[1] the packing and assembly of aromatic molecules in crystals,[2] and soot formation.[3-5] These interactions between aromatic molecules are strong enough so clusters can form via intermolecular forces. [5] The early chemistry of soot formation is now generally thought to be controlled by these nonbonding interactions. Understanding the formation of soot by aromatic molecules has received a great deal of research mainly due to the environmental implications.[3]

Aggregation of aromatic molecules is also of special interest in the field of astrochemistry, in particular for the study of the chemical composition of the outer sphere (photosphere) of carbon-rich stars. Evidence has been accumulating for years on the similarity between the chemical pathways observed

in combustion processes on Earth, which form benzene, polycyclic aromatic hydrocarbons (PAHs) and soot, with the high temperature conditions in circumstellar regions. [6-8] Observation from radio telescopes have determined that an envelope of carbon dust surround red-giants. These particular stars emit more carbon and less oxygen than most other stars and are the main carbon dust providers to local galaxies. [9] Theoretical studies of the growth of PAHs [6, 7, 10, 11] is commonly explained through hydrogen abstraction and acetylene addition (i.e., HACA mechanism [7]). PAHs could form and grow in a distinctive temperature range of 900-1500K, but the yields were rather low to explain the ubiquity of dust in carbon stars. The mechanism of aggregation of PAHs has not yet been addressed, and it is key to understanding the production of carbon grains from stars.

Similarities exist between the chemical and physical conditions of acetylene-rich flames and those pertaining to the upper atmosphere of carbon stars where dust forms. Both environments are rich in hydrogen, carbon, acetylene and oxygen, and have similar temperature regimes. Experiments in the laboratory on carbon dust formation mimicking the low-density, warm environments of carbon star winds have been undertaken. [12, 13] Plasma discharge experiments have shown that the growth of unsaturated hydrocarbons and aromatics proceeded through chain growth and ring formation. [14] More recently, the production of carbon clusters in ultra-high-vacuum gas aggregation sources shows that carbon growth mainly proceeds through the condensation of pure carbon chains and rings, where aromatics represent a minor population of carbon clusters. [15] Theoretical studies related to soot formation shows that the aggregation of PAHs through London dispersion interactions can be efficient for a large range of temperature from 400 K to 1600K. [16]

To advance our understanding of PAH aggregation into small carbon clusters, we conducted this extensive and systematic investigation through molecular dynamics (MD) and ab-initio simulations. We considered both homogenous (one molecular species) and heterogeneous (multi-molecular species) aggregation between 100K and 500K of three small aromatic molecules: benzene, naphthalene and anthracene. The literature on aggregation of polyaromatic hydrocarbons via simulations is extensive. [16-19] Moreover, the mechanism and the molecules involved in soot nucleation is unknown. [4, 20] This study provides insight into the mechanism of aggregation of small aromatic molecules. We believe that this is the first time that a systematic investigation of heterogenous mixtures of small aromatic molecules has been performed. In-depth analysis arrangements of the molecules within the cluster is performed and compared to ab-initio calculations of dimers and trimers. Finally, our results are discussed within the astrophysical context of dust formation in evolved stars.

Methods

Molecular Dynamics Simulation Details

Seven systems were constructed for the MD simulations. The composition of each system is detailed in Table 1. The number of molecules in each system was chosen to maintain a constant atomic density using periodic boundary conditions in a cubic cell of length 120 Å. The size of the simulation cell and number of molecules were chosen because we wanted to simulate a system where aggregation would occur on a timescale accessible to MD. We note that the density of the simulations is far higher than densities found in space.

Table 1: Composition of the mixture for the seven-system studied using molecular dynamics.

System	n(Benzene)	n(Naphthalene)	n(Anthracene)	n(atoms)
Benzene	60	0	0	720
Naphthalene	0	40	0	720
Anthracene	0	0	30	720
Benzene and Naphthalene	30	20	0	720
Benzene and Anthracene	30	0	15	720
Naphthalene and Anthracene	0	20	15	720
Benzene, Naphthalene and Anthracene	20	13	10	714

All systems were simulated with the molecular dynamics package NAMD 2.12 [21]. All simulations were conducted using the NVT ensemble (constant number of atoms, temperature and volume) with the Langevin thermostat [22, 23] and a dampening parameter of 5 ps to maintain constant temperature. Particle Mesh Ewald [24] was used to calculate electrostatic interactions. A timestep of 1 fs was used in all cases.

Initially the OPLS-AA force field [25] was chosen based on previously reports by Takeuchi who calculated the potential energy of PAHs clusters in static simulations using this force field [26, 27] . However, our dynamic simulations showed that aggregation was only observed when benzene was present using the OPLS-AA force field. Simulations involving naphthalene and anthracene did not show aggregation even at temperatures below the condensation point. We repeated the calculations using the parameters from CGenFF within the CHARMM36 force field [28] to describe bonding and non-bonding interactions and aggregation occurs for all systems. All results that follows use the CHARMM force field.

All systems were modelled for five temperatures: 100, 200, 300, 400 and 500 K. Each temperature was modelled five times using five different initial configurations, which is a method used in other MD studies. [29] Initial configurations were produced from randomly positioned molecules heated at 1000 K for 100, 200, 300, 400 and 500 ps. This method ensures the final configurations are not biased by the initial configuration. All simulations were performed for 10 ns. All seven systems displayed aggregation at 200 K, therefore these simulations were run until a single cluster was formed at this temperature, typically 100-200 ns. Each simulation trajectory was analysed to determine cluster sizes by searching atoms from different molecules within 4 Å. This distance was chosen as literature has shown that the maximum distance between two aromatic rings in a parallel displaced orientation would be no more than 3.9 Å. [30, 31]

Ab initio Simulation Details

High-level double-hybrid density functional theory (DHDFT) is used to investigate the geometries of dimer and trimer complexes. All geometry optimization, frequency, and energy calculations were performed using the Gaussian 09 program suite. [32] In the first approach all the degrees of freedom have been fully optimized at the M06-L/6-31G(2df,p) level of theory [33, 34] in order to obtain the local minima structures on the potential energy surface (PES). We were able to apply this approach for the benzene-benzene, benzene-naphthalene, and benzene-anthracene dimers in the parallel displaced (PD) configuration. In all cases harmonic vibrational frequencies have been calculated to confirm each

stationary point as an equilibrium structure (i.e., all real frequencies). In the second approach, we use the optimized monomers and scan the vertical intermolecular distance. The DHDFT PWPB95-D3 functional in conjunction with the Def2-QZVPP basis set [35, 36] was used to carry out the PES scans. Empirical D3 dispersion corrections [37, 38] are included using the Becke-Johnson [39] damping potential as recommended by Grimme *et. al.* [40] (denoted by the suffix D3). The Resolution of Identity (RI) approximation was used to speed up the calculations. This level of theory has been shown to give very accurate interaction energies for weakly bonded systems.[36] This approach is used to obtain the structures of all the dimers and trimers considered in the present work. It is worthwhile comparing the optimal intramolecular distances obtained via scanning only one coordinate (i.e., the R and R1 coordinates for T– and PD–structures, respectively) with those obtained via the fully optimizations. These results are gathered in Table S3 of the Supporting Information. The ring-to-ring distance R1 for fully optimized PD benzene dimer is 3.450 Å, whereas scanning this distance results in R1 = 3.620 Å. The difference between these two values amounts to 0.17 Å, that is 4.9% of the intramolecular bond distance. Similarly, the partial optimization leads to a bond length that is longer by 0.19 Å for the benzene-naphthalene dimer (i.e., 5.4% of the bond distance). For the benzene-anthracene dimer, the partial and full optimizations result in a similar intramolecular bond separation (the difference between the bonds being only 0.04 Å).

Results and Discussion

Clustering as function of temperature and mixture composition

The seven molecular systems were simulated at five different temperatures for 10 ns. The purpose of these simulations was to understand at which temperatures aggregation occurs within a practical MD timeframe for the different molecular systems. The maximum cluster size as a function of size is shown in the left-hand side of figures 1a-g. At 100 and 200 K, all molecular systems show evidences of cluster formation. For simulations involving only benzene there is very little aggregation at temperatures at 300 K and above. In the supplementary material table S1 shows all average number of clusters, average maximum cluster size and average cluster size for all replicates at each temperature. For simulations involving naphthalene and anthracene, clusters begin to form and are stable at higher temperatures of 300 K and 400 K. There is very little evidence of any clusters of more than two molecules forming for any molecular system at 500 K. For the same temperature, we find larger clusters in the heterogeneous mixtures compared to homogenous mixtures. For example, in the molecular system involving only benzene there is very little evidence of clusters forming at 300 K. However, in the heterogeneous molecular systems involving benzene and/or the polyaromatic molecules there is evidence of clusters forming at 300 K. Since the dominant interaction between these non-polar molecules is London dispersion forces, with its strength proportional to molecular weight, it is unsurprising that this would be the case.

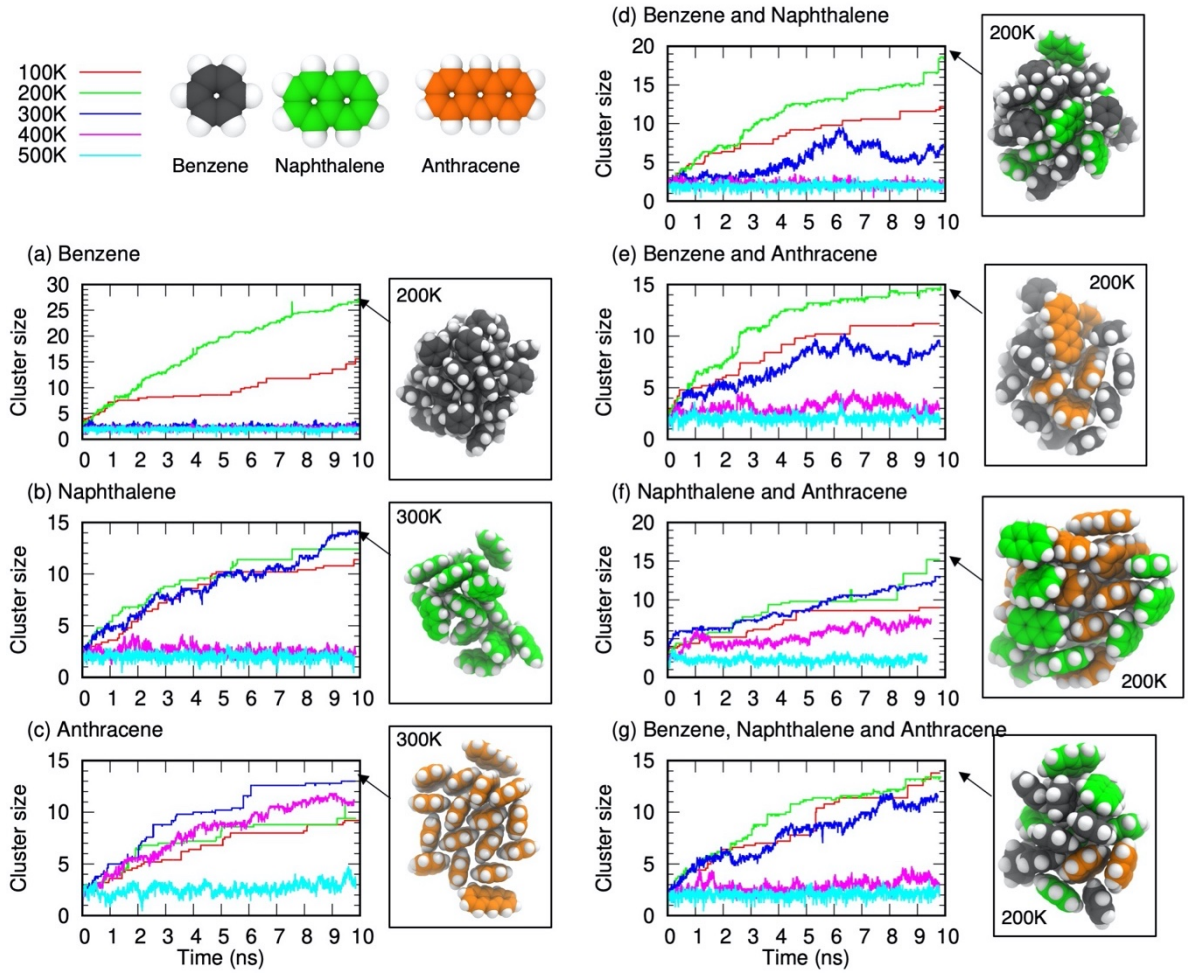

Figure 1: Maximum cluster size averaged over the five replicas at 100 K (red), 200 K (green), 300 K (blue), 400 K (pink) and 500 K (cyan) for a) 60 benzene molecules, b) 40 naphthalene molecules, c) 30 anthracene molecules, d) 30 benzene and 20 naphthalene molecules, e) 30 benzene and 15 anthracene molecules, f) 20 naphthalene and 15 anthracene molecules, and g) 20 benzene, 13 naphthalene and 10 anthracene molecules. An example of the final configuration of the largest cluster formed for each molecular system is shown on the right of each plot.

At 10 ns, the benzene homogenous mixture shows nearly half of the initial molecules aggregated into a single cluster at 200 K. There is very little evidence of any clusters forming above the 2-molecule threshold at 300, 400 or 500 K for the pure benzene molecular system. For the homogenous system of naphthalene, the bigger clusters are formed with increasing temperature from 100 K to 300 K, and no evidence of any cluster forming above 400 K (figure 1b). In the case of anthracene, clusters are observed at all temperatures except at 500 K (figure 1c). The calculation of the enthalpy change of sublimation from experimental studies for benzene ranges between temperatures of 193 [41]-279 K, [42] naphthalene from 274 -353 K,[43] and anthracene from 337 [44] - 488 K.[45] Our results are in agreement with those values, and show that at temperatures above the sublimation point there is no aggregation. At temperatures below the point of sublimation, for each homogenous molecular system, aggregation occurs.

Intermolecular forces are stronger in heterogenous mixtures compared to homogeneous mixtures. Heterogenous clusters are observed at temperatures in which homogeneous mixtures showed no aggregation. For example, in the simulation involving benzene and naphthene mixture, clusters can be seen at temperatures up to 300 K (Figure 1d), whereas in pure benzene mixtures clusters only formed for temperatures under 200 K (Figure 1a). Similarly, clusters are formed in benzene and anthracene mixtures at temperatures up to 300 K (Figure 1e). In the mixture containing the two largest aromatic hydrocarbons (naphthalene and anthracene, Figure 1f), the biggest clusters are observed at 200 and 300

K, while at low temperatures (100 K) and high temperature (above 400 K), aggregation is none or limited. Interestingly, the mixture of the three aromatics do not form cluster above 400 K (Figure 1g), however pure mixtures of anthracene clearly formed clusters at that temperature.

The final configuration of the largest clusters formed at the end of 10 ns simulation is shown in the right-hand side of Figure 1 a-g. In heterogenous mixtures, clusters rarely form of one single molecular type. Moreover, larger aromatic molecules are often found in the centre of the clusters surrounded by smaller molecules. One particular exception is the mixture of the three aromatic molecules, in which at the 10 ns stage no pattern appears evident and cluster are often randomly arranged.

Molecular dynamics simulations at 200 K for > 100 ns

To investigate the final configuration and geometries of homogeneous and heterogeneous mixtures over longer time scale, the seven systems were simulated at 200 K. The temperature of 200 K was chosen because all molecular systems showed clustering at this temperature based on the 10 ns simulations. All replica simulations were run until all molecules in the calculations clustered into a single particle. In the supplementary material table S2 shows the final MD time for the system to converge to form a single particle. Figure 2a shows the average cluster size across the five replicas for each mixture as a function of time. In the supplementary material figures S1-S7 show the variation of average cluster size, maximum cluster size and number of clusters for each of the replicates in each molecular system. All values in Figure 2a have been normalised by the total number of molecules in the system, therefore the value of 1 indicates a single cluster.

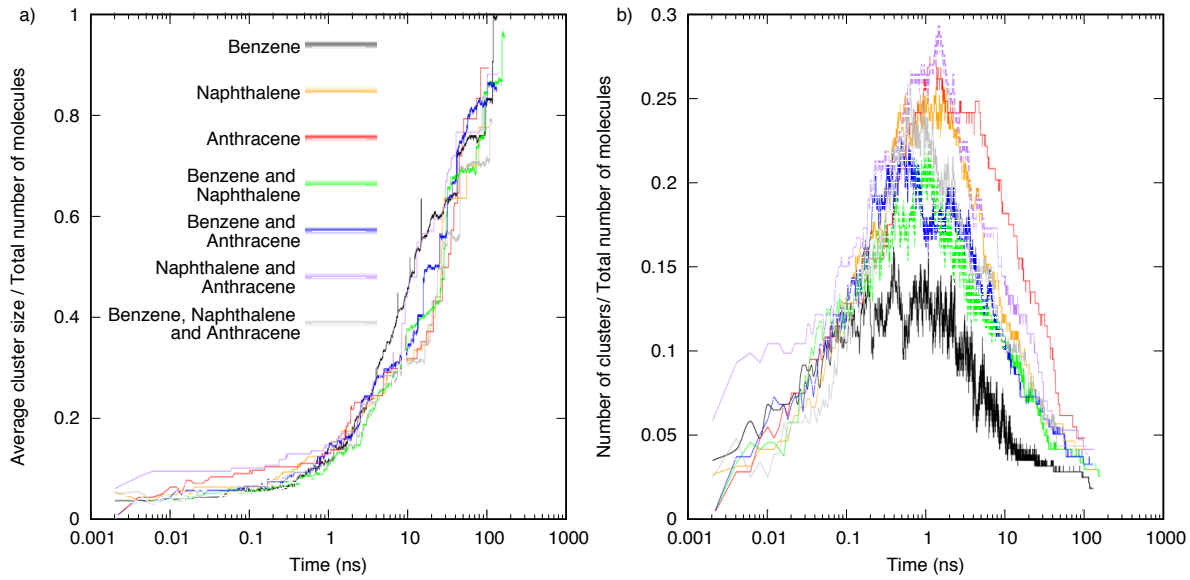

Figure 2: a) Average cluster size normalised by the total number of molecules in the system as a function of time (log scale) for all seven molecular systems at 200 K. b) Number of clusters per molecules in the system as a function of time (log scale) for all seven molecular at 200 K.

In all simulations, the mechanism of aggregation is through small clusters coalescing into larger clusters. A possible alternative mechanism of a small seed which inducing single molecule aggregation was never observed. Two distinct growth phases can be observed as demonstrated in figure 2. In the

first phase, a large number of small clusters quickly form within the first 1 ns. This is demonstrated by the small cluster size (Figure 2a) and large number of clusters (Figure 2b) at a time of 1 ns. In the second phase, after 1 ns, those smaller clusters aggregate, increasing the number of molecules comprising the cluster significantly (Figure 2a) and decreasing the number of clusters (Figure 2b). The alternative mechanism would display a constant small number of clusters, resulting in a much flatter curve for Figure 2b.

Structural analysis of clusters at 200 K

We analyse the arrangement of molecules within the clusters by pairs from the final configurations of the converged 200 K simulations. Three variables were used to describe the interactions between molecules: the distance between the centre of mass of the two molecules (r), and two angles (φ and θ) to describe the orientation with the two molecules with respect to each other. One angle is defined by the vectors normal to the ring plane of two molecules (φ), while the other (θ) angle is defined by two long axes of the molecules (see Figure 3a). The long axis in benzene is defined as the vector between the centre of mass and a carbon atom, while in naphthalene and anthracene it is defined through the longest length of the molecule. Due to the plane of symmetry through the ring plane, the range of φ and θ is 0 to 180°. The values plotted in the histogram have been corrected for sampling bias with a correction of $1/r^2$ for r, and $1/\sin\varphi$ and $1/\sin\theta$ for φ and θ respectively. The histograms have been coloured to illustrate angles between molecules at a large frequency (red), medium frequency (orange to yellow) and no frequency (black).

Figure 3b-g shows the structural analysis of molecule-molecule orientation for the entire length of the simulation. In the case of the benzene mixtures, molecules are between 4-6 Å apart and there is a large occurrence of configurations between the molecules in a T-configuration, φ=90°, and parallel-displaced (PD) configuration, φ=0° (see Figure 3b). Due to the symmetrical nature of benzene and rotation of the molecules, all possible θ values are observed with a slightly higher preference to 0°, 60°, 120° and 180° (Figure 3c). The most predominant configuration is more selective as the size of the PAHs increases. In the case of the naphthalene-naphthalene interactions, the perfect T-configuration is observed less frequently than in benzene as depicted by the yellow region at φ=90°, instead we find a configuration localised around a tilted configuration at φ=45° and 135° (Figure 3d). The long axes of the naphthalene molecules are almost aligned parallel with each other (θ=0° and 180°) as seen in Figure 3. Anthracene-anthracene interactions display the only two distinct configurations with respect to the φ angle (Figure 3f): PD configuration (at φ=0°) and tilted configuration (at φ=45°). As shown by the black region at φ=90° in Figure 3f, the T-configuration is not found in any anthracene cluster. The arrangement of naphthalene and anthracene at φ=45° and θ=0° observed in the clusters is consistent with herringbone pattern found in the crystal structure of anthracene[46] (Figure 4).

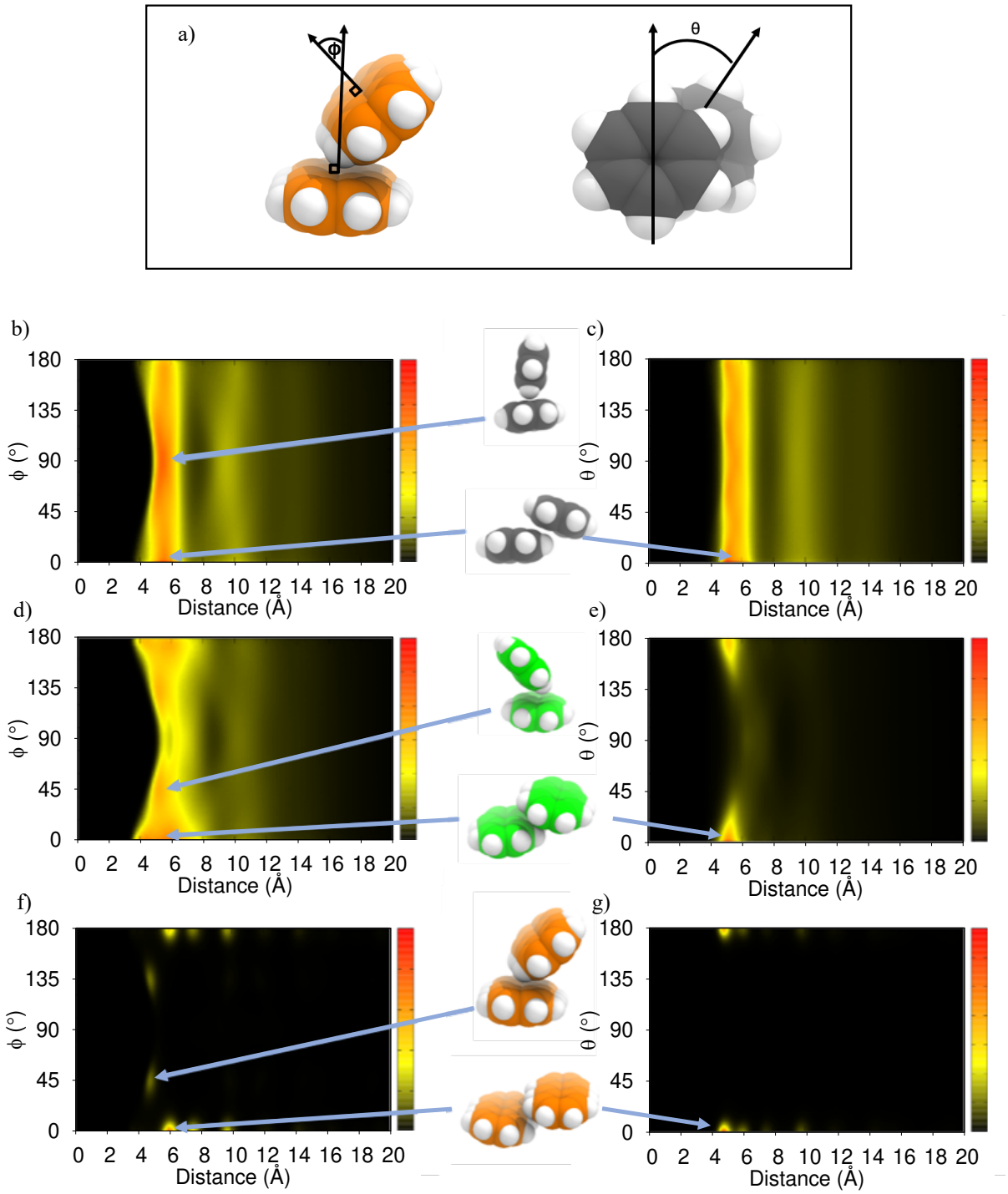

Figure 3: (a) Representation of naphthalene molecule with the vectors that define the φ angle and benzene molecule with the vectors that define the θ angle. Structural analysis of the orientation of molecules within a cluster for homogeneous systems for 200 K simulations as defined by two terms φ and θ as a function of distance, r. b-c) benzene-benzene interactions, d-e) naphthalene-naphthalene interactions and f-g) anthracene-anthracene interactions. The colours provide a measure of the number of molecules observed in the various configurations, from black indicating never observed to red indicating very frequent.

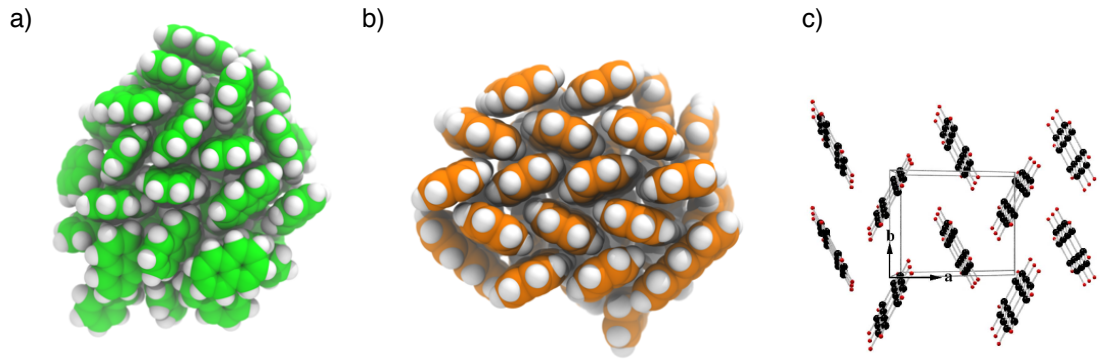

Figure 4: The final structures at 200 K of the a) naphthalene and b) anthracene clusters and c) the crystal structure of anthracene [47]

Heterogenous mixtures enables us to investigate the effect of another molecule in the structure of the cluster. In this paper we include r/φ analysis, while analysis of r/θ can be found in the supplementary material, figures S8-S9. Figure 5 shows the analysis for r/φ for all mixtures of two aromatic types. Moving down the first column of Figure 5 we see the benzene-benzene interaction in the presence of benzene, naphthalene and anthracene. They are very similar to one another, indicating that these interactions are identical in the one and two component mixtures. The same is also true for the anthracene-anthracene interaction in the third column. The benzene-benzene and anthracene-anthracene interactions are not significantly affected by the presence of other species. However, the naphthalene-naphthalene interaction behaves differently depending on which other molecules are in the mixture. The most significant change in the naphthalene-naphthalene interactions occurs in the presence of anthracene. Anthracene is coercing the naphthalene-naphthalene interaction to behave more like the anthracene-anthracene interaction, thereby restricting the rotation, removing any T-shaped type configurations and producing more herringbone pattern within the cluster.

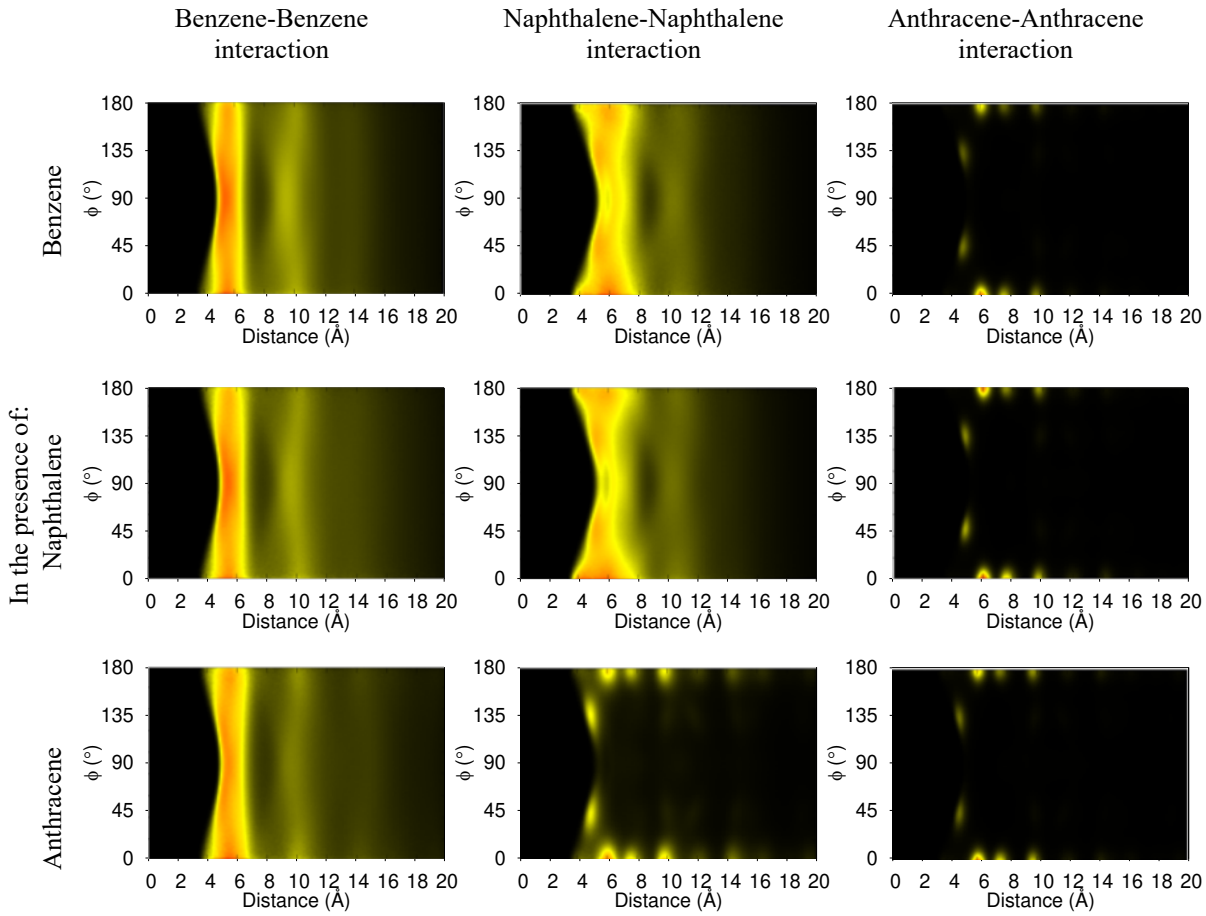

Figure 5: Structural analysis of the orientation of molecules within a cluster for heterogenous systems at 200 K as defined by φ angle as a function of distance, r. Benzene-benzene, naphthalene-naphthalene and anthracene-anthracene interactions described in the presence of benzene, naphthalene and anthracene. The colours provide a measure of the number of molecules observed in the various configurations, with black indicating none and red indicating many interactions at that configuration with graduations from yellow to orange.

Figure 6 shows the interactions in the three-component mixture. The benzene-benzene interaction and the anthracene-anthracene are the same as in the pure mixture and the two-component mixtures. Once again, the naphthalene-naphthalene interaction is modified by the presence of anthracene as in the anthracene and naphthalene mixture. Benzene-naphthalene interactions can rotate between T and PD configuration with no particular preference. Benzene-anthracene are dominated by tilted and PD configurations with limited T-configurations. Naphthalene-anthracene shows tilted and PD configurations with a total absence of the T-configuration, showing that anthracene restricts any T-arrangement.

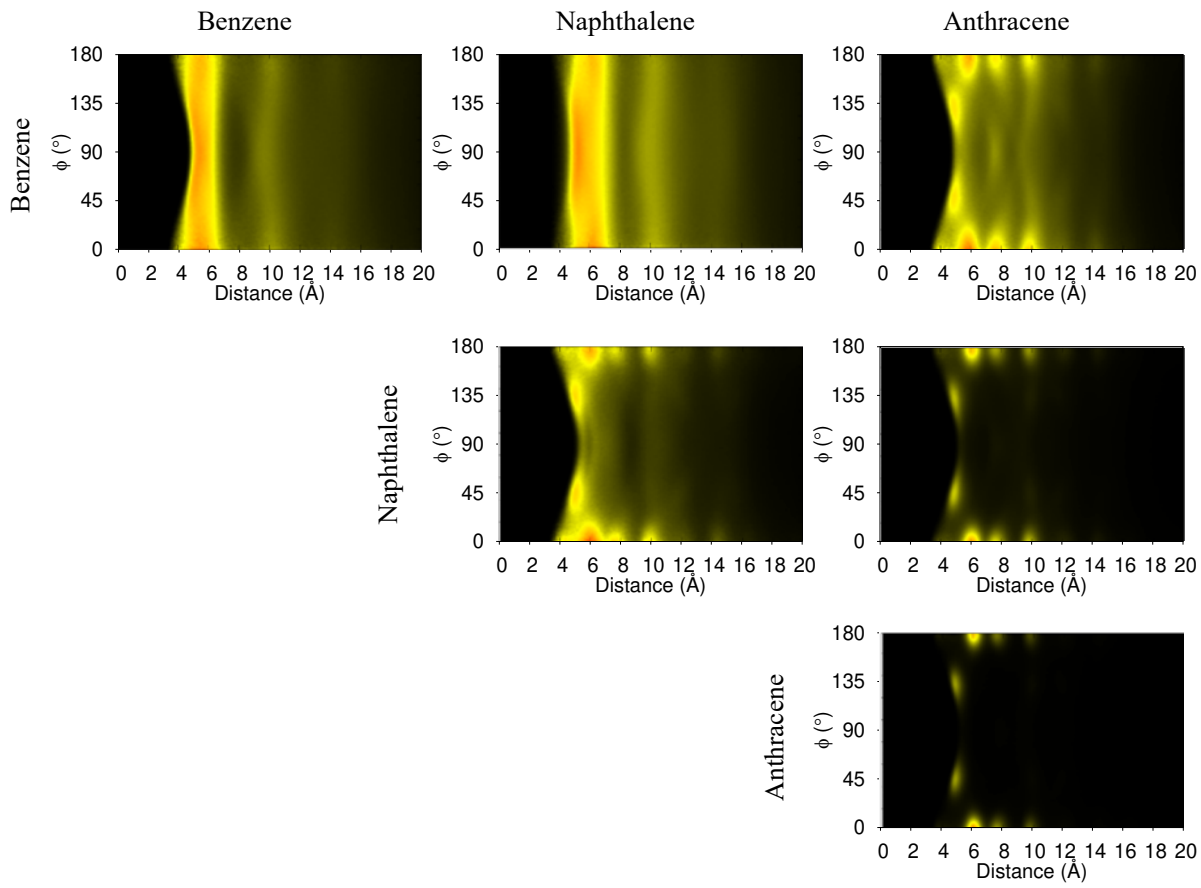

Figure 6: Structural analysis of the orientation of molecules within a cluster for heterogenous systems as defined by φ angle as a function of distance, r, for benzene-benzene, benzene-naphthalene, benzene-anthracene, naphthalene-naphthalene naphthalene-anthracene and anthracene-anthracene interactions. The colours provide a measure of the number of molecules observed in the various configurations, with black indicating none and red indicating many interactions at that configuration with graduations from yellow to orange

The final structure of each replicate of the three-component system from the longer simulations are displayed in Figure 7. All five final structures show a 'core' of anthracene in the centre of the cluster. Even when two smaller clusters merged, the anthracene cores in each cluster are seen to merge to form a new single core. As the interactions amongst aromatic molecules are primarily driven through the London dispersion forces, the strongest interactions will be the anthracene-anthracene interaction, and so the formation of the anthracene core aids to minimise the energy of each system. Naphthalene molecules fulfil the role of extending this anthracene core crystal. The naphthalene-anthracene interaction is weaker than the anthracene-anthracene interaction, and so naphthalene molecules are more mobile within the cluster than anthracene molecules. Benzene molecules provide the weakest interaction and remain mainly at the surface of the cluster.

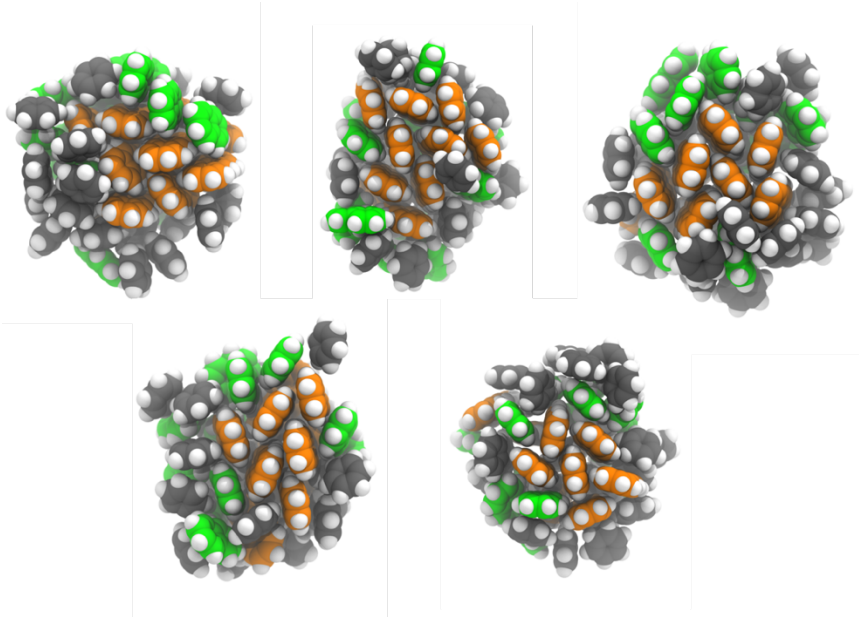

Figure 7: Final cluster of the three-species mixture for each of the five replica simulations. The centre of the cluster is predominantly made of anthracene molecules (orange), mostly surrounded by naphthalene (green) and benzene (grey) molecules in the outer surface.

Ab initio simulations of the π-π and σ-π binding energies

DHDFT calculations at the PWPB95-D3/Def2-QZVPP level of theory were performed on dimer and trimer systems to provide insight into the strength of the π-π interactions in PD configurations and σ-π interactions in T–shaped configurations. The following equation was used to determine the binding energies between two molecules:

$$\Delta E_e = E_e(A) + E_e(B) - E_e(AB) \qquad (1)$$

where ΔE_e is the binding energy of the dimer, $E_e(A)$ is the electronic energy of first molecule, $E_e(B)$ is the electronic energy of the second molecule and $E_e(AB)$ is the electronic energy of the complex. A similar calculation is performed for the trimers, however, the number of molecules in the equation would be adjusted accordingly. The equilibrium intermolecular distances and binding energies are shown in Figure 8 as well as in Table S3 in the Supporting Information. The binding energy for dimers with benzene in PD configuration increases from 11.9 kJ mol^{-1} in the case of benzene-benzene to 19.3 and 23.9 kJ mol^{-1} in the case of the benzene-naphthalene and benzene-anthracene, respectively (Figure 8a). We note that our results for the benzene-benzene PD dimer are in excellent agreement with estimated CCSD(T)/CBS. [48, 49]

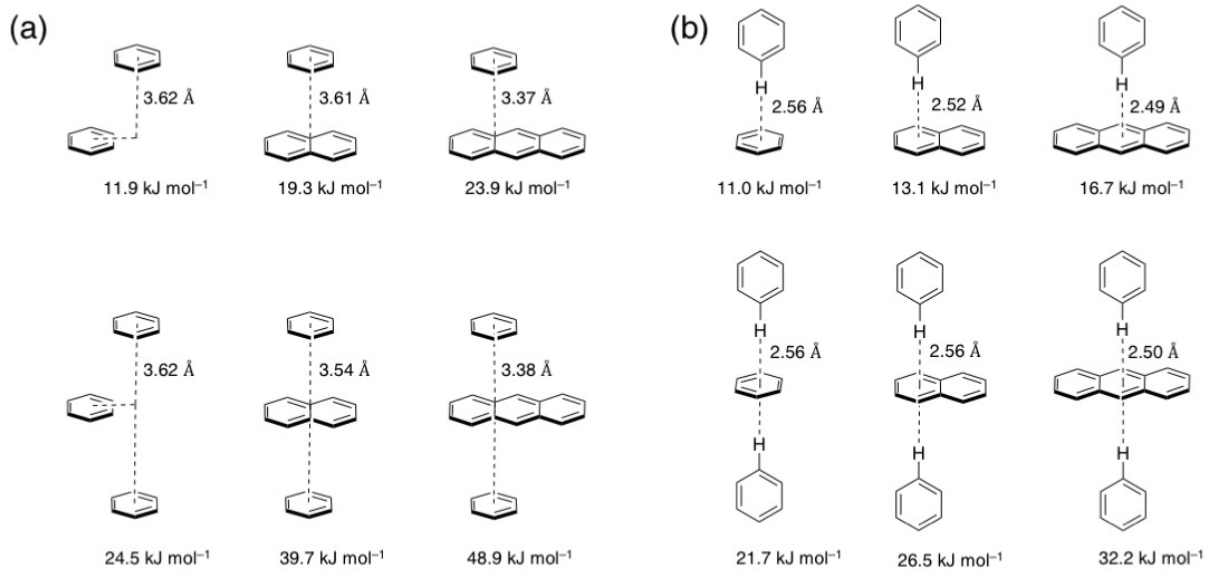

Figure 8. Parallel displaced (PD) (a) and T–shaped (b) configurations of dimers and trimers of benzene complexed with benzene, naphthalene, and anthracene. Equilibrium bond distances and PWPB95-D3/Def2-QZVPP binding energies are given (positive energies indicate an attractive interaction).

In the case of trimer complexes in PD configuration, we obtain binding energies of 24.5 for the benzene-benzene-benzene, 39.7 for the benzene-naphthalene-benzene, and 48.9 kJ mol^{-1} for the benzene-anthracene-benzene. Thus, similarly to the PD dimers the binding energy for the benzene-anthracene-benzene trimer is twice as large as that in the benzene trimer, showing that benzene prefers aggregation with other aromatics rather than itself. The increase in the binding energy in both the dimers and the trimers is also associated with shorter distances (see Figure 8). It should be noted that when moving from the dimers to the trimers the binding energy per bond (or per π-π interaction) slightly increases. This trends also apply to the T-shaped dimers and trimers. Table S3 in the supplementary material shows all binding energies for the PD and T–shaped configurations at equilibrium distances for the dimers and trimers shown in Figure 8 as well as for dimers of naphthalene-naphthalene and naphthalene-anthracene.

Our high-level DHDFT results show a significant and systematic increase in the π-π binding energies when benzene is complexed with benzene, naphthalene, and anthracene. The same trends are observed for the PD dimers and trimers. These results are consistent with the MD results which show that anthracene tends to aggregate in the centre of the aromatic clusters. Our ab initio results show that T-configurations are less stable than PD. This is consistent with our MD simulations where T-configurations are less frequent than PD.

Conclusions

In the present study, we investigate the aggregation of different aromatic molecules at temperatures similar to those found in the dust formation zone around stars. Regions close to the star (1300 K) are rich in hydrogen and acetylene [50] promoting the formation of small PAHs. While the mechanism of formation of aromatic molecules is under debate, one possible mechanism is the growth of benzene to naphthalene, anthracene and larger PAHs can be pictured via a sequence of HACA processes. [7] Dust formation from carbon stars may thus result from a first phase of pure chemical growth of aromatics in

the high-temperature, innermost regions followed by their aggregation and further growth once the gas temperatures have dropped to warm values (below 500K). Such a scenario leads to a population of small carbon clusters present at few stellar radii from the stellar surface.

The aggregation of PAHs into dust grains has not been considered in previous studies. The present study indicates the efficient clustering of small aromatic molecules through London dispersion forces at the warm temperatures (~ 300-500K) usually found at few stellar radii from the stellar surface. In our simulations, aggregation occurred for all systems containing larger polyaromatic molecules in this temperature range. Heterogeneous clustering will prevail at higher temperatures over homogeneous clustering, thereby facilitating the growth of non-planar aromatic clusters of small sizes, which may act as condensation nuclei. The structural analysis of our simulations reveals that benzene will orient itself in T-shaped and PD configurations. The structural behaviour of benzene molecules is due in part to the high degrees of symmetry within the molecule and is also further evidenced by the similar interaction energies between the two configurations. The polyaromatic molecules (naphthalene and anthracene) will not form a T-shaped configuration between molecules and will always form configurations localised around φ=45 and 135°, which is similar to the crystal packing of these systems as a solid. Moreover, in the molecular systems with all three molecules we observe evidence of a core of larger anthracene molecules surrounded by the smaller molecules. The intermolecular forces between anthracene molecules is the driving force to keep these molecules closer together in configurations close to the parallel displaced orientation.

The simulations presented in this study are the starting point for systematic and rigorous exploration of the aggregation of other aromatic molecules in homogenous and heterogenous mixtures. Recent laboratory studies, highlights the importance of small carbon chains in the dust formation process. [15] The role of carbon chains and the interaction with aromatic molecules in the production of small clusters in the dust formation zone of carbon stars needs further modelling investigation.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. This work was in part supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. This research was undertaken in part with the assistance of computational resources from the Pople high-performance computing cluster of the Faculty of Science at the University of Western Australia and with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government. AK gratefully acknowledges an Australian Research Council (ARC) Future Fellowship (Project No. FT170100373). I. S.-M. fellowship is funded by Australian Research Council (No. FT140100191). I.C. acknowledges funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC2013-SyG, Grant Agreement No. 610256 NANOCOSMOS. DT acknowledges MEAE and MESR for the travel support.

References

1. E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem.-Int. Edit.*, 2003, **42**, 1210-1250.
2. C. Janiak, *J. Chem. Soc.-Dalton Trans.*, 2000, DOI: 10.1039/b003010o, 3885-3896.
3. M. Frenklach, *Phys Chem Chem Phys*, 2002, **4**, 2028-2037.
4. H. Wang, *P Combust Inst*, 2011, **33**, 41-67.
5. H. Sabbah, L. Biennier, S. J. Klippenstein, I. R. Sims and B. R. Rowe, *J Phys Chem Lett*, 2010, **1**, 2962-2967.
6. I. Cherchneff, J. R. Barker and A. G. G. M. Tielens, *Astrophys J*, 1992, **401**, 269-287.
7. M. Frenklach and E. D. Feigelson, *Astrophys J*, 1989, **341**, 372-384.
8. P. Ehrenfreund and M. A. Sephton, *Faraday Discuss*, 2006, **133**, 277-288.
9. I. Cherchneff, *Eas Publications*, 2011, **46**, 177-189.
10. B. J. Cadwell, H. Wang, E. D. Feigelson and M. Frenklach, *Astrophys J*, 1994, **429**, 285-299.
11. P. Cau, *Astron Astrophys*, 2002, **392**, 203-213.
12. C. Jäger, F. Huisken, H. Mutschke, I. L. Jansa and T. Henning, *The Astrophysical Journal*, 2009, **696**, 706-712.
13. D. Fulvio, S. Góbi, C. Jäger, Á. Kereszturi and T. Henning, *The Astrophysical Journal Supplement Series*, 2017, **233**, 14.
14. C. S. Contreras and F. Salama, *Astrophys J Suppl S*, 2013, **208**.
15. L. Martínez, G. Santoro, P. Merino, M. Accolla, K. Lauwaet, J. Sobrado, H. Sabbah, R. J. Pelaez, V. J. Herrero, I. Tanarro, M. Agúndez, A. Martín-Jimenez, R. Otero, G. J. Ellis, C. Joblin, J. Cernicharo and J. A. Martín-Gago, *Nature Astronomy*, 2019, DOI: 10.1038/s41550-019-0899-4.
16. Q. Mao, A. C. T. van Duin and K. H. Luo, *Carbon*, 2017, **121**, 380-388.
17. M. Frenklach and A. M. Mebel, *Phys Chem Chem Phys*, 2020, **22**, 5314-5331.
18. H. Yuan, W. Kong, F. Liu and D. Chen, *Chem Eng Sci*, 2018, **195**, 748-757.
19. M. S. Celnik, M. Sander, A. Raj, R. H. West and M. Kraft, *P Combust Inst*, 2009, **32**, 639-646.
20. M. R. Kholghy, G. A. Kelesidis and S. E. Pratsinis, *Phys Chem Chem Phys*, 2018, **20**, 10926-10938.
21. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J Comput Chem*, 2005, **26**, 1781-1802.
22. R. L. Davidchack, R. Handel and M. V. Tretyakov, *J Chem Phys*, 2009, **130**.
23. R. L. Davidchack, T. E. Ouldridge and M. V. Tretyakov, *J Chem Phys*, 2015, **142**.
24. T. Darden, D. York and L. Pedersen, *J Chem Phys*, 1993, **98**, 10089-10092.
25. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *Journal of the American Chemical Society*, 1996, **118**, 11225-11236.
26. H. Takeuchi, *Journal of Physical Chemistry A*, 2012, **116**, 10172-10181.
27. H. Takeuchi, *Comput Theor Chem*, 2013, **1021**, 84-90.
28. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell, *J Comput Chem*, 2010, **31**, 671-690.
29. L. Pascazio, M. Sirignano and A. D'Anna, *Combustion and Flame*, 2017, **185**, 53-62.
30. C. R. Martinez and B. L. Iverson, *Chem Sci*, 2012, **3**, 2191-2201.
31. R. D. Zhao and R. Q. Zhang, *Phys Chem Chem Phys*, 2016, **18**, 25452-25457.
32. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenburg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian 09, (Gaussian, Inc., Wallington, CT, 2009)
33. Y. Zhao and D. G. Truhlar, *J Chem Phys*, 2006, **125**.

34. R. Ditchfield, W. J. Hehre and J. A. Pople, *J Chem Phys*, 1971, **54**, 724-+.
35. F. Weigend and R. Ahlrichs, *Phys Chem Chem Phys*, 2005, **7**, 3297-3305.
36. L. Goerigk and S. Grimme, *J Chem Theory Comput*, 2011, **7**, 291-309.
37. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J Chem Phys*, 2010, **132**.
38. S. Grimme, *Wires Comput Mol Sci*, 2011, **1**, 211-228.
39. A. D. Becke and E. R. Johnson, *J Chem Phys*, 2005, **123**.
40. S. Grimme, S. Ehrlich and L. Goerigk, *J Comput Chem*, 2011, **32**, 1456-1465.
41. C. G. Dekruif and C. H. D. Vanginkel, *J Chem Thermodyn*, 1977, **9**, 725-730.
42. A. W. Jackowski, *J Chem Thermodyn*, 1974, **6**, 49-52.
43. C. G. Dekruif, T. Kuipers, J. C. Vanmiltenburg, R. C. F. Schaake and G. Stevens, *J Chem Thermodyn*, 1981, **13**, 1081-1086.
44. X. Chen, V. Oja, W. G. Chan and M. R. Hajaligol, *J Chem Eng Data*, 2006, **51**, 386-391.
45. F. Emmenegger and M. Piccand, *Journal of Thermal Analysis and Calorimetry*, 1999, **57**, 235-240.
46. D. W. J. Cruickshank, *Acta Crystallogr*, 1956, **9**, 915-923.
47. K. Hummer, P. Puschnig, S. Sagmeister and C. Ambrosch-Draxl, *Mod Phys Lett B*, 2006, **20**, 261-280.
48. M. O. Sinnokrot, E. F. Valeev and C. D. Sherrill, *Journal of the American Chemical Society*, 2002, **124**, 10887-10893.
49. E. Miliordos, E. Apra and S. S. Xantheas, *Journal of Physical Chemistry A*, 2014, **118**, 7568-7578.
50. J. P. Fonfria, J. Cernicharo, M. J. Richter and J. H. Lacy, *The Astrophysical Journal*, 2008, **673**, 445-469.