Partitioning of polycyclic aromatic hydrocarbons in heterogeneous clusters

Kimberly Bowal[a], Jacob W. Martin[a,c], Markus Kraft[a,b,c,*]

[a]*Department of Chemical Engineering and Biotechnology, University of Cambridge, West Site, Philippa Fawcett Drive, Cambridge, CB3 0AS, United Kingdom*
[b]*School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, Singapore, 637459*
[c]*Cambridge Centre for Advanced Research and Education in Singapore (CARES), CREATE Tower, 1 Create Way, Singapore, 138602*

Abstract

The morphologies of heterogeneous clusters of polycyclic aromatic hydrocarbons (PAHs) are investigated using molecular modelling. Clusters of up to 100 molecules containing combinations of the different sized PAHs circumcoronene, coronene, ovalene, or pyrene are evaluated. Replica exchange molecular dynamics simulations using an all-atom force field parameterised for PAHs sample many configurations at high and low temperatures to determine stable low energy structures. The resulting cluster structures are evaluated using molecular radial distances and coordination numbers, and are found to be independent of initial configuration and the cluster sizes studied. Stable clusters consist of stacked PAHs in a core-shell structure, where the larger PAHs are found closer to the cluster core and the smaller PAHs are located on the cluster surface. This work provides novel insight into the molecular partitioning of heterogeneous aromatic clusters, with particular relevance to the structure of nascent soot particles.

*Corresponding author
Email address: mk306@cam.ac.uk (Markus Kraft)

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are stable fused aromatic molecules produced in high temperature processes, such as terrestrial combustion [1, 2] and interstellar supernova [3]. PAHs contribute to air pollution [4], affect the climate [5], and pose a significant health problem as a result of their toxic and carcinogenic properties [6]. Due to their stable bonding configurations at flame temperatures, PAHs are important intermediates in the formation of soot particles [7, 8]. An understanding of the PAH interactions involved in soot formation processes is required in order to effectively reduce the amount of anthropologically produced pollutants. The low species concentrations, short time frame, and small spatial extent of these processes limit experimental measurements and thus *in silico* tools provide valuable insight into soot nanoparticles [9, 10, 11, 12, 13].

In order to describe the behaviour and nature of soot particles it is necessary to understand the interactions of their components. Electronic structure calculations, such as symmetry adapted perturbation theory (SAPT(DFT)) and coupled cluster theory (CCSD(T)), provide accurate intermolecular binding energies and configurations of aromatic dimers. For small aromatics such as benzene, the T-shaped configuration (in which aromatic planes are perpendicular to each other) is the most stable, followed closely by the parallel-displaced configuration (aromatic planes are staggered face-to-face), while the sandwich configuration (aligned face-to-face) is the least favoured [14, 15, 16, 17]. This is because the hydrogen termination and π-aromatic network of these molecules creates a strong quadrupolar charge distribution that electrostatically stabilises the T-shaped configuration [18]. In contrast, the interactions between larger PAH dimers are dominated by dispersion interactions and the T-shaped configuration is substantially less stable than the parallel configurations [19, 20, 21].

Considering aromatic clusters containing more than two molecules requires

efficient parameterisation of the intermolecular interactions in order to achieve reasonable computational cost. Early approaches used the 12-6 Lennard-Jones power law functional form and atom-centred point charges fitted to the electrostatic potential [22, 2]. Improvements were made, first with a steeper exponential repulsion term [23], and then with high accuracy SAPT(DFT) calculations to generate an accurate force field, isoPAHAP [24, 10]. This PAH potential has been shown to be a transferable and highly accurate potential for systems containing planar pericondensed PAHs, for example reproducing SAPT(DFT) results of coronene dimer interaction energies, the second virial coefficients of benzene, and theoretical and experimental values of the graphite exfoliation energy [25]. Cluster stability and morphology are sensitive to the energies and potential functional form and therefore specific potentials, such as isoPAHAP, are better able to capture PAH system behaviour compared to generalised hydrocarbon potentials [26, 27, 22].

The dynamics and morphologies of homogeneous PAH clusters, containing a single PAH species, have been explored using molecular dynamics (MD) and global optimisation methods. This has allowed cluster properties such as melting points [11, 28, 29], surface accessibility [30], and gas interactions [13, 12] to be determined for PAH clusters. The morphologies of homogeneous PAH clusters have been found to be highly dependent on the type and number of PAHs present [31, 32, 22]. Small PAH clusters (for example, containing 2–6 coronene molecules) favour the formation of a single stack. Once the size of the cluster is approximately equal to the diameter of the PAH monomer (corresponding to 6 molecules for a coronene cluster), the stacks split into two and rearrange to form twisted stacks at 90° angles, known as a handshake motif [33, 22, 32, 34]. Larger clusters (for example, 12+ coronene molecules) adopt parallel stacks in a herringbone-like motif which match the bulk crystal structures [35]. These

stacked morphologies, predicted by global optimisation methods and MD studies using the isoPAHAP force field [22, 10, 27], have been confirmed recently using scanning tunnelling microscopy of coronene clusters [36].

Homogeneous clusters are often used as simple analogues of soot particles and interstellar dust even though these particles are known to contain a distribution of PAH molecule sizes, chiefly due to the increased complexity and computational cost of simulating systems containing more than one type of PAH. Only a few studies, by Wales *et al.* [22, 34], have been conducted examining heterogeneous PAH clusters. Global optimisations of heterogeneous clusters with binary combinations of coronene, hexabenzocoronene, or circumcoronene were performed using basin-hopping methods, first with an all-atom Lennard-Jones potential of clusters containing up to 10 PAHs [22] and then extended to larger clusters of up to 20 PAHs using a coarse-grained method that treats each molecule as a flattened ellipsoid [34]. In both studies, stable heterogeneous clusters showed similar structural motifs to those seen in minimised homogeneous PAH clusters, changing from single stack, handshake, and herringbone-like motifs as the number of PAHs increased. Within these stacked configurations, however, the two molecule types behaved differently, with the interactions between larger PAH interactions dominating and the smaller PAHs often capping the ends of the stacks. The coarse-grained results suggest that heterogeneous PAH clusters form stable structures in which the two molecule types are separated, although the morphology of this separation (manifesting in a core-shell or asymmetric configuration, for example) is unknown due to the fact that the clusters studied were not large enough to form more than two stacks.

Partitioning has been observed experimentally in mature soot particles, with high resolution transmission electron microscopy images showing a disordered particle core containing small fringes surrounded by larger stacked fringes [37,

38, 39, 40, 41, 42, 43, 44]. This partitioning may be physical in nature, caused by intermolecular interactions of a liquid-like particle at high temperatures in a flame. Chemical reactions, such as the carbonisation of the particle shell to form larger cross-linked molecules, may also be the source of this nanostructure. The simulation of large soot sized clusters would provide information into the contribution of intermolecular interactions in the formation of a partitioned core-shell nanostructure.

The aim of this work is to investigate the morphologies of heterogeneous PAH clusters, with a focus on understanding whether the core-shell structure of soot is due to physical partitioning of different sized PAHs. This is done by investigating the internal structure of heterogeneous clusters containing up to 100 PAHs to understand molecule partitioning at a nanoparticle size relevant to nascent soot. Replica exchange molecular dynamics allows molecular rearrangements and provides stable configurations of clusters consisting of binary and non-binary combinations of circumcoronene, coronene, ovalene, and pyrene molecules. The movement and stable orientation of molecules is evaluated qualitatively, as well as by the calculation of average radial distances and molecular coordination numbers, and related to experimentally observed soot particle nanostructures. The molecular interactions involved in stable heterogeneous structures provide insight into molecule size distributions and nanoparticle morphology not found in homogeneous PAH clusters.

2. Computational methods

Heterogeneous PAH clusters, considered here to be clusters containing two or more types of PAHs, are examined. The following molecule combinations are considered: clusters containing coronene (COR) and circumcoronene (CIR), clusters containing pyrene (PYR) and ovalene (OVA), and clusters containing all

four of the above PAHs. These molecules are selected because they span the range of PAH sizes observed in nascent soot particles, from 4–19 rings (0.5–1.1 nm) [45]. Clusters containing 32 molecules in equal molecular ratios are first assessed using several initial configurations. Large binary and non-binary clusters containing 100 molecules are then studied. In all cases, the notation AAA_xBBB_x will be used to describe the clusters studied, where AAA and BBB refer to the molecule types and x gives the number of corresponding molecules in the cluster. For example, $CIR_{50}COR_{50}$ indicates a cluster containing 100 molecules, made up of 50 circumcoronene and 50 coronene molecules.

Low energy systems indicate likely configurations of soot nanoparticles sampled from a flame. Replica exchange molecular dynamics (REMD) is a method able to rapidly determine low energy configurations by using higher energy parallel systems to explore new arrangements. REMD was developed to enhance sampling of a complex potential energy surface, based on the fact that the rate at which barrier-crossing events occur is increased with an increase in temperature. Derived from Monte Carlo parallel tempering [46], REMD involves simulating many independent isothermal systems, called replicas, in parallel across a range of temperatures [47]. At regular intervals, neighbouring replicas are able to exchange atomic coordinates based on a Boltzmann-weighted temperature dependent probability. This allows for more efficient sampling of the potential energy surface by permitting lower energy states to access higher energy configurations. As a result, after an exchange, each low energy configuration exchanged into a higher energy replica has a better opportunity to overcome energy barriers and move into a new lower energy region of phase space, while each swapped high energy configuration provides a low energy replica with a fresh configuration to sample.

In this work, REMD simulations are initialised in four independent cluster

configurations, defined by how the two molecule types are situated: mixed, janus, and two core-shell orientations (shown for a $CIR_{50}COR_{50}$ cluster in Figure 1(a-d)). All four initial configurations are simulated for the 32 molecule clusters, while only the mixed initial configuration is examined for the 100 molecule clusters. Clusters are created using the PACKMOL software [48] with the cluster radii selected to generate clusters of approximately 1.2 g/cm^3, corresponding to predicted soot densities [49]. Excess energy is removed by an energy minimisation step using the steepest descent algorithm, followed by the low-memory Broyden-Fletcher-Goldfarb-Shanno method [50].

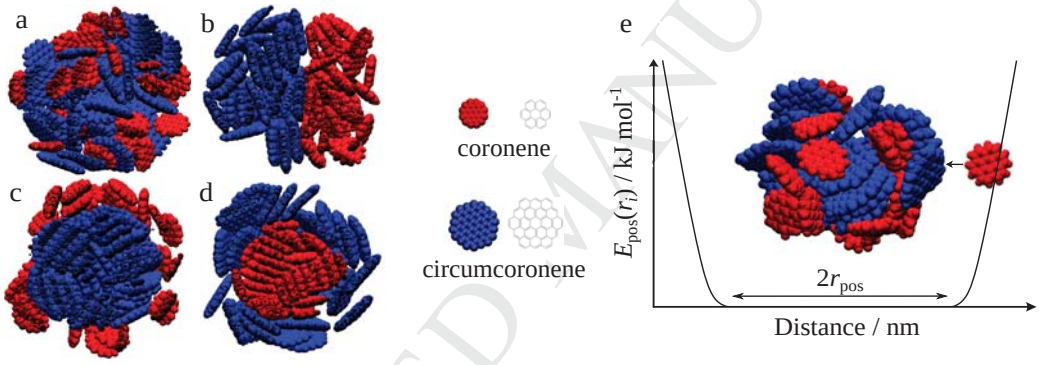

Figure 1: Initial cluster geometries are (a) mixed, (b) janus, (c) circumcoronene-core coronene-shell, and (d) coronene-core circumcoronene-shell cluster configurations, shown in cross-section. Circumcoronene molecules are shown as blue, coronene molecules are shown as red. (e) Schematic showing the use of a flat-bottomed position potential to maintain a constant cluster size. The potential does not act on molecules within the cluster, only on those that leave the cluster radius, r_{pos}. $CIR_{50}COR_{50}$ clusters are shown here and all other clusters are set up in the same way.

The canonical ensemble, defined by maintaining a constant number of atoms, system volume, and temperature, is sampled using chain of 10 Nosé-Hoover thermostats for temperature control. A velocity Verlet integrator [51] is used to advance the configuration in time and all simulations are conducted *in vacuo* without periodic boundary conditions. Intramolecular forces are determined using the

OPLS-AA force field [52] for molecular bonds, angles, dihedral and improper dihedral angles. This potential has been parametrised for liquid systems and used successfully in previous studies of PAH clusters [10, 53, 11, 54, 12]. The isoPAHAP intermolecular potential [24] is used to describe interactions between PAHs and intermolecular cut-offs are set to 3.0 nm.

A 5 ns trajectory of each system is simulated using a time step of 1 fs. Replica exchange attempts are made every 100 fs, since frequent exchanges are shown to increase efficiency without affecting the ensemble being sampled [55, 56]. Temperature distributions are 400−1500 K for the CIR_xCOR_x clusters, 200−800 K for the OVA_xPYR_x clusters, and 200−1500 K for the $CIR_xOVA_xCOR_xPYR_x$ clusters, selected so that replicas span the temperatures of solid and liquid phases. Forty replicas are used for the 32 molecule clusters and 60 replicas for the 100 molecule clusters in order to provide an overlap in potential energies of neighbouring replicas that ensures a reasonably consistent exchange rate and efficient mixing. Further information is found in the Supplementary Data. All minimisation and MD simulations are conducted using GROMACS 5.1.4 [57]. Purpose-made scripts are used to the process the output and VMD [58] provides visualisations.

Due to the size difference between molecule types, evaporation of the smaller PAHs from the cluster is observed at temperatures below the melting point of the larger PAHs. To access higher energy configurations without loss of molecules from the cluster, a position potential is implemented. This added potential acts on evaporated molecules to restrain molecules within a spherical volume. As shown in Figure 1(e), this flat-bottomed potential only applies a force on a molecule once it leaves the cluster, allowing all molecules within the cluster to move unaffected by the potential. The force applied acts towards the centre of the spherical volume and is independent of PAH size. The spherical position potential is as shown in

Equation 1.

$$E_{\mathrm{pos}}(r_i) = \frac{1}{2} k_{\mathrm{pos}} \left(|r_i - R| - r_{\mathrm{pos}} \right)^2 \mathrm{H}\left(|r_i - R| - r_{\mathrm{pos}} \right) \tag{1}$$

where E_{pos} is the position potential energy for a single atom i in kJ mol^{-1}, r_i is the coordinate position of atom i, R is the geometric centre of the cluster, r_{pos} is the position potential radius, $k_{\mathrm{pos}} = 1000$ kJ mol^{-1}nm^{-2} is the force constant, and H is the Heavyside step function. The position potential is necessary for minimum energy configurations to be attained since it allows the replicas to efficiently sample configurational space. Importantly, this applied potential is negligible at temperatures below the cluster melting point and low energy configurations are not affected by the position potential, shown in the following results.

For all clusters examined, molecular configurations are assessed by evaluating the average distance, r, of each molecule type from the geometric centre of the cluster. Additionally, molecular coordination numbers, CN, are calculated as the cumulative radial distribution function using molecule geometric centres, averaged over each molecule type. This provides an indication of the number of neighbours for each molecule type within the cluster. The maximum radial distance used in all calculations is 0.5 nm, selected to encompass the nearest interacting molecules only and provide CN values from zero to two, with $CN = 0$ corresponding to an isolated molecule, $CN = 1$ corresponding to a molecule interacting with a single other molecule, and $CN = 2$ corresponding to a molecule sandwiched between two others.

3. Results

Before considering the low energy configurations provided by REMD, standard MD simulations are conducted to evaluate the impact of the position potential. These simulations allow the dynamics of individual temperature systems to

be assessed and provide a confirmation that sufficient configurational sampling is achieved for the REMD simulations. In particular, it is important to assess whether high energy replicas are able to rigorously sample the configuration space and low energy replicas are unaffected by the position potential. Clusters are initialised in the four different configurations discussed previously and simulated at a high temperature (1600 K for CIR_xCOR_x clusters and 800 K for OVA_xPYR_x clusters) for 5 ns.

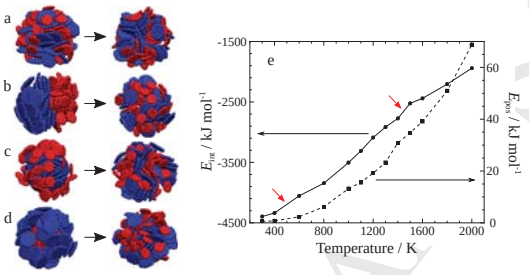

Figure 2: Initial and final configurations of $CIR_{50}COR_{50}$ clusters from high energy MD simulations. As in Figure 1, initial geometries are (a) mixed, (b) janus, and (c), (d) core-shell cluster configurations. Larger molecules (CIR, OVA) are coloured blue, and smaller molecules (COR, PYR) are red. (e) shows the energies of a $CIR_{16}COR_{16}$ cluster across a range of temperatures. E_{int} (circles) indicates the cluster intermolecular energy and E_{pos} (squares) indicates the energy contribution of the applied position potential. Red arrows indicate approximate melting point locations of the two cluster components. All energies shown are the average cluster energies for the second half of the simulation.

Figure 2(a)-(d) shows initial and final snapshots for the four starting configurations of $CIR_{50}COR_{50}$ clusters. It can be seen that regardless of initial configuration, the final clusters exhibit a well mixed configuration, which is expected of liquid clusters for which the kinetic energy of the molecules is higher than their intermolecular energy. This indicates that these high energy replicas are able to provide the configurational sampling required for REMD due to the stochastic nature of molecule movement. These results are important because they show that at high energies, the position potential keeps the cluster intact but allows mix-

ing to occur so that a particular structured morphology is not favoured due to the potential.

Figure 2(e) shows the average intermolecular energy (E_{int}) and position potential energy (E_{pos}) over a wide range of temperatures for a $CIR_{16}COR_{16}$ cluster. The sharper slopes of the change in the intermolecular energy between 1400–1500 K and 400–600 K (shown by red arrows) indicate the approximate melting point locations of the circumcoronene and coronene components, respectively. These are shifted to higher temperatures than expected based on results from analogous homogeneous clusters [11, 28], due to a pseudo-pressure effect caused by the external forces applied on the cluster by the position potential. Importantly, however, the energy contribution of the position potential is negligible at low energy replicas. This is crucial in allowing stable cluster configurations to be determined by the low energy simulations without influence from the position potential. These results show that REMD can use a position potential to sample the cluster configuration space at higher energy states and provide accurate low energy configurations that are not affected by the position potential.

REMD is applied to clusters initialised in four different configurations and the molecule size separation within the resulting structures is examined. Larger clusters containing 100 molecules, corresponding to the size of young soot particles, are then simulated. The resulting clusters are evaluated both qualitatively and quantitatively in order to provide an understanding of the interactions and structures of heterogeneous clusters.

Figure 3 shows the movement of the two molecule types over the duration of the REMD simulations for $CIR_{16}COR_{16}$ clusters, measured as the radial distance from the cluster centre over time. The low energy replicas show a distinct separation of the molecule types, with the larger molecules residing closer to the cluster centre and the smaller molecules situated farther away. This molecule type par-

titioning is not dependent on the initial cluster configurations, shown by dashed lines. The high energy replicas, shown as insets, do not show the same molecule type partitioning behaviour over time. Instead, the significant overlap between radial distance values of both molecule types throughout the simulation suggests good mixing for all molecules. The radial distance values of the high energy replicas show a larger spread than those seen in the low energy replicas, indicating that the molecules are much more mobile at high energy. Again, the behaviour of these replicas is independent of initial cluster configuration.

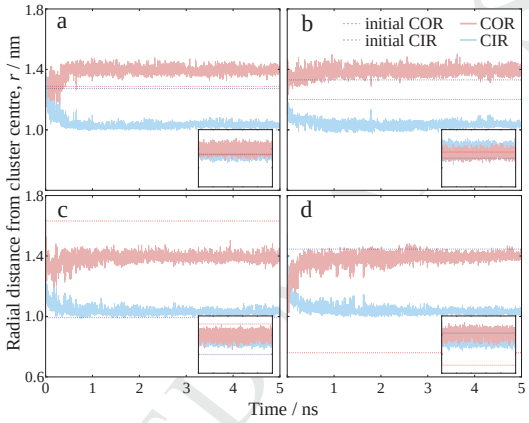

Figure 3: Radial distance of each molecule type over time for $CIR_{16}COR_{16}$ clusters at the lowest energy replica, initialised with different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. Dashed lines indicate the initial configuration radii. The inset figures show the corresponding high temperature replicas using the same axes scales. The behaviour in larger CIR_xCOR_x clusters, all OVA_xPYR_x clusters, and $CIR_xOVA_xCOR_xPYR_x$ clusters follow the same trends.

Initial and final cluster morphologies are shown in Figure 4. These cluster snapshots show clearly that the arrangement of molecule types in the final cluster configurations are similar, regardless of initial geometry or molecule types. In addition, the stable structures (shown on the right of the arrows) present a stacked morphology, indicating that dispersion interactions dominate. This stacked structure is expected due to the highly anisotropic nature of the molecule interaction

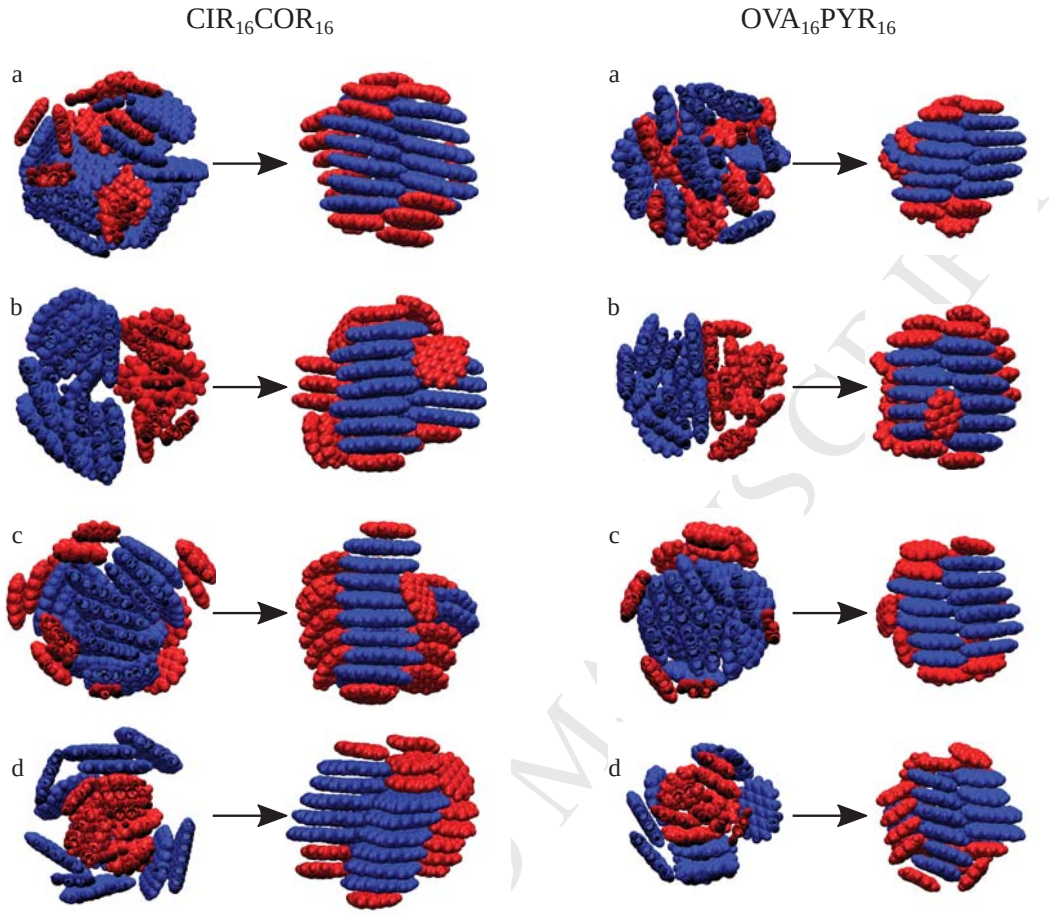

Figure 4: Initial and final configurations of $CIR_{16}COR_{16}$ and $OVA_{16}PYR_{16}$ clusters from the lowest energy replicas. The final stable configurations are similar regardless of initial configuration. As in Figure 1, initial geometries are (a) mixed, (b) janus, and (c), (d) core-shell cluster configurations, shown in cross-section. Larger molecules (CIR, OVA) are coloured blue, and smaller molecules (COR, PYR) are red.

strengths [59] and has been observed experimentally in PAH clusters [36]. As predicted by quantum calculations of PAH dimers, the slipped-parallel and sandwich molecule arrangements are dominant. As observed in previous work on PAH clusters [34], the slipped-parallel interactions are not dominant enough to produce the helical morphologies seen in theoretical discoid studies [60, 61]. Parallel stacks

show the tilted behaviour characteristic of a herringbone-like motif, which is the motif found in bulk crystal structures [35]. Some T-shaped interactions are seen between the two molecule types, highlighting that the electrostatic interactions of these aromatics do contribute [62].

A partitioning of the molecule types is evident, with the larger molecules forming parallel stacks in the core of the cluster and the smaller molecules capping the ends of these stacks and interacting more weakly in outer stacks or as individual molecules. The stacks formed by the larger PAHs contain more molecules than those consisting of the smaller PAHs, although none are larger than the limit due to the dissociation energy observed in analogous homogeneous clusters [22]. This suggests that the molecules in the core behave similarly to homogeneous clusters of the same size studied previously. The interactions of the smaller molecules around the outside of these parallel stacks indicate a clear priority to maximise carbon-carbon interactions, as observed in previous studies on homogeneous clusters [34], leading to configurations in which the larger molecules are located closer to the cluster centre. This strong molecule size partitioning was not seen in previous work where only small clusters consisting of two molecule pairs were considered [22].

Cluster snapshots provide visual results of the simulations but do not provide a complete analysis. In order to further understand the molecular distributions of these clusters, the radial distributions and coordination numbers are calculated to quantify partitioning and surface composition, respectively.

Figure 5 shows the distributions of the radial distance, measured from the centre of cluster to the centre of each molecule, for the two molecule types for all $CIR_{16}COR_{16}$ clusters over the final 3 ns of the simulations. The average radial distances, $\langle r \rangle$, from Table 1 are shown as dashed vertical lines. These figures provide an indication of the overall partitioning of the two molecule sizes over time and

allow the penetration of a given molecule type within the cluster centre to be quantified. All clusters show a similar structure in which the larger molecules are found at lower radial distances from the cluster centre than the smaller molecules. The spread of the histograms show movement of the molecules with time, but the two molecule types remain in distinct regions without significant overlap. This indicates that the larger molecules stay within the cluster core while smaller molecules are unable to penetrate into the cluster centre and are instead favoured in an outer shell. Table 1 provides the radial distances for all clusters studied and it is seen that in all cases, the larger molecules have a smaller radial distance than the smaller molecules in the cluster, indicating a core-shell structure. In addition, the difference between the molecule type radii corresponds to a single surrounding layer of smaller molecules.

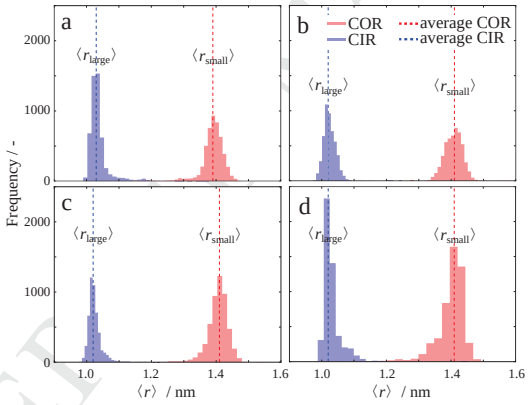

Figure 5: Molecular equilibrium radial distance distributions of $CIR_{16}COR_{16}$ configurations, considering the different initial geometries: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. Dashed vertical lines correspond to average radii and show that larger CIR molecules $\langle r_{\text{large}} \rangle$ are found closer to the cluster centre than smaller COR molecules $\langle r_{\text{small}} \rangle$.

Evaluation of the mixing between molecule types and the surface composition is provided by atomic distributions of the radial distances and coordination numbers for a single point in time, shown in Figure 6. An overall partition-

ing of the two molecule types is again seen, with the atoms within the smaller molecules occupying regions of larger radial distances than the atoms within the larger molecules. There is an overlap of these atomic radial distances, not seen in the equilibrium distributions, which indicates that the two molecule types do interact with each other within the clusters, although not to the extent that interdigited stacks are formed. All clusters show an arrangement similar to the circumcoronene-core coronene-shell initial geometry (shown in Supplementary Data Figure 11). The shape of the peaks supports a core-shell structure since the sharp peaks corresponding to the smaller molecule atoms suggest that these molecules are found in a layer around the larger molecules. The bimodal distribution seen for both molecule types in Figure 6(c), with a larger and broader peak at lower radii and a small sharp peak at high radii, can be explained by considering the stacked structures formed. As seen in Figure 4(c) this cluster contains a longer stack in the core compared to two smaller parallel stacks found in other configurations. This produces a cluster that is slightly more elongated than the other mostly spherical clusters and thus the smaller peaks found at the larger radii values correspond to the molecules at both ends of the molecular stacks.

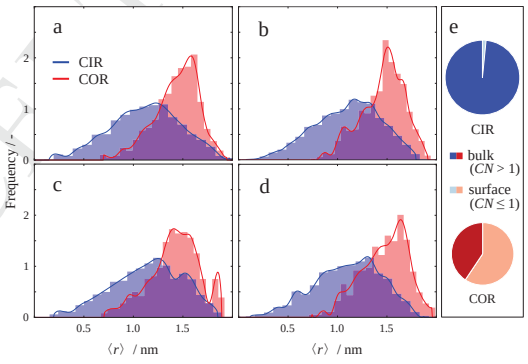

Figure 6: Normalised atomic radial distance distributions of final $CIR_{16}COR_{16}$ configurations, initialised in different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. (e) displays the bulk and surface molecule distributions for all $CIR_{16}COR_{16}$ clusters, determined by coordination numbers.

The extent of stacking of each molecule type can be expressed as a coordination number (*CN*), calculated as the average number of neighbouring molecules for each molecule within a PAH type. This value is also able to provide an assignment of surface and bulk molecules in the cluster, since molecules interacting closely with only one other molecule ($CN \leq 1$) are found on the cluster surface while molecules surrounded by more than one molecule ($CN > 1$) are indicative of bulk cluster molecules. The pie charts in Figure 6(e) highlight the proportion of molecules within each molecule type that are situated in the bulk or surface of the clusters. The darker colours for both CIR and COR molecules refer to proportion of molecules in the bulk and lighter colours show surface molecules. Nearly all of the larger CIR molecules interact with several near neighbours and are found within the cluster, while smaller COR molecules make up most of the surface molecules and display some bulk-like characteristics due to the formation of outer stacks. Table 1 provides the average *CN* values for all clusters studied at their final configuration. The *CN* values for the large molecules are near two in all cases, corresponding to each molecule experiencing close interaction with two other molecules, indicating a stacked structure. The smaller molecules have *CN* values around one, corresponding to surface molecules only located near one other molecule. Larger molecules are thus more stacked and less likely to be found on the surface of the cluster since they have higher *CN* values than the corresponding smaller molecules.

Large clusters containing 100 molecules are simulated to provide insight into soot sized nanoparticle behaviour. Final configurations of the 32 molecule clusters show that the stable configurations are not dependent on the initial structure, and thus these larger clusters are simulated only with mixed initial configurations. The resulting final configurations are shown in Figure 7(a) and (b). The morphologies of $CIR_{50}COR_{50}$ and $OVA_{50}PYR_{50}$ clusters are similar to those found for the

smaller clusters: tilted parallel stacks, with some T-shaped interactions between molecules around the surface. Molecular stacks dominate and the average molecular radial distances $\langle r \rangle$ and CN values indicate that the larger molecules are found closer to the cluster centre than the smaller molecules. This work suggests that unlike the changing stable morphologies found for small cluster sizes [63, 64, 22], clusters containing between 32 and 100 molecules form similar morphologies. The large cluster size allows the herringbone-like motif and partitioning between molecule sizes in a core-shell structure to be more clearly seen.

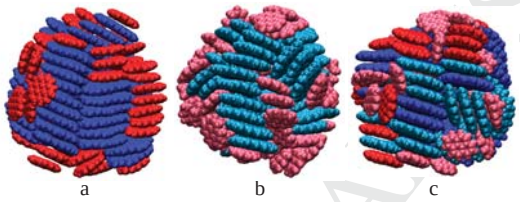

Figure 7: Final configurations of large heterogeneous PAH clusters, (a) $CIR_{50}COR_{50}$, (b) $OVA_{50}PYR_{50}$, and (c) $CIR_{25}OVA_{25}COR_{25}PYR_{25}$. The colour assignments are as follows: CIR is blue, OVA is cyan, COR is red, PYR is pink.

The PAH clusters considered so far are somewhat ideal since they contain only two molecule types. This allows for rigorous modelling but perhaps could lead to results indicative of a binary cluster only. To address this, a further REMD simulation of a large non-binary PAH cluster containing 25 molecules of circumcoronene, ovalene, coronene, and pyrene (100 molecules total) was simulated, initialised in a mixed configuration. The simulation set up follows the same procedure as the previous clusters, and the resulting replica exchange parameters are provided in the Supplementary Data. The resulting cluster configuration, shown in Figure 7(c), shows molecule size partitioning similar to that seen in the binary clusters. The largest CIR molecules are found within the cluster core along with the large OVA molecules, while the small COR molecules and smallest PYR molecules are located on the ends of the stacks containing the larger molecules

as well as in a perpendicular arrangement around the outside of the cluster. Figure 8 shows the distribution of the radial distance for the four molecules within the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster over the final 3 ns of the simulation. The average radial distances, $\langle r \rangle$, are shown as dashed vertical lines and show a partitioning of molecule types within the cluster. The *CN* values are also provided. These values indicate that the larger the molecule, the closer it is to the cluster core (lower $\langle r \rangle$ value) and the more highly stacked it is (higher *CN* value). In contrast to the binary clusters there is overlap between the molecule types, likely due to their similar sizes. However, a distinct core-shell structure is still observed in which the larger molecules form the core and the smaller molecules make up the shell.

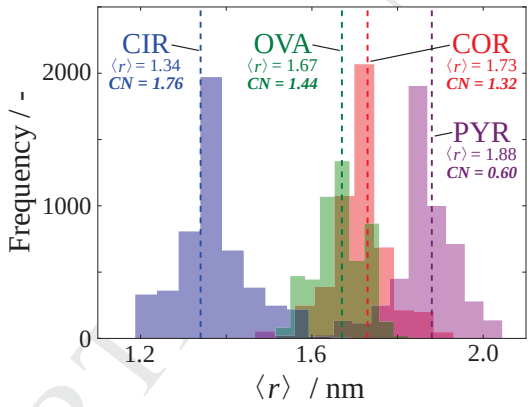

Figure 8: Molecular equilibrium radial distance distributions of $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ configurations, initialised in a mixed configuration. Average radii values are provided and shown as vertical dashed lines. Molecular *CN* values are also shown in bold italics.

Table 1 provides the binary cluster intermolecular energies, and it can be seen that this energy increases with component molecule size. The interaction energies of the large binary heterogeneous PAH clusters are between those of analogous clusters containing one type of PAH. For example, a cluster containing 50 coronene and 50 circumcoronene molecules has an approximate average intermolecular energy of $-14,000$ kJ/mol, while the energy of a cluster containing 100

coronene molecules is around $-9,400$ kJ/mol [11] and a cluster of 100 circumcoronene molecules is $-21,000$ kJ/mol [28]. Similarly, at -13442 kJ/mol, the intermolecular energy of the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster is approximately equal to that predicted by the mean of the four homogeneous 100 molecule cluster values.

Table 1: Average radial distances ($\langle r \rangle$, nm), coordination numbers (CN), and intermolecular energies ($\langle E_{\text{int}} \rangle$, kJ mol^{-1}) of heterogeneous PAH clusters from REMD simulations for different initial configurations and sizes. Subscripts refer to the large and small molecule types within each cluster. Properties are empirical equilibrium values, as indicated by angled braces, determined as the average over the final 3 ns of the simulation.

Structure	Initial configuration	$\langle r_{\text{large}} \rangle$	$\langle r_{\text{small}} \rangle$	CN_{large}	CN_{small}	$\langle E_{\text{int}} \rangle$
$CIR_{16}COR_{16}$	CIR-core COR-shell	1.02	1.41	2.00	1.56	-4590
$CIR_{16}COR_{16}$	COR-core CIR-shell	1.02	1.42	1.94	1.19	-4599
$CIR_{16}COR_{16}$	janus	1.02	1.41	2.00	0.88	-4601
$CIR_{16}COR_{16}$	mixed	1.03	1.39	2.00	1.25	-4602
$OVA_{16}PYR_{16}$	OVA-core PYR-shell	0.95	1.25	2.00	0.75	-2901
$OVA_{16}PYR_{16}$	PYR-core OVA-shell	0.95	1.19	1.94	0.94	-2901
$OVA_{16}PYR_{16}$	janus	0.87	1.21	2.00	0.88	-2907
$OVA_{16}PYR_{16}$	mixed	0.90	1.19	1.94	0.69	-2902
COR_{100} [11]	–	–	–	–	–	-9400
$CIR_{50}COR_{50}$	mixed	1.61	2.07	1.92	1.24	-14236
CIR_{100} [28]	–	–	–	–	–	-21000
PYR_{100} [11]	–	–	–	–	–	-6800
$OVA_{50}PYR_{50}$	mixed	1.40	1.73	1.78	0.82	-10244
OVA_{100} [28]	–	–	–	–	–	-12000

4. Discussion

The shape and interaction strength of anisotropic molecules govern the stable cluster morphologies they form [59]. In this way, the partitioning of molecule types within heterogeneous PAH clusters to form a core-shell structure may be caused by the different molecule sizes or molecule interaction strengths. It is difficult to assess their individual contributions since these two factors are coupled. Clusters of spherical Lennard-Jones (LJ) particles provide simple systems for which these components have been assessed independently and can provide insight into the PAH clusters studied here. LJ clusters made up of two different sized particles with the same interaction potential are known to configure in a core-shell structure, with larger particles on the surface and smaller particles in the core, in order to equalise surface and bulk interactions [65, 64, 66, 67]. In contrast, LJ clusters containing two particle types of the same size but different attractions partition in the opposite core-shell structure, in which the more strongly attractive particles make up the core and the more weakly attractive particles compose the shell [63]. Heterogeneous PAH clusters show a structure similar to the latter case, suggesting that the strength of intermolecular interactions plays a more dominant role in determining cluster structure than the molecular size differences.

A detailed examination of the average intermolecular interactions over the entire simulation period finds that the contribution from repulsion-dispersion dominates, with electrostatics contributing to less than 3% of the intermolecular interaction energy, for all clusters studied. The influential role of dispersion in these systems provides further understanding into the cluster structures and partitioning observed. The stacked configurations seen serve to maximise π-π interactions in contrast to electrostatics-dominated structures, which favour an edge-on structure such as a T-shaped configuration [68]. As an additive force, dispersion increases with increasing molecular weight due to the increased number of pair-wise in-

teractions. Thus large PAHs interact strongly with each other, while interactions between a large PAH and small PAH or between two small PAHs are less energetically favoured. This leads to an arrangement in which larger PAHs prefer to occupy interior sites in order to maximise the number of nearest neighbours between carbon atoms, and smaller PAHs are located on the surface of the cluster. The heterogeneous clusters examined have a large disparity between the molecule attraction strengths due to their different sizes. The results shown, in particular the molecule type partitioning at low energies, are more pronounced than would likely be the case between PAHs with similar interaction strengths. Therefore a cluster containing PAHs with similar sizes, for example the pericondensed PAHs circumpyrene and hexabenzocoronene with 14 and 13 aromatic rings respectively, would likely exhibit more mixing at lower energies. Evidence of this increased molecular mixing is seen visually in the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster configuration shown in Figure 7(c).

The minimum energy cluster configurations determined in this work indicate expected quenched soot structures, such as those obtained through experimental sampling methods in which soot is thermophoretically condensed onto a sampling grid for analysis. These heterogeneous clusters have diameters of 3−5 nm and C/H ratios of $2 - 2.6$, a size and composition representative of nascent soot particles [69, 70]. In addition, stable configurations contain molecular planes perpendicular to the cluster surface, a morphology that has been identified as typical of nascent soot [43]. This is distinct from the more graphitic structure of mature soot particles where the molecular planes are parallel to the particle surface.

Cluster structures show a core-shell structure in which stacked larger molecules dominate the core of the cluster. Recently, the first experimental work analysing molecule size with respect to soot particle radius at different soot maturities was published [71]. Using high resolution transmission electron microscopy, it was

seen that the molecules are slightly larger and more stacked at the core than at the surface within the youngest particles able to be sampled. This provides experimental support for the molecule size distributions of heterogeneous PAH cluster configurations determined in this work. The inverse morphology, where smaller fringes are found in the particle centre and larger fringes make up the surface layers, is reported qualitatively for mature soot particles using high resolution transmission electron microscopy [40, 37, 42]. It is known that soot particles form a graphitic shell as they mature and this work indicates that intermolecular interactions within a liquid-like particle are not responsible for the core-shell partitioning seen in mature particles. As only non-reactive molecular interactions are considered here, it is likely that chemical reactions such as carbonisation play a crucial role in determining mature soot particle structure. The identification of this inverse core-shell structure for nascent soot particles suggests their surface interactions are dominated by small PAHs. This provides a possible contributing factor for the enhanced reactivity of these young particles compared to mature soot [72].

Further work to better understand the transition from nascent to mature soot particles, likely dominated by carbonisation, and its impact on the core-shell structure is required. Curved PAH molecules have been very recently shown to form a significant component of soot particles [73] and our further work will explore the nature of curved molecule cluster morphologies. It is expected that similar stacked motifs will be observed, although molecule type partitioning in heterogeneous clusters may be less distinct since the interaction between molecules of different sizes is enhanced for these polar systems [74]. In addition, the stericity of these curved molecules may allow heterogeneity to further stabilise clusters, for example if smaller molecules are able to pack well into the concave surface of larger molecules. While REMD is able to provide stable configurations, it comes

at a very high computational cost. There is therefore still a need for efficient methods for sampling large heterogeneous aromatic clusters. We are currently exploring a Monte Carlo method for this purpose.

5. Conclusions

The morphology of large heterogeneous PAH clusters is explored using advanced molecular modelling. Replica exchange molecular dynamics simulations across a wide temperature range are able to show that clusters initialised with different geometries produce stable configurations independent of the initial cluster structure and size. Stable clusters present a parallel stacks motif, similar to the herringbone-like motif seen in bulk PAHs. The maximisation of carbon-carbon interactions dominates the arrangements formed, leading to smaller PAHs capping the ends of otherwise homogeneous stacks of larger PAHs. This agrees with previous work looking at interactions between PAH dimers and small heterogeneous clusters. For all systems studied, it is seen that heterogeneous PAH clusters favour a core-shell structure in which the larger molecules are located closer to the cluster centre. These results present the first simulations of soot sized heterogeneous PAH clusters and provide insight into the internal structure of non-carbonised nascent soot particles, useful for further understanding of soot formation and other similar systems.

Acknowledgements

This work was performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (http://www.hpc.cam.ac.uk), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England and funding from the Science and Technology Facilities Council. This work used the ARCHER UK National

Supercomputing Service (http://www.archer.ac.uk). K.B. is grateful to the Cambridge Trust and the Stanley Studentship at King's College, Cambridge for their financial support. This project is also supported by the National Research Foundation (NRF), Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

6. Supplementary Data

Replica temperatures were selected using an exponential temperature distribution

$$T_j = T_0 \exp(mj) \tag{2}$$

where T_j refers to the temperature (in K) at replica j, T_0 is the temperature (K) at replica 0, and m is a parameter which achieves the desired temperature range. For the 32 molecule systems studied, the replica index j spans from $0-39$ and $m = 0.035$. For the binary 100 molecule systems $j = 0-59$ and $m = 0.024$, and for 100 molecule systems containing four molecule types $j = 0-74$ and $m = 0.0275$.

In order to efficiently accept attempted replica exchanges, there must be adequate overlap of potential energy distributions between neighbouring temperature replicas [75]. Figure 9 shows the potential energy distribution of ten neighbouring replicas in a $CIR_{16}COR_{16}$ REMD simulation. The energy curves are broader for higher temperature replicas and it is clear that there is significant overlap between accessible energies of adjacent replica pairs, indicating good exchange rate probabilities between neighbouring replicas for this system.

The effectiveness of an REMD simulation relies on the proper exchange of states between replicas so that the low temperature states are able to sample the high temperature configurations and vice versa. Replica exchange acceptance is a good indication of the movement between replicas and it is found empirically and theoretically that an exchange acceptance of approximately 0.2 provided the

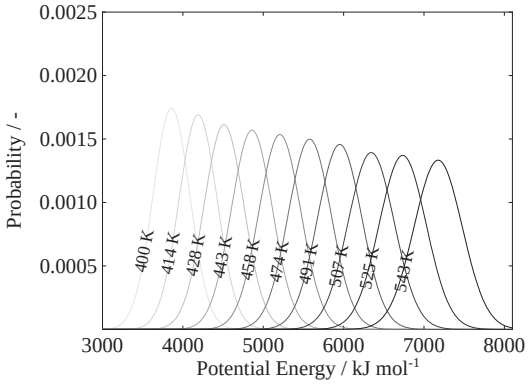

Figure 9: Potential energy distributions of neighbouring replicas from a 5 ns REMD simulation of $CIR_{16}COR_{16}$. Only the 10 lowest temperature replicas, corresponding to 400 – 543 K, are shown, but higher temperature replicas show the same trends.

best accuracy for a given computational time [76, 77]. Exchange acceptances for simulations conducted in this work are between 0.22 and 0.43, which indicates a good balance between equilibration within replicas and exchange between replicas. However the exchange acceptance is not a complete metric of REMD effectiveness since repeated exchanges between the same two replicas is treated in the same way as sequential exchanges between the full range of replicas. To ensure replicas are sampling the entire conformational space available, replica mixing is assessed by observing the movement of replica ensembles across the replica/temperature space. For all REMD simulations conducted, the replica ensembles show good sampling across replica conformations and temperatures, which is shown in a temperature trajectory for one replica ensemble in Figure 10. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures.

Figure 11 shows the normalised atomic radial distance distributions of $CIR_{16}COR_{16}$ clusters at their initial geometries. As previously, (a) corresponds to a mixed orientation, (b) janus, (c) circumcoronene-core coronene-shell, and (d) coronene-core circumcoronene-shell. This allows the molecule type separations at the start-

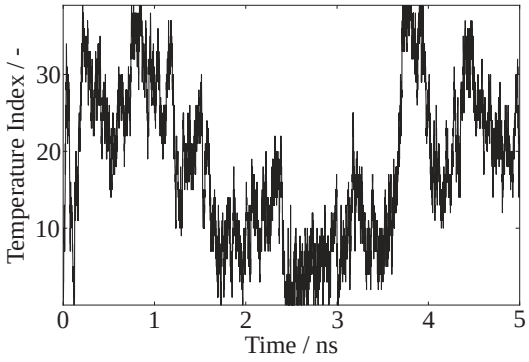

Figure 10: Temperature trajectory of the lowest energy replica for a $CIR_{16}COR_{16}$ cluster. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures. Similar configurational movement is seen for all replicas within each REMD system simulated.

ing geometries to be assessed and compared with those after the REMD simulations, shown in Figure 6.

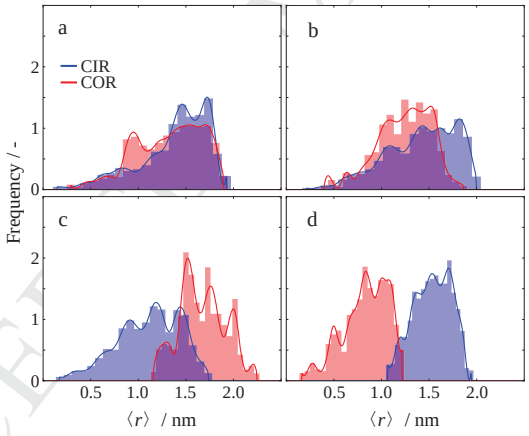

Figure 11: Normalised atomic radial distance distributions of initial $CIR_{16}COR_{16}$ configurations, initialised in different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell.

References

[1] L. B. Ebert, J. C. Scanlon, C. A. Clausen, Combustion tube soot from a diesel fuel/air mixture: Issues in structure and reactivity, Energy & Fuels 2 (4) (1988) 438–445. doi:10.1021/ef00010a009.

[2] J. D. Herdman, J. H. Miller, Intermolecular potential calculations for polynuclear aromatic hydrocarbon clusters, The Journal of Physical Chemistry A 112 (28) (2008) 6249–6256. doi:10.1021/jp800483h.

[3] C. Boersma, J. D. Bregman, L. J. Allamandola, Properties of polycyclic aromatic hydrocarbons in the northwest photon dominated region of NGC 7023. I. PAH size, charge, composition, and structure distribution, Astrophysical Journal 769 (2) (2013) 1–13. doi:10.1088/0004-637X/769/2/117.

[4] M. Guarnieri, J. R. Balmes, Outdoor air pollution and asthma, The Lancet 383 (9928) (2014) 1581–1592. arXiv:arXiv:1011.1669v3, doi:10.1016/S0140-6736(14)60617-6.

[5] E. J. Highwood, R. P. Kinnersley, When smoke gets in our eyes: The multiple impacts of atmospheric black carbon on climate, air quality and health, Environment International 32 (4) (2006) 560–566. doi:10.1016/j.envint.2005.12.003.

[6] IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures, in: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 92, 2010, pp. 1–868. doi:10.1002/14356007.a04.

[7] S. E. Stein, A. Fahr, High-temperature stabilities of hydrocarbons, The Journal of Physical Chemistry 89 (17) (1985) 3714–3725. doi:10.1021/j100263a027.

[8] R. A. Dobbins, R. A. Fletcher, H.-C. Chang, The evolution of soot precursor particles in a diffusion flame, Combustion and Flame 115 (3) (1998) 285–298. doi:10.1016/S0010-2180(98)00010-8.

[9] A. Violi, A. Venkatnathan, Combustion-generated nanoparticles produced in a benzene flame: A multiscale approach, The Journal of Chemical Physics 125 (5) (2006) 054302. doi:10.1063/1.2234481.

[10] T. S. Totton, A. J. Misquitta, M. Kraft, A quantitative study of the clustering of polycyclic aromatic hydrocarbons at high temperatures, Physical Chemistry Chemical Physics 14 (12) (2012) 4081–4094. doi:10.1039/C2CP23008A.

[11] D. Chen, T. S. Totton, J. Akroyd, S. Mosbach, M. Kraft, Size-dependent melting of polycyclic aromatic hydrocarbon nano-clusters: A molecular dynamics study, Carbon 67 (2014) 79–91. doi:10.1016/j.carbon.2013.09.058.

[12] D. Chen, T. S. Totton, J. Akroyd, S. Mosbach, M. Kraft, Phase change of polycyclic aromatic hydrocarbon clusters by mass addition, Carbon 77 (2014) 25–35. doi:10.1016/j.carbon.2014.04.089.

[13] P. Grančič, J. W. Martin, D. Chen, S. Mosbach, M. Kraft, Can nascent soot particles burn from the inside?, Carbon 109 (2016) 608–615. doi:10.1016/j.carbon.2016.08.025.

[14] P. Hobza, H. L. Selzle, E. W. Schlag, Potential energy surface for the benzene dimer. results of ab initio CCSD (T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced, The Journal of Physical Chemistry 100 (48) (1996) 18790–18794. doi:10.1021/jp961239y.

[15] E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, K. S. Kim, Understanding of assembly phenomena by aromatic-aromatic interactions: Ben-

zene Dimer and the substituted systems, Journal of Physical Chemistry A 111 (18) (2007) 3446–3457. doi:10.1021/jp068635t.

[16] M. O. Sinnokrot, C. D. Sherrill, Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations, Journal of Physical Chemistry A 108 (46) (2004) 10200–10207. doi:10.1021/jp0469517.

[17] R. Podeszwa, R. Bukowski, K. Szalewicz, Potential energy surface for the benzene dimer and perturbational analysis of π-π interactions, Journal of Physical Chemistry A 110 (34) (2006) 10345–10354. doi:10.1021/jp064095o.

[18] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, Origin of attraction and directionality of the π/π interaction: Model chemistry calculations of benzene dimer interaction, Journal of the American Chemical Society 124 (1) (2002) 104–112. doi:10.1021/ja0105212.

[19] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, High-level ab initio computations of structures and interaction energies of naphthalene dimers: Origin of attraction and its directionality, Journal of Chemical Physics 120 (2) (2004) 647–659. doi:10.1063/1.1630953.

[20] C. Gonzalez, E. C. Lim, A quantum chemistry study of the van der Waals dimers of benzene, naphthalene, and anthracene: Crossed (D_{2d}) and parallel-displaced (c_{2h}) dimers of very similar energies in the linear polyacenes, The Journal of Physical Chemistry A 104 (13) (2000) 2953–2957. doi:10.1021/jp9936421.

[21] R. Podeszwa, Interactions of graphene sheets deduced from properties of

polycyclic aromatic hydrocarbons, The Journal of Chemical Physics 132 (4) (2010) 044704. doi:10.1063/1.3300064.

[22] M. Rapacioli, F. Calvo, F. Spiegelman, C. Joblin, D. J. Wales, Stacked clusters of polycyclic aromatic hydrocarbon molecules, The Journal of Physical Chemistry A 109 (11) (2005) 2487–2497. doi:10.1021/jp046745z.

[23] D. E. Williams, Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon, Journal of Molecular Structure 485 (1999) 321–347. doi:10.1016/S0022-2860(99)00092-7.

[24] T. S. Totton, A. J. Misquitta, M. Kraft, A first principles development of a general anisotropic potential for polycyclic aromatic hydrocarbons, Journal of Chemical Theory and Computation 6 (3) (2010) 683–695. doi:10.1021/ct9004883.

[25] T. S. Totton, A. J. Misquitta, M. Kraft, Assessing the polycyclic aromatic hydrocarbon anisotropic potential with application to the exfoliation energy of graphite, The Journal of Physical Chemistry A 115 (46) (2011) 13684–13693.

[26] P. Grančič, R. Bylsma, H. Meekes, H. M. Cuppen, Evaluation of all-atom force fields for anthracene crystal growth, Crystal Growth & Design 15 (4) (2015) 1625–1633. doi:10.1021/cg5013507.

[27] L. Pascazio, M. Sirignano, A. D. Anna, Simulating the morphology of clusters of polycyclic aromatic hydrocarbons: The influence of the intermolecular potential, Combustion and Flame 185 (2017) 53–62. doi:10.1016/j.combustflame.2017.07.003.

[28] D. Chen, J. Akroyd, S. Mosbach, D. Opalka, M. Kraft, Solid-liquid transitions in homogenous ovalene, hexabenzocoronene and circumcoronene clus-

ters: A molecular dynamics study, Combustion and Flame 162 (2) (2015) 486–495. doi:10.1016/j.combustflame.2014.07.025.

[29] P. Bordat, R. Brown, Structure and molecular dynamics of crystalline and liquid anthracene and naphthalene: Possible transient rotator phase of naphthalene, The Journal of Chemical Physics 130 (12) (2009) 124501. doi:10.1063/1.3093065.

[30] D. Chen, J. Akroyd, S. Mosbach, M. Kraft, Surface reactivity of polycyclic aromatic hydrocarbon clusters, Proceedings of the Combustion Institute 35 (2) (2015) 1811–1818. doi:10.1016/j.proci.2014.06.140.

[31] H. Takeuchi, Structures, stability, and growth sequence patterns of small homoclusters of naphthalene, anthracene, phenanthrene, phenalene, naphthacene, and pyrene, Computational and Theoretical Chemistry 1021 (2013) 84–90. doi:10.1016/j.comptc.2013.06.029.

[32] M. Bartolomei, F. Pirani, J. M. C. Marques, Modeling Coronene Nanostructures: Analytical Potential, Stable Configurations and Ab Initio Energies, The Journal of Physical Chemistry C 121 (2017) 14330–14338. doi:10.1021/acs.jpcc.7b03691.

[33] J. Hernández-Rojas, F. Calvo, S. Niblett, D. J. Wales, Dynamics and thermodynamics of the coronene octamer described by coarse-grained potentials, Physical Chemistry Chemical Physics 19 (3) (2017) 1884–1895. doi:10.1039/c6cp07671h.

[34] J. Hernández-Rojas, F. Calvo, D. J. Wales, Coarse-graining the structure of polycyclic aromatic hydrocarbons clusters, Physical Chemistry Chemical Physics 18 (20) (2016) 13736–13740. doi:10.1039/c6cp00592f.

[35] R. Khanna, V. Sahajwalla, R. H. Hurt, An atomistic technique for simulating non-covalent interactions in large ensembles of high-molecular-weight polyaromatics, Carbon 43 (1) (2005) 67 – 77. doi:10.1016/j.carbon.2004.08.023.

[36] C. S. Wang, N. C. Bartelt, R. Ragan, K. Thürmer, Revealing the molecular structure of soot precursors, Carbon 129 (2018) 537–542. doi:10.1016/j.carbon.2017.12.005.

[37] T. Ishiguro, Y. Takatori, K. Akihama, Microstructure of diesel soot particles probed by electron microscopy: First observation of inner core and outer shell, Combustion and Flame 108 (1) (1997) 231–234. doi:10.1016/S0010-2180(96)00206-4.

[38] M. Alfè, B. Apicella, R. Barbella, J. N. Rouzaud, A. Tregrossi, A. Ciajolo, Structure-property relationship in nanostructures of young and mature soot in premixed flames, Proceedings of the Combustion Institute 32 I (2009) 697–704. doi:10.1016/j.proci.2008.06.193.

[39] W. Merchan-Merchan, S. Granados Sanmiguel, S. McCollam, Analysis of soot particles derived from biodiesels and diesel fuel air-flames, Fuel 102 (2012) 525–535. doi:10.1016/j.fuel.2012.04.029.

[40] M. L. Botero, D. Chen, S. González-Calera, D. Jefferson, M. Kraft, HRTEM evaluation of soot particles produced by the non-premixed combustion of liquid fuels, Carbon 96 (2016) 459–473. doi:10.1016/j.carbon.2015.09.077.

[41] A. Raj, M. Sander, V. Janardhanan, M. Kraft, A study on the coagulation of polycyclic aromatic hydrocarbon clusters to determine their collision efficiency, Combustion and Flame 157 (3) (2010) 523–534. doi:10.1016/j.combustflame.2009.10.003.

[42] B. Apicella, P. Pré, M. Alfè, A. Ciajolo, V. Gargiulo, C. Russo, et al., Soot nanostructure evolution in premixed flames by High Resolution Electron Transmission Microscopy (HRTEM), Proceedings of the Combustion Institute 35 (2) (2015) 1895–1902. doi:10.1016/j.proci.2014.06.121.

[43] M. R. Kholghy, A. Veshkini, M. J. Thomson, The core-shell internal nanostructure of soot - A criterion to model soot maturity, Carbon 100 (2016) 508–536. doi:10.1016/j.carbon.2016.01.022.

[44] P. Parent, C. Laffon, I. Marhaba, D. Ferry, T. Regier, I. Ortega, et al., 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 101 (2016) 86–100. doi:10.1016/j.carbon.2016.01.040.

[45] M. L. Botero, E. M. Adkins, S. González-Calera, H. Miller, M. Kraft, PAH structure analysis of soot in a non-premixed flame using high-resolution transmission electron microscopy and optical band gap analysis, Combustion and Flame 164 (2016) 250–258. doi:10.1016/j.combustflame.2015.11.022.

[46] K. Hukushima, K. Nemoto, Exchange Monte Carlo method and application to spin glass simulations, Journal of the Physical Society of Japan 65 (6) (1996) 1604–1608. doi:10.1143/JPSJ.65.1604.

[47] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chemical Physics Letters 314 (1999) 141–151. doi:10.1016/S0009-2614(99)01123-9.

[48] L. Martinez, R. Andrade, E. G. Birgin, J. M. Martinez, PACKMOL: A package for building initial configurations for molecular dynamics simu-

lations, Journal of Computational Chemistry 30 (13) (2009) 2157–2164. doi:10.1002/jcc.21224.

[49] T. S. Totton, D. Chakrabarti, A. J. Misquitta, M. Sander, D. J. Wales, M. Kraft, Modelling the internal structure of nascent soot particles, Combustion and Flame 157 (5) (2010) 909–914. doi:10.1016/j.combustflame.2009.11.013.

[50] R. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM Journal of Scientific Computing 16 (1995) 1190–1208. doi:10.1137/0916069.

[51] L. Verlet, Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, Physical Review 159 (1) (1967) 98–103. doi:10.1103/PhysRev.159.98.

[52] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, The Journal of Physical Chemistry B 105 (28) (2001) 6474–6487. doi:10.1021/jp003919d.

[53] S.-H. Chung, A. Violi, Peri-condensed aromatics with aliphatic chains as key intermediates for the nucleation of aromatic hydrocarbons, Proceedings of the Combustion Institute 33 (1) (2011) 693–700. doi:10.1016/j.proci.2010.06.038.

[54] P. Elvati, A. Violi, Thermodynamics of poly-aromatic hydrocarbon clustering and the effects of substituted aliphatic chains, Proceedings of the Combustion Institute 34 (1) (2013) 1837–1843. doi:10.1016/j.proci.2012.07.030.

[55] D. Sindhikara, Y. Meng, A. E. Roitberg, Exchange frequency in replica

exchange molecular dynamics, The Journal of Chemical Physics 128 (2) (2008) 01B609. doi:10.1063/1.2816560.

[56] D. J. Sindhikara, D. J. Emerson, A. E. Roitberg, Exchange often and properly in replica exchange molecular dynamics, Journal of Chemical Theory and Computation 6 (9) (2010) 2804–2808. doi:10.1021/ct100281c.

[57] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, et al., Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25. arXiv:arXiv:1503.05249v1, doi:10.1016/j.softx.2015.06.001.

[58] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, Journal of Molecular Graphics 14 (1) (1996) 33–38. doi:10.1016/0263-7855(96)00018-5.

[59] S. N. Fejer, D. Chakrabarti, D. J. Wales, Self-Assembly of anisotropic particles, Soft Matter 7 (7) (2011) 3553–3564. doi:10.1039/c0sm01289k.

[60] S. N. Fejer, D. J. Wales, Helix self-assembly from anisotropic molecules, Physical Review Letters 99 (8) (2007) 086106. doi:10.1103/PhysRevLett.99.086106.

[61] D. Chakrabarti, S. N. Fejer, D. J. Wales, Rational design of helical architectures, Proceedings of the National Academy of Sciences 106 (48) (2009) 20164–20167. doi:10.1073/pnas.0906676106.

[62] O. I. Obolensky, V. V. Semenikhina, A. V. Solov'yov, W. Greiner, Interplay of electrostatic and van der Waals forces in coronene dimer, International Journal of Quantum Chemistry 107 (6) (2007) 1335–1343. doi:10.1002/qua.21253.

[63] M. Mravlak, T. Kister, T. Kraus, T. Schilling, Structure diagram of binary Lennard-Jones clusters, The Journal of Chemical Physics 145 (2) (2016) 024302. doi:10.1063/1.4954938.

[64] J. P. K. Doye, L. Meyer, Mapping the magic numbers in binary Lennard-Jones clusters, Physical Review Letters 95 (6) (2005) 063401. doi:10.1103/PhysRevLett.95.063401.

[65] J. P. K. Doye, L. Meyer, The structure of binary Lennard-Jones clusters: The effects of atomic size ratio, ArXiv Condensed Matter e-prints cond-mat/0604250.

[66] G. Rossi, A. Rapallo, C. Mottet, A. Fortunelli, F. Baletto, R. Ferrando, Magic polyicosahedral core-shell clusters, Physical Review Letters 93 (10) (2004) 105503. doi:10.1103/PhysRevLett.93.105503.

[67] S. Cozzini, M. Ronchetti, Local icosahedral structures in binary-alloy clusters from molecular-dynamics simulation, Physical Review B - Condensed Matter and Materials Physics 53 (18) (1996) 12040–12049. doi:10.1103/PhysRevB.53.12040.

[68] C. Gonzalez, E. C. Lim, Evaluation of the Hartree-Fock Dispersion (HFD) model as a practical tool for probing intermolecular potentials of small aromatic clusters: Comparison of the HFD and MP2 intermolecular potentials, Journal of Physical Chemistry A 107 (47) (2003) 10105–10110. doi:10.1021/jp030587e.

[69] C. Russo, A. Tregrossi, A. Ciajolo, Dehydrogenation and growth of soot in premixed flames, Proceedings of the Combustion Institute 35 (2) (2015) 1803–1809. doi:10.1016/j.proci.2014.05.024.

[70] R. A. Dobbins, Soot inception temperature and the carbonization rate of precursor particles, Combustion and Flame 130 (3) (2002) 204–214. doi:10.1016/S0010-2180(02)00374-7.

[71] M. L. Botero, Y. Sheng, J. Akroyd, J. Martin, J. A. Dreyer, W. Yang, et al., Internal structure of soot particles in a diffusion flame, Carbon 141 (2019) 635–642. doi:10.1016/j.carbon.2018.09.063.

[72] J. Camacho, Y. Tao, H. Wang, Kinetics of nascent soot oxidation by molecular oxygen in a flow reactor, Proceedings of the Combustion Institute 35 (2) (2015) 1887–1894. doi:10.1016/j.proci.2014.05.095.

[73] J. W. Martin, K. Bowal, A. Menon, R. I. Slavchov, J. Akroyd, S. Mosbach, et al., Polar curved polycyclic aromatic hydrocarbons in soot formation, Proceedings of the Combustion Institutedoi:10.1016/j.proci.2018.05.046.

[74] J. W. Martin, R. I. Slavchov, E. K. Yapp, J. Akroyd, S. Mosbach, M. Kraft, The polarization of polycyclic aromatic hydrocarbons curved by pentagon incorporation: The role of the flexoelectric dipole, Journal of Physical Chemistry C 121 (48) (2017) 27154–27163. doi:10.1021/acs.jpcc.7b09044.

[75] D. A. Kofke, On the acceptance probability of replica-exchange Monte Carlo trials, The Journal of Chemical Physics 117 (15) (2002) 6911–6914. doi:10.1063/1.1507776.

[76] N. Rathore, M. Chopra, J. J. De Pablo, Optimal allocation of replicas in parallel tempering simulations, The Journal of Chemical Physics 122 (10) (2005) 24111–174903. doi:10.1063/1.1831273.

[77] A. Kone, D. A. Kofke, Selection of temperature intervals for parallel-tempering simulations, The Journal of Chemical Physics 122 (20) (2005) 206101. doi:10.1063/1.1917749.

Supplementary Data: Partitioning of polycyclic aromatic hydrocarbons in heterogeneous clusters

Kimberly Bowal[a], Jacob W. Martin[a,c], Markus Kraft[a,b,c,*]

[a]*Department of Chemical Engineering and Biotechnology, University of Cambridge, West Site, Philippa Fawcett Drive, Cambridge, CB3 0AS, United Kingdom*
[b]*School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, 637459, Singapore*
[c]*Cambridge Centre for Advanced Research and Education in Singapore (CARES), CREATE Tower, 1 Create Way, 138602, Singapore*

[*]Corresponding author
Email address: mk306@cam.ac.uk (Markus Kraft)

Replica temperatures were selected using an exponential temperature distribution

$$T_j = T_0 \exp(mj) \tag{1}$$

where T_j refers to the temperature (in K) at replica j, T_0 is the temperature (K) at replica 0, and m is a parameter which achieves the desired temperature range. For the 32 molecule systems studied, the replica index j spans from $0-39$ and $m = 0.035$. For the binary 100 molecule systems $j = 0-59$ and $m = 0.024$, and for 100 molecule systems containing four molecule types $j = 0-74$ and $m = 0.0275$.

In order to efficiently accept attempted replica exchanges, there must be adequate overlap of potential energy distributions between neighbouring temperature replicas [1]. Figure S1 shows the potential energy distribution of ten neighbouring replicas in a $CIR_{16}COR_{16}$ REMD simulation. The energy curves are broader for higher temperature replicas and it is clear that there is significant overlap between accessible energies of adjacent replica pairs, indicating good exchange rate probabilities between neighbouring replicas for this system.

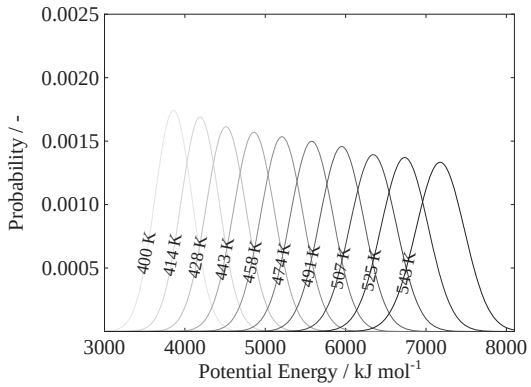

Figure S1: Potential energy distributions of neighbouring replicas from a 5 ns REMD simulation of $CIR_{16}COR_{16}$. Only the 10 lowest temperature replicas, corresponding to 400 – 543 K, are shown, but higher temperature replicas show the same trends.

The effectiveness of an REMD simulation relies on the proper exchange of states between replicas so that the low temperature states are able to sample the

high temperature configurations and vice versa. Replica exchange acceptance is a good indication of the movement between replicas and it is found empirically and theoretically that an exchange acceptance of approximately 0.2 provided the best accuracy for a given computational time [2, 3]. Exchange acceptances for simulations conducted in this work are between 0.22 and 0.43, which indicates a good balance between equilibration within replicas and exchange between replicas. However the exchange acceptance is not a complete metric of REMD effectiveness since repeated exchanges between the same two replicas is treated in the same way as sequential exchanges between the full range of replicas. To ensure replicas are sampling the entire conformational space available, replica mixing is assessed by observing the movement of replica ensembles across the replica/temperature space. For all REMD simulations conducted, the replica ensembles show good sampling across replica conformations and temperatures, which is shown in a temperature trajectory for one replica ensemble in Figure S2. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures.

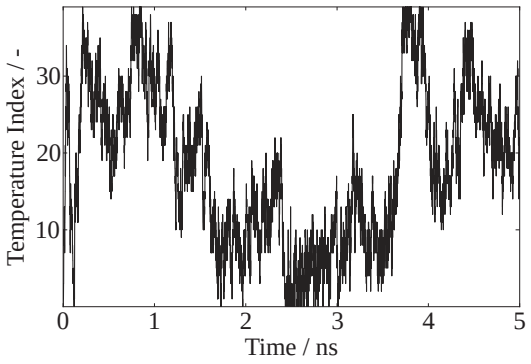

Figure S2: Temperature trajectory of the lowest energy replica for a $CIR_{16}COR_{16}$ cluster. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures. Similar configurational movement is seen for all replicas within each REMD system simulated.

Figure S3 shows the normalised atomic radial distance distributions of $CIR_{16}COR_{16}$

clusters at their initial geometries. As previously, (a) corresponds to a mixed orientation, (b) janus, (c) circumcoronene-core coronene-shell, and (d) coronene-core circumcoronene-shell. This allows the molecule type separations at the starting geometries to be assessed and compared with those after the REMD simulations, shown in Figure 6.

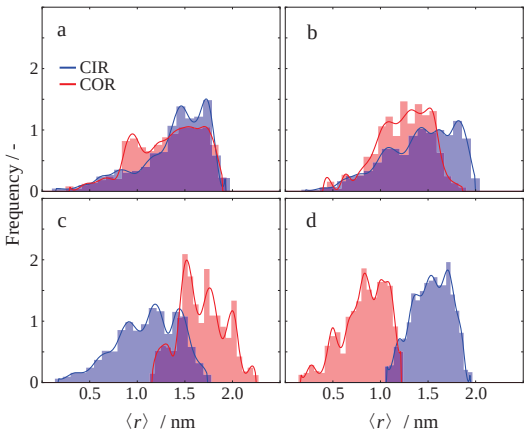

Figure S3: Normalised atomic radial distance distributions of initial $CIR_{16}COR_{16}$ configurations, initialised in different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell.

References

[1] D. A. Kofke, On the acceptance probability of replica-exchange Monte Carlo trials, The Journal of Chemical Physics 117 (15) (2002) 6911–6914. doi:10.1063/1.1507776.

[2] N. Rathore, M. Chopra, J. J. de Pablo, Optimal allocation of replicas in parallel tempering simulations, The Journal of Chemical Physics 122 (10) (2005) 24111–174903. doi:10.1063/1.1831273.

[3] A. Kone, D. A. Kofke, Selection of temperature intervals for parallel-tempering simulations, The Journal of Chemical Physics 122 (20) (2005) 206101. doi:10.1063/1.1917749.