Investigations into the Formation of Soot Particles

Tim Totton
Churchill College

Supervised by Dr Markus Kraft

12 June 2009
Preface

This dissertation is submitted for the Certificate of Post-Graduate Study. The work described in this report was carried out in the Department of Chemical Engineering and Biotechnology, University of Cambridge, between October 2008 and June 2009. The work is the result of my own research, unaided except as specifically acknowledged in the text, and it does not contain material that has already been used to any substantial extent for a comparable purpose. This dissertation contains 9,987 words on 40 pages.

________________________________________________________________________
Student                                                                 Date

Acknowledgements

I would like to thank Drs Alston Misquitta, Dwaipayan Chakrabarti and Markus Kraft for their guidance throughout the year. This work has been funded by the EPSRC, Shell Research Ltd. and Churchill College.
Summary

Soot formation and growth are complex processes for which there is much scope to further understanding. This report details investigations into the formation of soot particles as a consequence of intermolecular interactions between polycyclic aromatic hydrocarbon (PAH) molecules. This has been achieved using a combination of quantum chemistry and a molecular geometry optimisation code based on simple model potentials.

A brief introduction is given outlining some experimental and theoretical studies into soot formation processes. Following this a review of the theory of quantum chemistry is given followed by a more detailed look at the theory of intermolecular interactions and ways in which simple model potentials can be built and parameterised with ab initio quantum chemistry results.

Work undertaken to investigate PAH molecular clusters containing either pyrene (C_{16}H_{10}) or coronene (C_{24}H_{12}) is then presented. Clusters containing up to fifty molecules have been considered and comparison of TEM-style projections of the resulting minimum energy structures with experimental TEM images of soot particles reveal a number of similarities.

A comparison of two literature potentials, a Lennard-Jones plus point charges potential [1] and the W99 force field [2], with SAPT(DFT) ab initio results [3] is then presented. Whilst these potentials were both found to predict equilibrium intermolecular separations in stacked PAH geometries well, it is only the W99 force field which accurately predicts well depths and the repulsive walls in these configurations. When non-stacked geometries are considered both potentials are shown to severely over-predict repulsion due to their isotropic nature and having been parameterised using experimental crystallographic data.

In an attempt to overcome the deficiencies of these potentials a new model potential has been built which accounts for anisotropy in the electron distribution around atoms within benzene molecules. The steps taken to do this are presented along with a comparison of the new potential to SAPT(DFT) results and the W99 force field. The new potential is found to match the accuracy of the W99 force field in stacked configurations whilst also accurately predicting non-stacked configurations.

The report concludes with a consideration of further work. This includes the development of a general PAH anisotropic atom-atom potential as well as a coarse-grained potential. Possible applications of these potentials are given in the context of developing understanding of soot formation and growth processes.
Contents

1 Introduction 3

2 Background 3

3 Theory 5

3.1 Introduction to Quantum Chemistry 5

3.1.1 The Schrödinger Equation 5

3.1.2 Hartree-Fock Theory 6

3.1.3 Density Functional Theory 7

3.1.4 Basis Sets 8

3.2 Intermolecular Interactions 8

3.2.1 Ab initio Intermolecular Methods 9

3.2.2 Intermolecular Perturbation Theory 11

3.2.3 Intermolecular Basis Sets 13

3.2.4 Building an Intermolecular Model Potential 13

3.2.5 Energy Landscapes and Basin-Hopping 16

3.3 Computer Software 17

4 Results and Discussion 18

4.1 Investigating PAH clusters 18

4.2 Comparison of Intermolecular Potentials 19

4.3 Building an Anisotropic Benzene Potential 22

4.3.1 Molecular Properties 23

4.3.2 Dispersion Model 24

4.3.3 Short-Range Energies 26

4.3.4 The Overall Potential 28

5 Conclusions 29

6 Future Work 30

7 Nomenclature 33

Report produced in L\LaTeX
1 Introduction

Soot produced from incomplete combustion of hydrocarbon fuels is ubiquitous and is generally regarded as a pollutant. As well as being responsible for the discolouration of buildings [4] some now see soot as a major contributor to global warming because when deposited on snow and ice, it increases the absorption of solar radiation [5]. Of particular notoriety are the adverse health effects associated with inhalation of soot particles and there is evidence of causal links between soot inhalation and various respiratory disorders [6].

However, as well as being regarded as an unwanted pollutant, soot is deliberately made in vast quantities in the form of carbon black. Carbon black is produced in controlled conditions to give a high surface area to volume ratio, and is used extensively as an industrial feedstock in pigments for inks, coatings and plastics, and to give mechanical strength to rubber and plastic products [7].

Despite the prevalence of soot, the fundamental mechanisms of soot particle inception in flame environments are not well understood. This dissertation reports on work aimed at furthering understanding of soot formation and growth processes. In order to achieve this, effort has been channelled into developing a better understanding of the fundamental chemistry of soot particles and in particular the intermolecular forces involved between polycyclic aromatic hydrocarbon (PAH) molecules. These species are known to be amongst the precursor molecules involved in soot formation. Computational quantum chemistry methods have been used to determine the electronic structure and the intermolecular potentials existing between these molecules which determine the structures adopted by molecular clusters. This dissertation reports how such an approach can be used to develop accurate intermolecular force field models and how these can be implemented into a potential energy surface search scheme to find minimum energy PAH cluster configurations. Such clusters can be considered as the embryos of ‘primary soot particles’.

The structure of the remainder of this dissertation has been organised to reflect the specialised nature of this field and how it relates to the field of soot research. Firstly a brief introduction and review of literature relating to studies of soot particle inception from the combustion community is given. Secondly a general introduction to the theory of quantum chemistry is presented. This is then extended to the field of intermolecular forces. In the final sections the results achieved during the first year of research are reported and discussed with appropriate conclusions drawn. Finally a discussion is given of planned future work.

2 Background

Over the last few decades there have been several theoretical and experimental investigations aimed at furthering understanding of the mechanisms governing particle formation. Experimental morphological studies have been undertaken using high resolution transmis-
sion electron microscopy (HRTEM), X-ray diffraction and dark field transmission electron microscopy (DFTEM) [8–10]. These studies present evidence of crystallinity in mature soot particles which is thought to arise from the stacking of polycyclic aromatic hydrocarbon (PAH) molecules. A recent study by Vander Wal et al. [11] presents evidence indicating that soot particles can have an amorphous core surrounded by a more crystalline, or graphitic, outer shell. It is thought that amorphous regions may be formed initially through the random packing of PAH molecules which form the nucleus for further condensation of PAH molecules from the gas phase. Nevertheless it is clear that fuel type and the combustion environment can dramatically affect the type of soot produced [12]. Low temperatures tend to lead to more amorphous soots, whilst high temperatures lead to greater crystallinity. This is thought to arise due to the ability of PAH molecules to overcome the potential barrier and rearrange into graphitic-like stacked structures, either in planar form or curved. The latter occur due to the presence of 5-member rings.

There are a number of processes which have been identified as affecting the overall growth rate and morphology of soot particles:

- Condensation of PAH molecules onto the ‘primary particle’.
- Surface reactions with the gas phase leading to growth (and also depletion through oxidation) of PAH molecules within the primary particle.
- Coagulation of primary particles forming aggregates, for which the extent of particle sintering becomes important.

The process of primary particle inception in flame environments is less well understood. The presence of stacked PAH molecular structures in experimental HRTEM images of soot particles has led some to suggest that the intermolecular binding of PAH molecules may be responsible for particle inception. This hypothesis has provoked a large number of theoretical studies on the stability and relative orientation of PAH molecules present in dimers and larger stacks in flame environments [13–18]. Currently, many numerical simulations of soot formation in flames consider the dimerisation of molecules as small as pyrene (C_{16}H_{10}) [18, 19] to be the particle inception step, however the validity of this assumption is still debated [20].

In an early study this hypothesis was investigated by assuming equilibrium was met in the PAH dimerisation process. Equilibrium PAH dimer concentrations were calculated by using estimated monomer concentrations and dimerisation equilibrium constants calculated from model dispersive and electrostatic potentials [17]. It was concluded that because dimer concentrations were less than the observed number densities of the earliest soot particles, homogeneous nucleation of PAH molecules may not be as important as chemical growth in particle inception. In [21] the authors argue that the assumption of equilibrium in the PAH dimerisation steps may not hold in a flame environment, and that the dimerisation process may proceed irreversibly. In a later analysis by Miller [15] a kinetic perspective
was taken, in which the lifetimes of the dimers under flame conditions were calculated by assuming that the energy of the initial collision could be either accommodated by the PAH molecules themselves or removed by molecules in the bath. It was concluded that dimer lifetimes were only large enough compared to the characteristic time for chemical growth to be a competitive particle inception process for relatively large PAH species (>800 amu). However this analysis neglected the conversion of the collisional kinetic energy into internal molecular degrees of freedom. More recently, Schuetz and Frenklach [18] used a molecular dynamics approach using a semi-empirical potential to calculate pyrene dimer lifetimes in flame conditions. They found that the deposition of energy into internal rotations in the colliding pair greatly extended lifetimes and hence dimer formation of smaller PAH molecules may be a competitive particle inception process.

As PAH molecules are considered to be such important chemical precursor species in the soot formation process, it becomes necessary to gain a better understanding of their intermolecular chemistry. This is difficult to do experimentally and hence a more tractable approach is to employ computational quantum chemistry to calculate intermolecular potentials \textit{ab initio}.

3 Theory

3.1 Introduction to Quantum Chemistry

3.1.1 The Schrödinger Equation

The goal of quantum chemistry is to solve, or at least reach an approximate solution of the time-independent, non-relativistic Schrödinger wave equation (Eqn. 1). The resulting solution can then be used to describe the electronic behaviour of atoms and molecules. A comprehensive guide to this field is given by Atkins and Friedman [22].

\[ \hat{H}\Psi_i(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N, \vec{R}_1, \vec{R}_2, \ldots, \vec{R}_M) = E_i \Psi_i(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N, \vec{R}_1, \vec{R}_2, \ldots, \vec{R}_M). \] (1)

Here \( \hat{H} \) is the Hamilton operator for a molecular system consisting of \( M \) nuclei and \( N \) electrons in the absence of magnetic or electric fields. \( \Psi_i \) is the wavefunction of the \( i^{th} \) state of the system which is dependent upon \( 3N \) spatial coordinates and \( N \) spin coordinates associated with the electrons and \( 3M \) spatial coordinates of the nuclei. These are collectively termed \( \{ \vec{x}_k \} \) for the electrons and \( \{ \vec{R}_K \} \) for the nuclei. \( \Psi_i \) is an eigenfunction which contains all the information which can be known about the quantum system being investigated. Finally, \( E_i \) represents the value of the energy of the \( i^{th} \) state of the system.

Unfortunately, the Schrödinger equation cannot be solved exactly for multi-electron atoms and molecules and hence approximate methods are required. A key approximation employed for these many-body problems is the Born-Oppenheimer approximation. In this
approximation nuclei movements are neglected, such that their kinetic energy is zero and the
potential energy due to nucleus-nucleus repulsion becomes a constant. The classical justifi-
cation for this is based on the large mass of a proton compared to the mass of an electron,
meaning nuclear velocities are expected to be very small in comparison with electronic ve-
locities. Thus it can be assumed that electrons are always in their most stable configuration
for a given nuclear position.

Using this approximation the Hamiltonian can be reduced to give a constant nuclear
repulsion term and the so-called electronic Hamiltonian containing terms for the electron
kinetic energy, electron-electron repulsion and electron-nucleus repulsion. This is given in
atomic units as

\[
\hat{H}_{\text{elec}} = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} = \hat{T} + \hat{V}_{\text{Ne}} + \hat{V}_{\text{ee}}.
\]

The solution of the electronic Hamiltonian is the electronic wave function, \( \Psi_{\text{elec}} \) and the
electronic energy, \( E_{\text{elec}} \). The total energy, \( E_{\text{tot}} \) is given by the sum of \( E_{\text{elec}} \) and the constant
nuclear repulsion term.

### 3.1.2 Hartree-Fock Theory

The Hartree-Fock (HF) approximation is the foundation of all \textit{ab initio} many-electron wave-
function methods. In this approach the molecular wavefunctions are called the ‘Slater de-
terminants’. These Slater determinants approximate the true \( N \)-electron wavefunction by
taking the antisymmetrised product of \( N \) one-electron wavefunctions. This is also known
as the orbital approximation as each electron occupies its own orbital. The product must
be antisymmetrised to ensure the Pauli exclusion principle is obeyed, which states that all
electrons are indistinguishable, and that no two electrons can occupy the same spin state.

The one-electron wavefunctions making up the Slater determinant are termed spin or-
bitals, \( \varphi_i \), composed of one spatial orbital and one of the two possible spin functions. The
‘best’ spin orbitals are found by solving the Hartree-Fock equations where

\[
\hat{f} \varphi_i = \varepsilon_i \varphi_i, \quad i = 1, 2, \ldots, N.
\]

\( \hat{f} \) is the Fock operator which is an effective one-electron operator dependent upon the kinetic
energy of the electron, the potential energy due to electron-nucleus attraction and an average
repulsive term, known as the Hartree-Fock potential, based on the electron moving in the
mean field of the other electrons. This is known as the ‘mean field’ approximation and
replaces the complicated two-electron repulsion operator.

It should be noted that the Hartree-Fock potential found within the Fock operator, itself
depends on the spin orbitals so a pseudo-eigenfunction problem is set up which can only be
solved iteratively. This technique is called the self-consistent field (SCF) procedure, which starts from a ‘guessed’ set of orbitals from which the HF equations are solved giving a new set of orbitals for the next iteration.

The electronic Hamiltonian in the HF approximation becomes the sum of the Fock operators, operating on the Slater determinant giving the Hartree-Fock energy. The difference between the true ground state energy and the HF energy is termed the ‘correlation’ energy. This arises due to the mean field approximation and mostly represents the instantaneous repulsion of electrons. Many ‘post-Hartree-Fock’ methods, such as Møller-Plesset perturbation theory and coupled cluster theory, have been developed to try to recover the correlation energy to give a better description of the electronic structure of molecules. These methods are typically very computationally expensive, but coupled cluster theory in particular at CCSD(T) level has been used to calculate some of the most accurate electronic structures possible with quantum chemistry.

3.1.3 Density Functional Theory

Density functional theory (DFT) marks an alternative approach to wavefunction methods based on Hartree-Fock (HF) theory. In DFT methods the ground state energy is calculated directly from the electron density, \( \rho(\vec{r}) \). The energy now depends on only three (spatial) degrees of freedom rather than \( 3N \) spatial degrees of freedom associated with wavefunction methods, leading to a large computational benefit. The problem is that the exact form of the functional (i.e. a function of a function) required to calculate the energy directly from the density is unknown.

In order to use DFT practically the Kohn-Sham approach is adopted \([23]\) in which as much information as possible is computed exactly, leaving only a small part of the total energy to be determined using an approximate functional. This involves a similar procedure as with HF theory whereby a set of Kohn-Sham (KS) orbitals give a Slater determinant and are solved using a corresponding set of Kohn-Sham equations in an iterative manner.

The ground state energy is then separated into the kinetic energy, \( T \), the energy due to the nuclei, \( E_{Ne} \), the classical electrostatic electron-electron repulsion energy, \( J \), and the remainder, called the exchange-correlation energy, \( E_{XC} \). Thus the ground state energy of the system is given as:

\[
E_0[\rho(\vec{r})] = T[\rho(\vec{r})] + E_{Ne}[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{XC}[\rho(\vec{r})].
\] (4)

All terms, bar the exchange-correlation energy can in principle be determined exactly, and thus the goal of DFT based in the Kohn-Sham approach is to find suitable exchange-correlation functionals which can as accurately as possible recover the exchange-correlation energy.

There are a wide variety of exchange-correlation functionals which have been proposed
3 THEORY

[24]. Behind these functionals there are two major schemes used to calculate $E_{\text{XC}}$: the local density approximation (LDA) and the generalized gradient approximation (GGA). The LDA assumes the electron density varies slowly on an atomic scale and treats the density as a uniform electron gas with the same density, calculating the density only at the coordinate where the functional is evaluated. The GGA functionals are termed ‘semi-local’ because they additionally take into account the gradient of the electron density. In practice the GGA based exchange-correlation functional is split into separate terms for the exchange and correlation terms and approximations for the two terms are sought individually. As the exchange contributions dominate it is necessary to ensure that an accurate expression for the exchange functional is found. The HF theory can be used to give the exact exchange and as a consequence so-called hybrid functionals have been generated which mix the exact exchange energy from HF theory with the DFT exchange. Such hybrid functionals have become the mainstay of current DFT methods (e.g. B3LYP, PBE0 and B97-1).

3.1.4 Basis Sets

The Linear Combination of Atomic Orbitals (LCAO) approach is used to write the molecular KS orbitals (§ 3.1.3), $\{\varphi_i\}$, or the HF orbitals (§ 3.1.2), as a sum of predefined basis functions, $\{\eta_\mu\}$, which are commonly centred on each atomic nucleus and can be thought of as atomic orbitals.

$$\varphi_i = \sum_{\mu=1}^{L} c_{\mu} \eta_\mu.$$  (5)

The KS orbitals could only be expressed exactly in the LCAO approach if $L \to \infty$. The basis functions themselves typically have a Gaussian form and are known as Gaussian-type-orbitals (GTO). The simplest basis set is called the minimal basis set and in this case only one function (or contracted function) is used to represent each one electron molecular orbital. Significant improvements can be gained by using more than one basis function for each orbital. Such basis sets using two or three basis functions for each orbital are termed double-zeta and triple-zeta basis sets respectively.

3.2 Intermolecular Interactions

Intermolecular potentials are used to describe the interactions between molecules in ‘weakly bound’ molecular systems where a variety of individual interactions determine the overall intermolecular potential. The starting point for understanding intermolecular forces is the many-body expansion, which partitions the $N$-body interaction energy [25]. The interaction energy of a cluster of rigid molecules (such as PAH molecules), $U_{ABC...}(\Omega_{ABC...})$ where $\Omega_{ABC...}$
describes the cluster geometry, can be written in the many-body expansion as

\[ U_{ABC...}(\Omega_{ABC...}) = \sum_{X<Y} U_{XY} + \sum_{X<Y<Z} \Delta U_{XYZ} + \cdots, \]

where \( U_{XY} \) describe two-body interactions, in which each pair of molecules is treated as if no other molecules are present. The next term, \( \Delta U_{XYZ} \), describe three-body corrections, i.e. the correction that must be added to the sum of two-body terms to describe each set of three molecules correctly. This sequence continues for higher-body corrections, up to \( N \).

The two-body interactions are well understood and can be evaluated easily, however the many-body corrections are computationally very demanding to calculate. Most of these corrections are very small and can be ignored, the exception to this is the three-body polarization (induction) energy correction, which is often large. For PAH molecules however, polarization is not a large term and in this work pairwise interactions only have been considered when developing atom-atom potentials.

In principle the intermolecular potential can be found using \textit{ab initio} quantum chemistry methods but in practice these techniques are too computationally expensive when dealing with systems of many molecules. Consequently force field models are often used, in which the intermolecular interactions are constructed by assembling simple functions that describe the different contributions to the interaction on an atom pair basis. The parameters for such force fields are often found by fitting to the results of \textit{ab initio} calculations or to experimental data, e.g. heats of formation and crystallographic data.

Intermolecular interactions are broadly split into two groups: long-range and short-range. Long-range interactions persist at distances where the molecular wavefunctions no longer overlap to any significant extent and such interactions include electrostatic interactions, induction and dispersion. Short-range interactions occur due to overlap of the molecular wavefunctions and contributions to these interactions principally come from exchange-repulsion, charge transfer (attractive donor-acceptor interaction) and penetration (arising from the electron clouds of each molecule penetrating into others).

3.2.1 \textit{Ab initio} Intermolecular Methods

In this work \textit{ab initio} methods have been used to calculate accurate intermolecular energies and from these energies force fields have been developed. The first step in this process is calculate the electronic structures of the constituent molecules and can be done with the \textit{ab initio} methods described in the previous section (§ 3.1). This information is required to calculate molecular electronic properties and it is most common to use the computationally cheaper DFT methods for this task.

Explicitly calculating intermolecular energies is more difficult. The use of pure DFT to calculate intermolecular potentials is extremely limited as intermolecular separations and
binding energies, can be highly dependent on the functional used. There are also currently no practical and quantitative functionals which give interaction energies between molecules that behave like $R^{-6}$ at long-range which means they fail to describe the dispersion energy.

Wavefunction methods can be used by employing the supermolecule method. This is simple in concept and has the advantage of being easy to use at any level of theory with any basis set. If a system of two molecules, $A$ and $B$, are considered then the interaction energy, $U_{AB}$, can be expressed in terms of the ‘supermolecule’ energy, $W(A, B)$, and the energies of the isolated molecules, $W_A$ and $W_B$:

$$U_{AB} = W(A, B) - W_A - W_B,$$

(7)

However this approach has some difficulties. Firstly the interaction energy cannot be improved variationally. The variational principle can be applied to wavefunction methods to gain better estimates of the true energy for the individual molecules and the supermolecule energy calculations. However because the interaction energy is a difference between the supermolecule energy and the monomer energies, there is no way of knowing a priori whether the difference will increase or decrease as the individual terms are improved.

Using the supermolecule method also introduces ‘basis set superposition error’ (BSSE) [26]. This occurs because the basis set used is never complete and when two molecules are brought close together their basis functions will overlap and as a consequence each molecule can borrow some of the other’s basis functions. This effectively increases the basis set available and hence when the variational procedure is performed a spuriously low energy is obtained for the supermolecule compared to the constituent molecules. There are correction schemes available such as the ‘counterpoise method’ [27] but successful use depends on having a suitably accurate basis set, which adds computational cost.

The most serious shortcoming of the supermolecule method using standard wavefunction methods, in the context of building intermolecular force fields, is that the only output is the total intermolecular energy. No physical analysis is possible and this makes it hard to parameterise an atom-atom potential comprising terms describing the different attractive and repulsive forces.

Despite these shortcomings the supermolecule method has been used extensively with $ab\ initio$ methods such as MP2 and CCSD(T), and some studies have investigated PAH molecules. For example, naphthalene dimers have been investigated using estimates of the CCSD(T) interaction energy at the basis set limit using the second-order Møller-Plesset perturbation interaction energy near saturation and the CCSD(T) correction term obtained using a medium size basis set [28].

However some problems have been noted with Møller-Plesset perturbation theory, particularly when applied to weakly bound systems which show a significant amount of $\pi$-bonding (such as stacked PAH clusters). Compared to the more reliable coupled cluster calcula-
tions at CCSD(T) level, MP2 calculations have been shown to considerably overestimate attraction between molecules [28–30].

The problems of MP2 theory in weakly bound systems with π-bonding should be noted when assessing the results published by Appel et al. [19]. They studied various PAH dimers in order to find the size of PAH molecule required to form stable dimers at typical flame temperatures. They concluded that at temperatures of 1500K pyrene dimers would be stable as the calculated interaction energy was larger than the average internal energy of the two molecules. However the MP2 calculations will have overestimated the intermolecular binding energy and thus underestimated the size of PAH molecule required to form stable dimers in flame conditions. This shortcoming provides a strong motivation to use the best methods to remove this sort of error and uncertainty.

3.2.2 Intermolecular Perturbation Theory

An alternative approach to the supermolecule method is to use intermolecular perturbation theory. As intermolecular forces are relatively weak, they becomes ideal candidates for use with perturbation theory, where the perturbation is taken as the molecular interaction operator. Perturbation theory also provides an ideal framework for generating force fields, as it provides the interaction energy as a sum of physically understandable components. The energies are recovered at different orders of perturbation, but typically only first- and second-order energies are calculated. First-order energies include electrostatic and exchange energies, whilst second-order energies account for induction and dispersion energies.

Using Rayleigh-Schrödinger perturbation theory the overall Hamiltonian for the system comprises the Hamiltonian for the combined non-interacting system, $\hat{H}_0$, and the interaction operator, $\hat{V}$:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} = \hat{H}(A) + \hat{H}(B) + \lambda \hat{V}. \quad (8)$$

$\lambda$ is a parameter chosen to control the amount of perturbation added. Providing the perturbation is sufficiently weak expressions for the perturbed wavefunction and energy can be expressed as a power series in $\lambda$:

$$\Psi = \lim_{n \to \infty} \sum_{i=0}^{n} \lambda^i \Psi^{(i)}, \quad (9)$$

$$E = \lim_{n \to \infty} \sum_{i=0}^{n} \lambda^i E^{(i)}. \quad (10)$$

The order of the perturbation is determined by the exponent of the parameter, $\lambda$. The interaction energy operator describing the interactions between the electrons and nuclei of
A with those of $B$ can be written as,

$$
\hat{V} = \sum_{a \in A} \sum_{b \in B} e_a e_b r_{ab},
$$

(11)

where $e_a$ and $e_b$ are the charges in a.u. of particles $a$ and $b$ and $r_{ab}$ is the distance between these particles.

The expressions for the interaction energies which can be derived using perturbation theory are valid for all intermolecular separations. However, they are only of practical use when studying long-range intermolecular interactions because it can be assumed that the monomer wavefunctions do not overlap significantly. This allows the dimer wavefunction to be represented without antisymmetrising the product of the monomer wavefunctions. Furthermore, the interaction operator can be expanded in a multipole series and the interaction energy components can be written in terms of the properties of the unperturbed monomers. However, long-range intermolecular perturbation theory fails to account for electron exchange and hence cannot account for the repulsive exchange energies.

In general the unperturbed monomer wavefunctions required for the perturbation theory expansion are not known exactly. Only approximations are available using either Hartree-Fock theory or density functional theory (§ 3.1). This requires a different partitioning of the Hamiltonian (Eqn. 9). Also, to calculate short-range energies antisymmetrisation of the dimer wavefunction must be ensured. These requirements have led to the development of new short-range perturbation theories [26].

A recently developed perturbation theory for short-range interactions is known as symmetry-adapted perturbation theory (SAPT) [31, 32]. This method, based on HF theory (§ 3.1.2), uses a perturbation expansion in the correlation of each monomer as well as the intermolecular interaction. It is formally an iterative procedure in which projection techniques are used to ensure that the energy is calculated from correctly-antisymmetrised wavefunctions. This method is accurate but is very computationally demanding.

The most recent development in symmetry-adapted perturbation theory involves the use of density functional theory, hence the acronym SAPT(DFT) [33, 34]. First-order components of the intermolecular potential, such as electrostatics and exchange, are calculated using a Hamiltonian based on the Kohn-Sham operators (§ 3.1.3), using asymptotically-corrected functionals to ensure the correct exchange-correlation potential at large distances. Second-order components of the intermolecular potential, the induction and dispersion, are calculated using Kohn-Sham linear-response theory; this is possible because second-order energies are response energies. The resulting procedure is much simpler and faster than SAPT.

Recently SAPT(DFT) calculations of the benzene dimer [35] have been shown to compare very favourably with CCSD(T) calculations [36]. SAPT(DFT) has also been used in the context of some other simple PAH dimers: naphthalene (C$_{10}$H$_8$), anthracene (C$_{14}$H$_{10}$) and
3 THEORY

pyrene (C\textsubscript{16}H\textsubscript{10}) \cite{3}.

### 3.2.3 Intermolecular Basis Sets

The basis set requirements for intermolecular interactions are different from those used for standard monomer energy calculations. Basis sets that get the molecular properties correct are required and these tend to be bases augmented with diffuse functions and off-bond functions. The aug-cc-pVTZ or Sadlej-pVTZ basis sets are generally adequate for molecular properties and first-order interaction energies, but ‘mid-bond’ or ‘far-bond’ functions are required for accurate second-order interaction energies, such as dispersion and induction \cite{37}. These extra basis functions contribute very significantly to the convergence of the second-order energies. In this work, so called ‘monomer-centred’ basis sets have been used which construct the molecular orbitals of each monomer using the basis of the monomer, as opposed to ‘dimer-centred’ basis sets, which construct the molecular orbitals of each monomer in the entire dimer basis. When only first-order energies are important, a monomer-centred (MC) basis has been used owing to reduced computational cost. By contrast, when second-order energies are required the more expensive monomer-centred-plus (MC+) basis, which includes mid-bond functions, has been used.

### 3.2.4 Building an Intermolecular Model Potential

In this work the approach adopted has been to fit simple mathematical expressions to \textit{ab initio} SAPT(DFT) results. Repulsion and dispersion terms are universal and in a model potential are commonly represented in atom-atom form. Other terms such as electrostatic interactions are often added on an atom pair basis. The overall interaction energy of a molecular cluster is then given by the sum over all atom-atom interactions:

\[
U = \sum_{A} \sum_{A<B} \sum_{a \in A} \sum_{b \in B} U_{ab},
\]

(12)

where \(U_{ab}\) denotes the atom-atom interaction potential, the indices \(A, B\) are for molecules, and the indices \(a, b\) run over all the atomic interaction sites within a molecule.

There have been many different forms suggested for the term \(U_{ab}\), perhaps the simplest being the Lennard-Jones (L-J) 12-6 potential:

\[
U_{\text{LJ}} = 4\epsilon_{ab} \left[ \left( \frac{\sigma_{ab}}{R_{ab}} \right)^{12} - \left( \frac{\sigma_{ab}}{R_{ab}} \right)^{6} \right],
\]

(13)

where \(\epsilon_{ab}\) and \(\sigma_{ab}\) are the well-depth and size parameters describing the interaction between atoms \(a\) (in molecule \(A\)) and \(b\) (in molecule \(B\)). \(R_{ab}\) is the separation between atom \(a\) in molecule \(A\) and atom \(b\) in molecule \(B\). The \(R^{-12}\) term represents repulsion but is typically too steep. The \(R^{-6}\) term represents dispersion but in order to get the correct well-depth the
coefficient has to be much larger than the true dispersion coefficient and consequently the long-range description is also wrong. However the L-J potential is computationally efficient.

A better mathematical form to model repulsive and dispersive atom-atom interactions is the ‘exp-6’ potential:

\[ U_{\text{exp-6}} = K \exp \left[ -\alpha_{ab}(R_{ab} - \rho_{ab}) \right] - \frac{C_{ab}}{R_{ab}^6} \] \quad (14)

In this form there is an explicit dispersion coefficient, \( C_{ab} \), while the repulsive part of the interaction is in exponential or ‘Born-Mayer’ form. \( K \) is used to set the energy unit for the repulsive term (e.g. 1 milliHartree). The parameter \( \rho_{ab} \) is the separation at which the repulsion reaches a value of \( K \), so it is a size parameter. \( \alpha_{ab} \) describes the ‘hardness’ of the repulsion, and is typically around 2 bohr\(^{-1}\). However there are disadvantages to this potential as well. In this form the dispersion tends to \(-\infty\) as \( R \to 0 \), whilst the repulsion remains finite. Consequently, as the atom-atom separation is reduced a maximum is reached and the potential subsequently dives to \(-\infty\). This normally happens at values of \( R \) well inside the repulsive region, but it can lead to difficulties in some simulations when molecules approach too closely.

In general the long-range theory of dispersion energy between two atoms is given in the form

\[ U_{\text{disp}} = \frac{-C_6}{R^6} - \frac{C_7}{R^7} - \frac{C_8}{R^8} - \ldots, \] \quad (15)

where the \( C_n \) are the (angular-dependent) dispersion coefficients.

To alleviate the problem of the potential tending to \(-\infty\) at short-range ‘damping functions’ can be used whereby the dispersion energy is described by a modified term:

\[ U_{\text{disp}} = -f_6(R) \frac{C_6}{R^6} - f_7(R) \frac{C_7}{R^7} - f_8(R) \frac{C_8}{R^8} - \ldots, \] \quad (16)

where the \( f_n(R) \) are the damping functions. The damping functions are chosen to satisfy the following conditions:

- \( f_n(R) \to 1 \) as \( R \to \infty \), to recover the long-range formula.
- \( f_n(R) \to R^n \) as \( R \to 0 \), to suppress the singularity.

Using such functions in simple atom-atom potentials has not been common due to the added computational cost, however without their use the potential energy surface created from isotropic atom-atom dispersion potentials can be uncharacteristically ‘bumpy’. This is due to the atom-atom potentials typically being fitted to only a very small subset of all possible configurations (such as crystal structures) and consequently in certain configurations undamped dispersion models can give highly inaccurate results.

The most widely used damping functions are those proposed by Tang and Toennies [38]:
\[ f_n(R) = 1 - \exp(-\beta R) \sum_{k=0}^{n} \frac{(\beta R)^k}{k!} \]  

(17)

The parameter \( \beta \) in the dispersion damping functions can be determined solely from the molecular vertical ionisation energies, \( I_A \) and \( I_B \) [39]:

\[ \beta = \sqrt{2I_A} + \sqrt{2I_B} \]  

(18)

It is common to add an additional term, \( U_{es} \), to model potentials to describe the electrostatic interaction explicitly, often using partial charges assigned to atomic sites:

\[ U_{es} = \frac{q_a q_b}{R_{ab}}. \]  

(19)

Here (in a.u.) \( q_a \) and \( q_b \) represent the partial charges associated to atom \( a \) on molecule \( A \) and atom \( b \) on molecule \( B \) respectively. The point charges can be obtained through a variety of techniques which fit them in such a way as to reproduce the electrostatic potential of the molecule as accurately as possible. A more detailed description of the electrostatics can be given by using a distributed multipole analysis (DMA) [26, 40] in which multipoles up to arbitrary ranks (e.g. dipoles or quadrupoles) are centred on atom sites and are derived from the molecular wavefunction.

The expressions used in the model potentials so far treat atoms as spherical and are ‘isotropic’. Some measure of the anisotropy of individual atoms within the molecular environment can be added by the inclusion of higher moments in a distributed multipole model. However this only fixes the electrostatic interaction on the overall potential. In reality it is often found that the other terms, repulsion and, to a lesser extent, dispersion, are also dependent on the relative orientation of the atoms \( a \) and \( b \). Due to the steepness of the repulsion term, the intermolecular potential is very sensitive to any anisotropy at small separations. The repulsive term in the ‘exp-6’ model (Eqn. 14) can be adjusted to include this anisotropy by allowing the parameters, \( \alpha_{ab} \) and \( \rho_{ab} \), to be orientation-dependent. Dispersion anisotropy can be corrected by introducing orientation-dependent dispersion coefficients, \( C_{ab} \), although it is more common and computationally efficient to absorb the dispersion anisotropy in the Born-Mayer term, though this is an approximation.

The dispersion coefficients themselves can be calculated using the Williams-Stone-Misquitta (WSM) method [39]. This method calculates distributed (frequency-dependent) polarisabilities from the electron density of a given molecule using a constrained density-fitting algorithm. These polarisabilities in turn can be used to calculate anisotropic or isotropic dispersion coefficients. In this work a model potential has been developed with isotropic \( C_6 \) dispersion coefficients and any anisotropy has been accommodated within the Born-Mayer term by allowing \( \rho_{ab} \) to vary with relative orientation. This parameter is known as the shape
function, $\rho_{ab}(\Omega_{ab})$, where $\Omega_{ab}$ represents some measure of relative orientation. Using Stone’s definition \[26\] the shape function of a pair of sites is the sum of the shape functions of the individual sites:

$$\rho_{ab}(\Omega_{ab}) = \rho^a(\theta_a, \phi_a) + \rho^b(\theta_b, \phi_b), \quad (20)$$

where

$$\rho^a(\theta_a, \phi_a) = \sum_{l\kappa} \rho^a_{l\kappa} C_{l\kappa}(\theta_a, \phi_a), \quad (21)$$

with a similar expression for $\rho^b(\theta_b, \phi_b)$. Here the angles $\theta_a$ and $\phi_a$ are the polar coordinates describing the site-site vector from $a$ to $b$ in the local axis system of site $a$. It should be noted that in the implementation of these potentials it is common to define local axis systems for each molecule which can take advantage of molecular symmetry rather than relying on a global axis system. $C_{l\kappa}(\theta_a, \phi_a)$ are the renormalised spherical harmonics (in the Racah definition) and $\rho^a_{l\kappa}$ represents a constant associated with each spherical harmonic. In this work only three spherical harmonics have been used to model anisotropy:

$$C_{00} = 1 \quad (22)$$
$$C_{10} = \cos \theta \quad (23)$$
$$C_{20} = \frac{1}{2}(3 \cos^2 \theta - 1) \quad (24)$$

### 3.2.5 Energy Landscapes and Basin-Hopping

Having obtained a suitable model potential it is possible to explore the potential energy landscapes of molecular systems. In order to study the morphology of molecular clusters such systems must be optimized in terms of minimizing the overall potential energy. A computationally efficient method of scanning the potential energy surface (PES) of a molecular system is the ‘basin-hopping’ method \[41\], which is designed to search for the global minimum on a PES. This approach uses a simple transformation of the energy landscape which does not change the global minimum, or the relative energies of the local minima:

$$\tilde{E}(\vec{R}) = \min\{E(\vec{R})\}, \quad (25)$$

where ‘min’ signifies that an energy minimization is performed which locates a local minimum on the PES \[42\]. The transformed energy, $\tilde{E}(\vec{R})$, at any point, $\vec{R}$, becomes the energy of the structure obtained by minimization. Each local minimum is, therefore, surrounded by a plateau of constant energy consisting of all the neighboring geometries from which that particular minimum is obtained (Figure 1). The PES is thus transformed into a discrete set of energy levels, which, when combined with a search scheme can be used to find the associated minima. In the simplest basin-hopping scheme steps are proposed by perturbing the current coordinates and carrying out a minimization from the resulting geometry. A step
Figure 1: A schematic diagram illustrating the effects of the energy transformation for a one-dimensional example. The solid line is the energy of the original surface and the dashed line is the transformed energy $\tilde{E}$.

is accepted if the energy of the new minimum, $E_{\text{new}}$, is lower than the starting point, $E_{\text{old}}$. If $E_{\text{new}}$ is greater than $E_{\text{old}}$ then the step is accepted or rejected based on an accept-reject scheme such as the Metropolis criterion. Such a scheme allows the search to escape from traps and survey other regions of the PES.

3.3 Computer Software

- **Gaussian03**: Gaussian03 [43] is a commercial quantum chemistry software package. It has many capabilities but has been used to calculate the electron structure of molecules. There are a number of methods available, including DFT and many wavefunction methods.

- **CamCASP**: CamCASP [44] is a suite of programs designed to calculate the interaction energies between pairs of molecules, molecular properties including multipoles and frequency-dependent polarisabilities in single-site and distributed form. As an input it takes molecular orbitals calculated using DALTON [45] which is another quantum chemistry program capable of calculating electronic structures. From the molecular orbitals CamCASP performs the SAPT(DFT) intermolecular energy calculations. It can also calculate dispersion coefficients using the WSM method [39]. The program provides an overlap model which can be used to fit short-range energies using an interface with ORIENT (see below) to a Born-Mayer potential. In order to calculate multipoles CamCASP calls a program called GDMA [46] which derives multipoles of a specified rank from the wavefunction calculated in either Gaussian03 or DALTON.

- **ORIENT**: ORIENT [47] is a program used to carry out calculations of various kinds for an assembly of interacting molecules. In this work the program was used to calculate energies between molecules using atom-atom potentials. The program enables the user to specify arbitrary potentials and can handle anisotropic potentials and multipoles.
The program was also used during the fitting stage when seeking to parameterise potentials with *ab initio* results.

- **GMIN**: GMIN [48] has been used to scan the potential energy landscapes of molecular systems. It uses the basin-hopping method [41] to search for local minima. This code treats molecules as rigid bodies and has recently been extended to allow more efficient energy scans and to permit the use of both isotropic and anisotropic site-site potentials [42].

## 4 Results and Discussion

The results are presented in three sections. The first section details investigations into the structures of pyrene can coronene clusters. The second section makes a comparison of two isotropic atom-atom potentials found in the literature to *ab initio* results for two PAH dimers. In the third section, current work on developing an anisotropic benzene potential is presented with the hope of overcoming some deficiencies in the available isotropic potentials.

### 4.1 Investigating PAH clusters

As PAH molecules are relatively inflexible, it is reasonable to treat the atomic interaction sites within a rigid-body framework. As such the GMIN basin-hopping code [48] has been used to study clusters of coronene (C_{24}H_{12}) and pyrene (C_{16}H_{10}) molecules based on a simple isotropic Lennard-Jones plus point charges atom-atom potential. The Lennard-Jones contribution has been parameterised for simple organic molecules [1] (Table 1). Molecular geometries and point charges were taken from [49].

**Table 1**: Parameters for the Lennard-Jones potential [1]

<table>
<thead>
<tr>
<th>Atom Pair</th>
<th>$\epsilon$ / kJ mol$^{-1}$</th>
<th>$\sigma$ / Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C</td>
<td>0.3926</td>
<td>3.475</td>
</tr>
<tr>
<td>C H</td>
<td>0.1435</td>
<td>3.208</td>
</tr>
<tr>
<td>H H</td>
<td>0.0543</td>
<td>2.937</td>
</tr>
</tbody>
</table>

Minimum energy configurations for pyrene clusters containing two, five and ten molecules are shown in Figure 2 and similarly for coronene in Figure 3. These global minima are usually characterized by high symmetry, in agreement with the principle of maximum symmetry [50].

Larger molecular clusters were also considered up to the size of 50 molecules (Figure 4). In these larger systems a number of low-lying minima were located, but because of the stochastic nature of the potential energy surface (PES) search, and the high dimensional nature of the configurational space, it is hard to find the true global minima. However, in the context of flame environments, where soot formation is kinetically controlled and
4 RESULTS AND DISCUSSION

(a) $E/N = -27.24 \text{ kJ mol}^{-1}$ (b) $E/N = -45.09 \text{ kJ mol}^{-1}$ (c) $E/N = -51.04 \text{ kJ mol}^{-1}$

Figure 2: Low energy structures of pyrene molecular clusters with, (a) two, (b) five and (c) ten molecules.

(a) $E/N = -47.45 \text{ kJ mol}^{-1}$ (b) $E/N = -78.87 \text{ kJ mol}^{-1}$ (c) $E/N = -92.64 \text{ kJ mol}^{-1}$

Figure 3: Low energy structures of coronene molecular clusters with, (a) two, (b) five and (c) ten molecules.

occurs rapidly, it is unlikely that configurations would arrange to find the global minimum structure. Rather it is likely a local minimum is found, which due to further condensation of PAH molecules on the particle surface, becomes trapped on the PES with high energy barriers preventing the inner core of the cluster rearranging to the global minimum structure. Figure 5(a) shows a projection of a cluster of 50 coronene molecules to indicate how it might appear using transmission electron microscopy and comparison can be made to experimental HR-TEM images (Figure 5(b)). The separation between molecular planes is approximately 3.6 Å. Whilst the experimental HR-TEM images indicate the presence of PAH structures much larger than coronene, it can be noted that there are similarities in the type of structure seen and the layer separations between stacked PAH molecules.

4.2 Comparison of Intermolecular Potentials

Ab initio SAPT(DFT) calculations [3] for naphthalene (C$_{10}$H$_8$) and anthracene (C$_{14}$H$_{10}$) dimers have been used to assess the accuracy of currently available isotropic atom-atom potentials. Two potentials have been compared; the Lennard-Jones plus point charges potential (§ 4.1) and the W99 force field [2], each having been parameterized using experimental data.
RESULTS AND DISCUSSION

(a) $E/N = -65.76 \text{ kJ mol}^{-1}$  
(b) $E/N = -105.46 \text{ kJ mol}^{-1}$

Figure 4: Low energy molecular clusters of (a) 50 pyrene molecules and (b) 50 coronene molecules.

(Table 1 and Table 2). The W99 force field has a modified ‘exp-6’ form and adds an explicit point charge term. In atomic units it is given as

$$U_{W99} = B_{ab} \exp(-C_{ab} R_{ab}) - A_{ab} R_{ab}^6 + q_a q_b R_{ab}. \quad (26)$$

The monomer geometries were optimized using the Gaussian03 program [43] using DFT with the B3LYP functional and the 6-31G* basis set as used by Podeszwa et al. [3]. Partial atomic charges, needed for the potentials, were computed using Gaussian03 to fit the electrostatic potential at points selected according to the Merz-Singh-Kollman scheme using DFT with the PBE0 functional and Sadlej-pVTZ basis set.

<table>
<thead>
<tr>
<th>Atom Pair</th>
<th>$A$ / kJ mol$^{-1}$</th>
<th>$A^b$</th>
<th>$B$ / kJ mol$^{-1}$</th>
<th>$C$ / Å$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C</td>
<td>1701.73</td>
<td>270363</td>
<td>3.60</td>
<td></td>
</tr>
<tr>
<td>C H</td>
<td>688.27</td>
<td>58551</td>
<td>3.58</td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>278.37</td>
<td>12680</td>
<td>3.56</td>
<td></td>
</tr>
</tbody>
</table>

The SAPT(DFT) results were obtained using the PBE0 functional with asymptotic correction and the aug-cc-pVDZ monomer-centred ‘plus’ (MC+) basis set. The specific naphthalene and anthracene dimer configurations are shown in Figure 6 and Figure 7, where the slipped parallel, graphite-type and crossed geometries represent stacked dimer configurations and the interaction energy has been calculated at various interplanar spacings. The T-shaped configuration energies have been calculated at various separations of the monomer centres of mass along the vector normal to the plane of the lower monomer.
RESULTS AND DISCUSSION

(a) A TEM-style projection of a cluster of 50 coronene molecules

(b) Experimental HR-TEM images of an aggregate sampled from a diesel engine. Indicated are length scales of structures within a primary particle [51]

Figure 5: Comparison of TEM-style projection of coronene cluster with experimental HR-TEM of soot aggregate

The ORIENT program [47] was used to calculate the intermolecular interaction energies for the model potentials and this was automated using a short python script so as to calculate the energies over a range of separations. Figure 8 and Figure 9 show the potential curves for all the configurations (graphite, slipped parallel, crossed and T-shape) for naphthalene and anthracene dimers.

Both potentials accurately model the equilibrium separation of the dimers for the stacked configurations chosen. However, the Lennard-Jones potential tends to over-predict the potential well depth and also gives a repulsive wall which is too ‘hard’, i.e. repulsion is over-predicted at short dimer separations. In contrast the W99 force field compares well to the SAPT(DFT) calculations for all stacked configurations.

Both potentials fail to accurately predict the total interaction energy for the T-shaped configuration where the equilibrium separation is typically over-predicted by 0.3-0.4 Å and
the well depth is under-predicted by 5-10 kJ mol$^{-1}$. This is a significant error and is due to the potentials being isotropic in nature and having been parameterised with experimental crystallographic data. For PAH molecules stacked configurations are energetically more favourable than other possible configurations and hence are observed in crystal structures. As the W99 potential has been parameterised with such data [2], it is unsurprising that it models these configurations well, although to see such good comparison with the SAPT(DFT) results remains impressive. As both potentials are isotropic they do not take account of higher electron density associated with the $\pi$-bonding above and below the molecular plane in the stacked configurations. Being fitted to such configurations merely means they expect the same electron density to be present around each atom in all orientations and so they over-predict repulsion for configurations where the electron density between the molecules is in reality lower, such as the T-shape.

When molecules exhibit such anisotropy in terms of electron density around the constituent atoms, an isotropic potential cannot be made to accurately predict interaction energies in all dimer configurations. The correct prediction of intermolecular energies and separations for PAH molecules in non-stacked configurations is important when seeking to study the process of soot formation as interacting molecules will not always be stacked relative to one another. These observations have prompted the development of an anisotropic benzene potential.

### 4.3 Building an Anisotropic Benzene Potential

The new anisotropic benzene potential has been built by fitting potential terms to \textit{ab initio} SAPT(DFT) calculations for which the CamCASP suite of programs [44] was used. Due to the ongoing nature of the development of this software, only relatively small molecules can be investigated currently. Whilst it is hoped eventually a general anisotropic PAH potential can
be developed, the current limitations have meant initial work has been focused on benzene.

4.3.1 Molecular Properties

In the first stage the Gaussian03 program [43] was used to calculate the molecular geometry of the benzene molecule using DFT with the B3LYP functional and the 6-31G* basis set. CamCASP was then used in conjunction with DALTON [45] to calculate the KS orbitals and molecular properties using the asymptotically corrected PBE0 functional in the Sadlej-pVTZ MC basis. An additional required input to this calculation was the vertical ionisation energy of benzene which was taken as 891.9 kJ mol$^{-1}$ [52].

As an output from the molecular properties calculation a distributed multipole model was obtained in which carbon atoms were given to rank 4 (hexadecapoles) and hydrogen atoms were given to rank 1 (dipoles). The multipole model provided the means of representing the electrostatic term in the new potential. As an alternative electrostatic model (Eqn. 19), point charges were calculated using Gaussian03 with the Merz-Singh-Kollman (ESP) scheme.

Figure 8: Naphthalene dimer potentials

(a) Graphite

(b) Slipped Parallel

(c) Crossed

(d) T-shape
4 RESULTS AND DISCUSSION

Figure 9: Anthracene dimer potentials

(a) Graphite

(b) Slipped Parallel

(c) Crossed

(d) T-shape

using DFT with the PBE0 functional and Sadlej-pVTZ basis set.

4.3.2 Dispersion Model

Also from the initial molecular properties calculations using CamCASP, distributed frequency-dependent polarisabilities were obtained. From these CamCASP was used to implement the Williams-Stone-Misquitta (WSM) method [39] and a number of dispersion models were calculated, the simplest of which was the isotropic $C_6$ model, similar to that used to model dispersion in the W99 potential. In contrast to the W99 potential the dispersion models include damping using the Tang and Toennies damping function (Eqn. 17) with a $\beta$ value, calculated from the vertical ionisation potential of benzene (Eqn. 18), of 3.1152 Å$^{-1}$.

One hundred random benzene dimer geometries were generated for which dispersion energies were calculated using each dispersion model. Figure 10 shows the results of these calculations with model energies being plotted against SAPT(DFT) energies. The SAPT(DFT) energies were calculated in the MC+ Sadlej-pVTZ basis set and represent the standard to
which the model dispersion energies are compared. The dispersion models presented are anisotropic unless given the suffix ‘iso’, in which case they are isotropic.

The general trend seen in the dispersion models is that as higher terms in the dispersion energy series (Eqn. 15) are added, energies become more accurate. Adding anisotropy to the dispersion model also tended to produce more accurate results with the $C_{12}$ anisotropic model most closely matching the SAPT(DFT) results. Fortunately the deviation from the SAPT(DFT) results is approximately linear for all the dispersion models, and hence it becomes possible to introduce a scaling factor by which the simplest $C_{6,iso}$ model dispersion energies can be scaled to recover more accurately the SAPT(DFT) energies. In order to find the scaling coefficients a function of the following form was minimised:

$$
\Delta = \sum_i w(i) \left[ E_{\text{disp,tot}}(i) + \sum_{a \in A, b \in B} \xi_{ab} f_6(R_{ab}) C_{6,iso}^{ab} \right]^2,
$$

where $i$ labels the configurations, $E_{\text{disp,tot}}$ is the total dispersion energy given by SAPT(DFT) (the sum of the second-order dispersion, $E_{\text{disp,2}}$ and exchange-dispersion energies), $E_{\text{ex-disp}}^{(2)}$. The coefficients $\xi_{ab}$ are determined by a least-squares fit and $w$ is a weight, which will generally be energy dependent. In the general case the scaling coefficient would depend on the atom pairs, but in this work the simplest possible fit has been considered: all configurations
are weighted equally and a single constant of proportionality is used. The appropriate scaling factor for the damped isotropic $C_6$ model was found to be 1.372 for the range of dispersion energies greater than -20 kJ mol$^{-1}$. This energy range covers the most likely values of benzene dimer dispersion energies.

4.3.3 Short-Range Energies

A Born-Mayer term (Eqn. 14) was parameterised to fit the short-range SAPT(DFT) energies and to accommodate anisotropy by allowing an orientational-dependent shape parameter, $\rho$. The short-range energies were taken to be the first-order exchange energy, $E_{\text{exch}}^{(1)}$, and the first-order penetration, $E_{\text{pen}}^{(1)}$, defined as the difference between the SAPT(DFT) calculated electrostatic energy, $E_{\text{elst}}^{(1)}$, and the asymptotic (long-range) electrostatic energy given by the multipole or point charge model, $E_{\text{elst,asym}}$. A decision was made to neglect the induction energy at long-range (as this is small for PAH molecules) and include short-range induction contributions in the Born-Mayer term.

In a first attempt to get parameters for the Born-Mayer potential 500 random benzene geometries were chosen and SAPT(DFT) calculations in the monomer-centred (MC) basis were performed to find, $E_{\text{exch}}^{(1)}$, $E_{\text{pen}}^{(1)}$ and the overlap, $S$, which is the total density overlap of the two monomers. The ‘overlap model’ was then used to parameterise the Born-Mayer potential. This model assumes that the first-order energies, $E^{(1)}$, comprising the first-order exchange and penetration, are directly proportional to the overlap of the charge distributions.

$$E^{(1)} \equiv e \approx K_0 S,$$

(28)
Here $K_0$ is the scaling parameter. The quantity minimised is

$$\chi^2 = \sum_p w(e_p) \left[ \frac{K_0 S_p}{e_p} - 1 \right]^2 ,$$

(29)

where $w(e)$ is a weight that depends on the energy $e$. The simplest weighting scheme is to have $w(e) = 1$, however in practice a scheme is required which biases towards a known important energy, $e_0$, in order that outliers do not overly influence the fitting procedure. The following weighting scheme was employed for this fitting procedure:

$$w(e) = \exp \left[ -\alpha (\ln(e/e_0))^2 \right]$$

(30)

The following parameters were used: $e_0 = 20.0$ kJ mol$^{-1}$ and $\alpha = 1/\ln 10$.

An improvement is made by performing a second stage of fitting. In the second stage the total density overlap is distributed such that it is split into contributions from all atom pairs and individual atom pair scaling coefficients can be obtained, i.e.,

$$S = \sum_{a \in A, b \in B} S^{ab} ,$$

(31)

Here $S^{ab}$ is the density overlap associated with individual atoms, rather than molecules as a whole. The splitting takes place by using a density fitting algorithm which splits each molecular density into a sum of densities each representing individual atoms within each molecule. Having distributed the densities, the overall first-order energies for given dimer
geometries are given by,
\[ e = \sum_{a \in A, b \in B} K_{ab} S_{ab}, \]  
where a set of scaling coefficients, \( K_{ab} \), are found. These coefficients are permitted to have different values depending on the atom types involved. However in order to prevent unphysical values they are constrained by imposing a penalty function that suppresses values that are too different from \( K_0 \). Thus, the quantity minimised is now
\[ \chi^2 = \sum_p w(e_p) + \sum_{ab} \lambda (K_{ab} - K_0)^2, \]  
with
\[ \chi^2 = \sum_p w(e_p) \left[ \sum_{a \in A, b \in B} K_{ab} S_{ab} - e_p \right]^2. \]  
Typically small values of \( \lambda \) are chosen, and in this case \( \lambda = 10^{-8} \), thus ensuring the values of \( K_{ab} \) are only mildly constrained.

Having the atom pair scaling coefficients, the energy associated which each atom pair based on the atom pair overlap can be found and from this the simple Born-Mayer expression can be used to fit the energy contribution to give a separation dependent term to model short-range repulsive energies in the model potential.

\[ e_{ab} = K_{ab} S_{ab} = K \exp[-\alpha_{ab} (R_{ab} - \rho_{ab} (\Omega_{ab}))]. \]  
Here the pre-exponential factor, \( K \), was set to 1 milliHartree (2.6255 kJ mol\(^{-1}\)).

This whole fitting procedure was carried out in the CamCASP and ORIENT programs. Having fitted the first-order exchange and penetration to the Born-Mayer term, a further stage of fitting was undertaken using ORIENT. CamCASP was used to generate 100 random benzene geometries and to calculate the corresponding SAPT(DFT) induction energies in the Sadlej-pVTZ MC+ basis.

In order to account for short-range induction energies the previously calculated Born-Mayer parameters were relaxed within certain constraints and fitted to equally weighted energies comprising \( E^{(1)} \) and \( E_{\text{ind, tot}} \) (this includes the second-order induction, \( E^{(2)}_{\text{ind}} \), and exchange-induction \( E^{(2)}_{\text{ex, ind}} \) energies).

\[ \text{4.3.4 The Overall Potential} \]

Finally the procedure was repeated by combining the Born-Mayer term and the \( C_6 \) damped isotropic dispersion model and relaxing all parameters within contraints to fit equally weighted SAPT(DFT) energies for \( E^{(1)} \), \( E_{\text{ind, tot}} \) and \( E_{\text{disp, tot}} \) calculated for the 100 random benzene geometries and also 27 extra geometries relating to specific benzene dimer configurations
The final parameters for the benzene anisotropic potential, based on a point charge electrostatic model, are given in Table 3. The point charges in a.u. used for carbon and hydrogen atoms are -0.1111 and 0.1111 respectively. The root mean squared residual energy (i.e. $E_{\text{SAPT(DFT)}} - E_{\text{Potential}}$) given for the new potential when compared with the 127 SAPT(DFT) energies was found to be 0.6 kJ mol$^{-1}$. A plot of energies calculated with the new potential against the SAPT(DFT) energies calculated for the 100 random benzene dimer geometries is given Figure 12. The scatter in the results is likely to be due to the damped $C_6$ isotropic dispersion model and for which there cannot be further improvement without going to a more complex dispersion model. Figure 13 shows a comparison of the new benzene potential with SAPT(DFT) energies and the W99 potential for the chosen configurations. The new potential matches the SAPT(DFT) results in all configurations, especially the T-shaped configuration where the W99 potential is notably poor.

## 5 Conclusions

In this theoretical study the formation of soot has been investigated by considering the intermolecular interactions of PAH molecules. In the first part of the work a basin-hopping global optimization scheme was used to investigate the structure of PAH molecular clusters thought to be precursors of larger soot aggregates. This used a simple isotropic Lennard-Jones plus point charges potential found in the literature. Clusters of up to 50 coronene and pyrene molecules were studied and a TEM-style projection of a 50 molecule coronene cluster was made and compared to experimental TEM images of soot aggregates. Whilst

<table>
<thead>
<tr>
<th>Atom Pair</th>
<th>$l_a\kappa_a$</th>
<th>$l_b\kappa_b$</th>
<th>$\rho$ / Å</th>
<th>$\alpha$ / Å$^{-1}$</th>
<th>$C_6$ / kJ mol$^{-1}$ Å$^6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C</td>
<td>00 00</td>
<td>2.5165</td>
<td>3.3702</td>
<td>1753.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 10</td>
<td>0.0317</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 00</td>
<td>0.0317</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 20</td>
<td>-0.7905</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 00</td>
<td>-0.7905</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C H</td>
<td>00 00</td>
<td>2.7154</td>
<td>3.2413</td>
<td>734.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 10</td>
<td>-0.0305</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 00</td>
<td>-0.0169</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 20</td>
<td>-0.1234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 00</td>
<td>-0.1425</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>00 00</td>
<td>2.3504</td>
<td>2.8158</td>
<td>311.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 10</td>
<td>-0.1263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 00</td>
<td>-0.1263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 20</td>
<td>0.0999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 00</td>
<td>0.0999</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the PAH molecules in the soot aggregate appeared to be larger, similarities in structure were observed.

In the second part of the study a comparison was made between two isotropic atom-atom potentials, namely the Lennard-Jones plus point charges potential and the W99 force field, with SAPT(DFT) results for naphthalene and anthracene dimers reported by Podeszwa and Szalewicz [3]. Both potentials were found to correctly predict equilibrium dimer separations in the stacked configurations studied but the Lennard-Jones plus point charges potential tended to over-predict the well depth and also over-predicted the repulsive wall. The W99 potential predicted the potential curves for the stacked configurations remarkably accurately, but as with the Lennard-Jones plus point charges potential it tended to over-predict the potential for the T-shape configuration and also failed to give the correct equilibrium separation. This failing was found to be due to the isotropic nature of the potential and because it had been fitted to experimental data which biased towards stacked configurations.

In the final section a new benzene anisotropic potential was built using SAPT(DFT) energy calculations. The potential was found to have a r.m.s. residual of 0.6 kJ mol\(^{-1}\) when compared with SAPT(DFT) energies calculated for 127 dimer geometries. When compared with the W99 potential for specific configurations the new potential was found to correctly predict all configurations, unlike the W99 force field which only matched SAPT(DFT) energies for the stacked configurations.

6 Future Work

The general aims of the research are to investigate the formation and growth of soot particles. It is hoped that this research will give greater physical insights into mechanisms involved in these processes and also provide more detailed models for the group’s current soot model [53–55]. There are important questions which it is hoped this study will help to answer:-

- What are the mechanisms for soot particle inception?
- How ‘easy’ is it for small gas phase species present in flames to penetrate into soot particles?
- How ‘liquid-like’ are soot particles in flames and to what extent will primary particles sinter when forming an aggregate?

The current soot model uses a primitive inception model in which two pyrene molecules bind to form a ‘primary soot particle’. This assumption is still a matter of debate, and it is hoped that by studying the intermolecular forces of PAH molecules a more detailed understanding of soot particle inception can be established.

The ease of transport of small molecules (such as hydrogen, oxygen and acetylene) through a particle is important when seeking to establish the growth and oxidation rates
of soot particles. The current model assumes all PAH molecules within a soot particle are equally likely to undergo chemical reactions with the gas phase. This is simplistic and a detailed study of the potential energy surface (PES) of PAH clusters is required to establish the ease with which molecules can move through the cluster. The ability of PAH molecules to rearrange in a particle can also be studied by examining barrier heights between local minima on the PES.

To be able to examine the PES of primary soot particles it becomes necessary to look at molecular clusters containing many PAH molecules. At 50 molecules (the size of coronene and pyrene clusters studied in this work), the clusters are 3-4 nm in diameter which is comparable to primary soot particles. However these simulations, usually set for 1000 ‘basin-hopping’ iterations, took many days to run and simulations with larger clusters are not feasible at present. A possible route around this problem is to coarse-grain the potential to a molecule-molecule potential allowing larger molecular systems to be studied.

The planned future work can be divided into short and long term projects. In the short term, work will continue in the development of a general anisotropic PAH atom-atom potential. This will be used to model small clusters, but also to act as an accurate validation for coarse-grained PAH potentials. It is hoped that the benzene potential already developed can be simply altered to apply to large PAH molecules. If this cannot be achieved then new anisotropic potentials for larger PAH molecules, such as naphthalene, anthracene and pyrene, will be required. This will allow specification of different carbon atom environments (other than the one specified in the current benzene potential). It is then hoped that a common set of parameters can be found which give accurate atom-atom potentials for all PAH molecules with a small number of different carbon environments. The new potentials will be used to study PAH clusters using GMIN. This will allow a systematic study of the potential energy surfaces of both small single-type molecular clusters and small ensembles of different PAH molecules.

A longer term aim is to develop a coarse-grained potential for dealing with intermolecular interactions of many PAH molecules. Currently larger clusters using an atom-atom potential require much longer computational time as the number of atom-atom potential calculations, \( N_{a-a} \) performed during each iteration is given by

\[
N_{a-a} = \frac{k(k-1)n^2}{2}
\]

where \( k \) is the number of molecules and \( n \) is the number of atoms in each molecule.

Reducing the intermolecular potential from atom-atom interactions to molecule-molecule interactions would significantly reduce the number of calculations required for large clusters. There already exist some coarse-grained potentials in the literature which have even been parameterised for some PAH molecules [56]. These potentials will act as the starting point for further coarse-grained work with PAH molecules.
It is unclear what kind of coarse-grained potential will be required to be able to form a general PAH potential. Most current coarse-grained potentials choose one potential for the whole molecule and parameterise it accordingly. However PAH molecules come in many different shapes and sizes and it will not be possible to produce a new potential for every type of PAH molecule expected in flame environments. A possible solution to this problem is to break PAH molecules up into blocks (such as six membered rings) and produce a coarse grain potential for each block which can then be combined in arbitrary ways to make up the shape of larger molecules. This will add computational cost to the potential as in general more than one site will be considered in each molecule but this approach will provide an easy means to generalise to all PAH molecules and will allow a closer match to accurate atom-atom potentials already developed.

Once coarse-grained potentials have been developed and larger PAH systems can be studied, there will be opportunity to answer some of the current questions associated with soot formation and growth. Many of the questions will be answered by developing a detailed understanding of the PES. It may be found that the GMIN code, even with coarse-grained potentials, will not be able to handle large clusters of molecules in reasonable computation time. In this case other methods will be developed such as a Monte-Carlo algorithm whereby many random molecular configurations are sampled and energies calculated, and thereby building up a picture of the PES. This approach differs from the GMIN code by avoiding searching for a local minimum in each step, and saves a lot of computation time.

In order to study the dynamics of the rearrangement of molecules when two primary particles aggregate, it is hoped that the coarse-grained potentials can be implemented in a molecular dynamics code. If this turns out to be too computationally intensive an alternative approach may be to fix an inner core of PAH molecules to create a ‘particle’ and arbitrarily move molecules around the particle surface to investigate how the potential energy of the system varies. This approach would locate surface minima and would also give an indication of relative barrier heights for rearrangement around the surface.
## 7 Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_i$</td>
<td>Basis function coefficient of $i$th spin orbital</td>
</tr>
<tr>
<td>$e_i$</td>
<td>Charge of particle $i$</td>
</tr>
<tr>
<td>$f_i$</td>
<td>Dispersion damping function for dispersion model level $i$</td>
</tr>
<tr>
<td>$\hat{f}$</td>
<td>One-electron Fock operator</td>
</tr>
<tr>
<td>$r_{ij}$</td>
<td>Separation between charged particles $i$ and $j$</td>
</tr>
<tr>
<td>$q_i$</td>
<td>Point charge associated with atom centre $i$</td>
</tr>
<tr>
<td>$w$</td>
<td>Energy weight</td>
</tr>
<tr>
<td>$C_i$</td>
<td>Dispersion coefficient for dispersion model level $i$</td>
</tr>
<tr>
<td>$C_{ij}$</td>
<td>Renormalised spherical harmonic corresponding to parameters $i$ and $j$</td>
</tr>
<tr>
<td>$E_i$</td>
<td>Energy corresponding to $i$</td>
</tr>
<tr>
<td>$\hat{E}$</td>
<td>Basin-hopping transformed energy</td>
</tr>
<tr>
<td>$\hat{H}$</td>
<td>Molecular Hamilton operator</td>
</tr>
<tr>
<td>$I_i$</td>
<td>Vertical ionisation energy for molecule $i$</td>
</tr>
<tr>
<td>$J$</td>
<td>Classical electrostatic electron-electron repulsion energy</td>
</tr>
<tr>
<td>$K$</td>
<td>Scaling coefficient in overlap model and Born-Mayer pre-exponential factor</td>
</tr>
<tr>
<td>$R_{ij}$</td>
<td>Separation between atomic centres $i$ and $j$</td>
</tr>
<tr>
<td>$S$</td>
<td>Density overlap</td>
</tr>
<tr>
<td>$\hat{T}$</td>
<td>Molecular kinetic energy operator</td>
</tr>
<tr>
<td>$U_i$</td>
<td>Intermolecular energy associated with $i$</td>
</tr>
<tr>
<td>$\hat{V}$</td>
<td>Molecular and Intermolecular potential energy operator</td>
</tr>
<tr>
<td>$W_i$</td>
<td>Energy of molecule $i$</td>
</tr>
<tr>
<td>$Z$</td>
<td>Nuclear charge</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Born-Mayer ‘hardness’ coefficient</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Tang and Toennies damping function parameter</td>
</tr>
<tr>
<td>$\epsilon_i$</td>
<td>One-electron spin orbital energy of $i$th state</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Basis function</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Zenith angle</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Perturbation expansion parameter</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Scaling parameter for $C_6$ dispersion model</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Born-Mayer shape parameter</td>
</tr>
<tr>
<td>$\rho(\vec{r})$</td>
<td>Electron density at position $\vec{r}$</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Azimuth angle</td>
</tr>
<tr>
<td>$\varphi_i$</td>
<td>One-electron spin orbital of $i$th state</td>
</tr>
<tr>
<td>$\Psi_i$</td>
<td>Molecular wavefunction of $i$th state</td>
</tr>
</tbody>
</table>
REFERENCES

References


REFERENCES


REFERENCES


