JCTC Journal of Chemical Theory and Computation

A First Principles Development of a General Anisotropic Potential for Polycyclic Aromatic Hydrocarbons

Tim S. Totton,[†] Alston J. Misquitta,*[,‡] and Markus Kraft[†]

Department of Chemical Engineering and Biotechnology, University of Cambridge, New Museums Site, Pembroke Street, Cambridge CB2 3RA, United Kingdom, and Department of Physics, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge, CB3 0HE, United Kingdom

Received September 15, 2009

Abstract: Standard empirical atom−atom potentials are shown to be unable to describe the binding of polycyclic aromatic hydrocarbon (PAH) molecules in the variety of configurations seen in clusters. The main reason for this inadequacy is the lack of anisotropy in these potentials. We have constructed an anisotropic atom−atom intermolecular potential for the benzene molecule from first principles using a symmetry-adapted perturbation theory based on density functional theory (SAPT(DFT)), interaction energy calculations and the Williams−Stone−Misquitta method for obtaining molecular properties in distributed form. Using this potential as a starting point, we have constructed a transferable anisotropic potential to model intermolecular interactions between PAHs. This new potential has been shown to accurately model interaction energies for a variety of dimer configurations for four different PAH molecules, including certain configurations which are poorly modeled with current isotropic potentials. It is intended that this potential will form the basis for further work on the aggregation of PAHs.

1. Introduction

Polycyclic aromatic hydrocarbon (PAH) molecules have often been invoked as intermediates in the chemistry of soot formation and growth.[1] The presence of stacked PAH molecular structures in experimental high-resolution transmission electron microscopy (HRTEM) images of soot particles[2−4] 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 the relative orientation of PAH molecules present in dimers and larger stacks in flame environments.[1,5−9] Currently, many numerical simulations of soot formation in flames consider the dimerization of molecules as small as pyrene ($C_{16}H_{10}$)[9,10] to be the particle inception step, however, the validity of this assumption is still debated.[11]

The aggregation of PAH molecules has traditionally been modeled using atom−atom potentials, which approximate the total interaction energy, U, as sum over all pairwise atomic interactions between molecules:

$$U = \sum_{\mathrm{A}} \sum_{\mathrm{A<B}} \sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{B}} U_{ab}(R_{ab}, \Omega_{ab}) \qquad (1)$$

Here $U_{ab}(R_{ab},\Omega_{ab})$ denotes an atom−atom interaction potential. The indices A and B are for molecules, and the indices a and b run over all the atomic sites within these molecules. In general, the interaction potential depends upon the atom−atom separation, R_{ab}, and the relative molecular orientation, described in some way by Ω_{ab}. Often, however, orientational dependence is removed as a simplification, and such potentials are 'isotropic', i.e., the atoms in a molecule are considered to be spherically symmetric. Previous studies into the intermolecular chemistry of PAHs have been largely based on computationally convenient model potentials, such as isotropic Lennard-Jones 12-6 (eq 2) and exp-6 potentials (eq 3).[1,12,13] Explicit electrostatic models are often added to these forms, the simplest being based upon partial atom-centered point charges, q_a and q_b (eq 4).

* Corresponding author. E-mail: am592@cam.ac.uk.

[†] Department of Chemical Engineering and Biotechnology, University of Cambridge.

[‡] Department of Physics, University of Cambridge.

10.1021/ct9004883
Published on Web 02/12/2010

$$U_{ab}^{\mathrm{LJ}} = 4\varepsilon_{ab}\left[\left(\frac{\sigma_{ab}}{R_{ab}}\right)^{12} - \left(\frac{\sigma_{ab}}{R_{ab}}\right)^{6}\right] \quad (2)$$

$$U_{ab}^{\exp-6} = B_{ab}\exp(-C_{ab}R_{ab}) - \frac{A_{ab}}{R_{ab}^{6}} \quad (3)$$

$$U_{ab}^{\mathrm{elst}} = \frac{q_a q_b}{R_{ab}} \quad (4)$$

In recent years, with the advance of computational power, the theoretical understanding of intermolecular interactions has developed significantly yet empirical potentials have remained largely unchanged. Current isotropic literature potentials have typically been parametrized to be applicable to a wide range of organic molecules.[14−17] Such potentials are typically parametrized with heats of sublimation and crystallographic data and, while transferable, are only accurate for the configurations they were parametrized for, and often fail at others. For example, consider two widely used potentials: the Williams W99 potential[15,16] based on the exp-6 form and the 12-6 Lennard-Jones potential,[14] both including point charges (the latter parametrized from an earlier form of the Williams potential). The performance of these isotropic potentials has been examined for naphthalene and anthracene dimer orientations shown in Figures 1 and 2. In Figures 3 and 4, we show cross sections of the potential energy surface at these orientations, where separation, R, is between the centers of mass of the monomers. The reference energies are taken from ab initio SAPT(DFT) calculations performed by Podeszwa and Szalewicz.[18]

The W99 potential performs remarkably well for stacked PAH geometries, while the Lennard-Jones plus point charges potential tends to overestimate well depths by 5−10 kJ mol^{-1}. However, both potentials show substantial errors for the T-shape dimer with an error in equilibrium separation of 0.3−0.4 Å and with an underbinding of as much as 7 kJ mol^{-1} in the case of the W99 potential.

Being isotropic, the potentials cannot accurately model the atom−atom interactions where there is significant anisotropy in the electron distribution around constituent atoms, such as in PAHs where there is significant π-bonding. These potentials also suffer from being required to possess too large a degree of transferability to make them sufficiently accurate for the specific system of interest, and development of anisotropic potentials empirically is precluded due to insufficient experimental data.

To accurately model dimers in all orientations either accurate ab initio methods must be used directly (on-the-fly methods) or new anisotropic atom−atom potentials are required, parametrized using ab initio results. Currently, most ab initio methods are prohibited due to high computational expense. Density functional theory (DFT) is the only method which is computationally feasible, but currently there are no practical and quantitative functionals which correctly predict intermolecular dispersion energies. In reality, the size of PAH systems (which can consist of many hundreds of atoms) and the complexity of the calculations restrict us to using model atom−atom potentials. In the context of PAHs and investigating soot structure at a molecular level, there are a number of requirements for a potential:

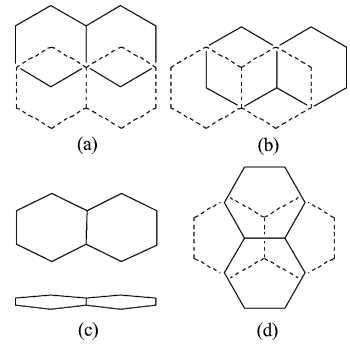

Figure 1. Naphthalene: (a) Slipped-parallel, symmetry C_{2h}; (b) Graphite-type, symmetry C_i; (c) T-shape, symmetry C_{2v}; (d) Crossed, symmetry D_{2d}.

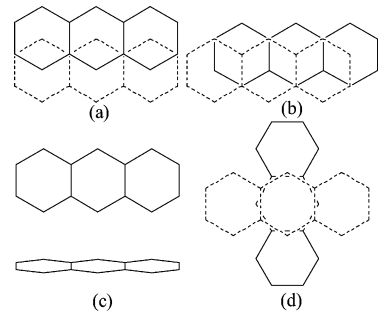

Figure 2. Anthracene: (a) Slipped-parallel, symmetry C_{2h}; (b) Graphite-type, symmetry C_i; (c) T-shape, symmetry C_{2v}; (d) Crossed, symmetry D_{2d}.

• Accuracy: The potential is expected to be accurate for the variety of dimer configurations which are expected to be sampled in a flame environment. In particular, the potential must correctly predict barriers on the potential energy surface (PES) of the molecular cluster.

• Transferability: In a flame environment, typically large ensembles of different PAH molecules exist of varying size (C_6−C_{400}),[20] and consequently, it is very important that any potential developed can be easily transferable to different PAHs.

• Simplicity: Large PAH clusters need to be studied requiring extensive calculations. This will limit the functional complexity of the potential expression used to model interactions.

In the context of PAHs, there have been several studies using ab initio methods. For example, coupled cluster calculations at CCSD(T) level have been used to study naphthalene dimers,[21] while MP2 level calculations have been used to obtain dimer interaction energies for various PAHs.[10] However, Møller−Plesset perturbation theory is inadequate to study intermolecular interactions between systems with a significant amount of π-bonding (such as PAH clusters). Compared to the more reliable CCSD(T) calculations, MP2 calculations have been shown to considerably overestimate attraction between molecules, in some cases by almost a factor of 2,[21−23] throwing into doubt some earlier studies of soot particle inception.[10] However, CCSD(T) is computationally demanding and is not ideally suited for potential development due to the inability to decompose the overall interaction energy into physically significant contributions. This makes it hard to parametrize

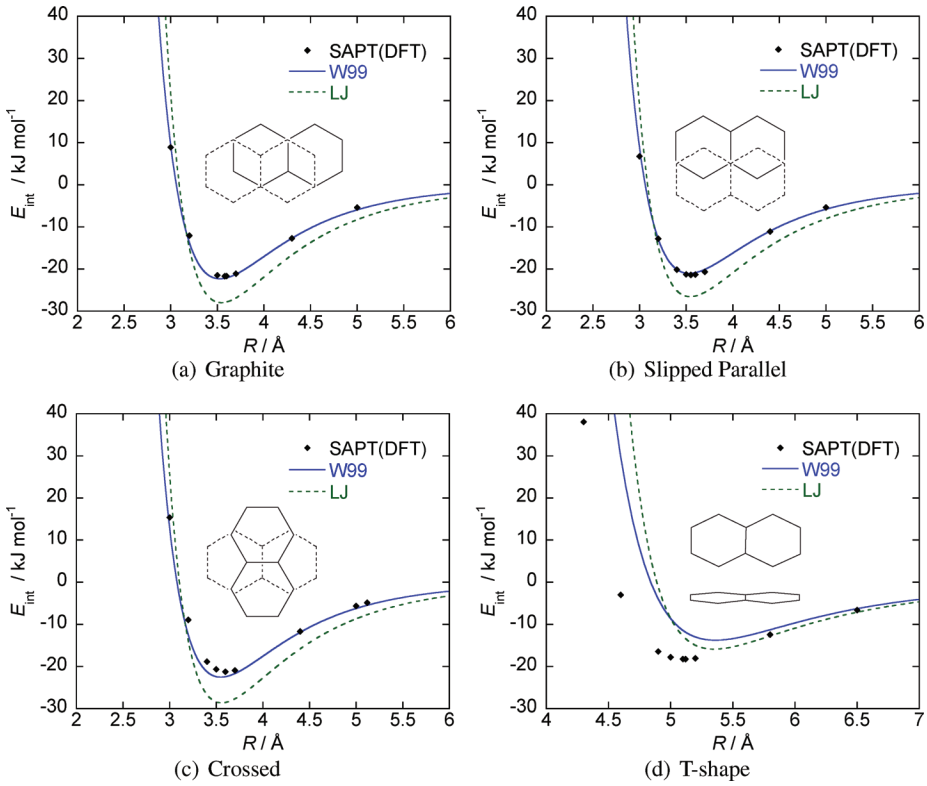

Figure 3. Comparison of isotropic model potentials with SAPT(DFT) energies for different naphthalene dimers. A key to the geometries is given in Figure 1. Model potential energies have been calculated using the Orient[19] program.

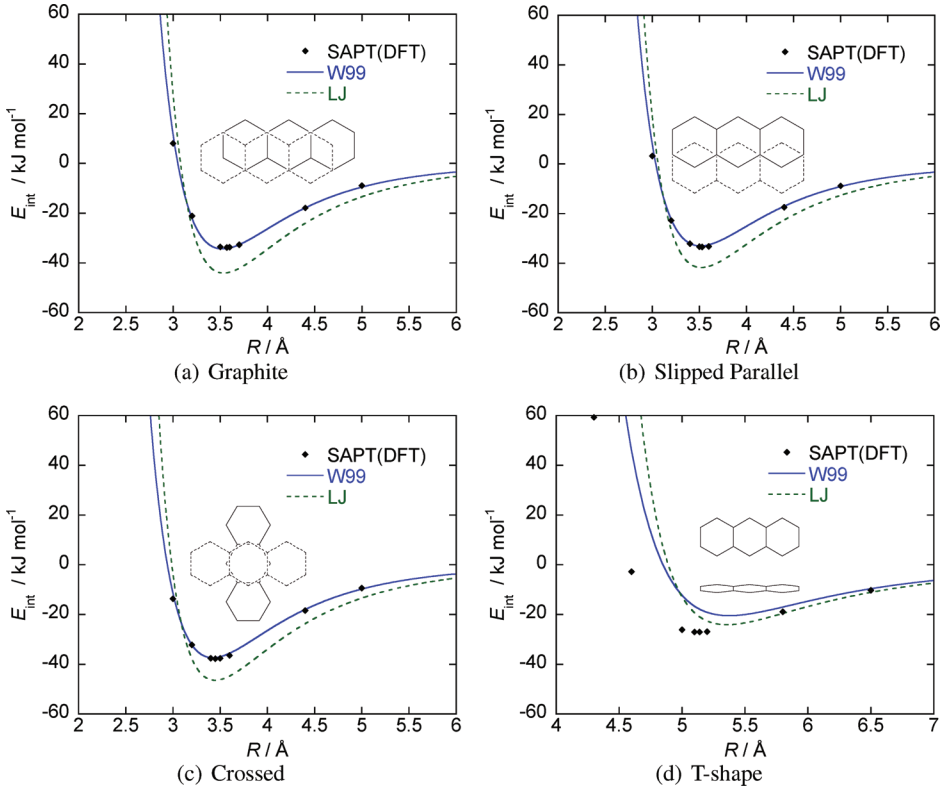

Figure 4. Comparison of isotropic model potentials with SAPT(DFT) energies for different anthracene dimers. A key to the geometries is given in Figure 2. Model potential energies have been calculated using the Orient[19] program.

analytic potentials comprised of multiple terms, where each term describes a different interaction, such as the dispersion or repulsion.

By contrast, intermolecular perturbation theory provides an ideal framework for the development of model potentials because it provides the interaction energy as a sum of

physically significant contributions. This allows the separate parametrization of different terms representing different interactions within multi-term model potentials. The development of symmetry-adapted perturbation theories has enabled both long- and short-range interactions to be accurately calculated, and the recent development of SAPT(DFT)[24−31] has made possible highly accurate studies of intermolecular interactions at a level comparable to CCSD(T),[32,33] with modest computational resources.

This methodology has already been used to develop intermolecular potentials. Misquitta et al.[34] have developed an anisotropic potential to predict the crystal structure of the 1,3-dibromo-2-chloro-5-fluorobenzene ($C_6BrClFH_2$) molecule, giving results in excellent agreement with experiment. Similarly, a potential derived from SAPT(DFT) interaction energies has been used to study the potential energy surface of the cyclotrimethylene trinitramine (RDX) dimer.[35] A benzene potential has been constructed using SAPT(DFT) energy calculations of 491 dimer geometries.[33] However, in addition to the usual atomic sites, this potential also contains off-atomic sites, and it is difficult to see how the parameters for off-atomic sites can be transferred easily to larger PAHs.

The potential form we have chosen to use is

$$U_{ab} = \underbrace{G\exp\left[-\alpha_{ab}\left(R_{ab}-\rho_{ab}(\Omega_{ab})\right)\right]}_{short-range} \underbrace{-f_6(R_{ab})\frac{C_{6,\mathrm{iso}}}{R_{ab}^6}+E_{\mathrm{elst}}(\mathrm{model})}_{long-range}, \qquad (5)$$

where the first term is a Born−Mayer term describing short-range interactions, the second is an isotropic, damped dispersion term, and the third term is an appropriate electrostatic model. These terms will be described in detail later. This form of the potential remedies two of the major deficiencies of the exp-6 potential in eq 3: (i) the short-range term now includes a shape-function, ρ_{ab}, that models the anisotropy of the interacting sites through a dependence on the relative orientation, Ω_{ab}, of the two sites, and (ii) the singularity in the dispersion term is removed by the damping function, $f_6(R_{ab})$ (defined later). A more elaborate potential functional form could have been chosen, for example, we could have included higher order terms and anisotropy in the dispersion model and the anisotropy in the damping function, but the resulting potential would be unnecessarily computationally demanding and would not be usable in most simulation programs without a significant degree of modification.

We begin this article with a description of the methods we have used to parametrize this potential for benzene. We fit energies and molecular properties to the best ab initio data provided by SAPT(DFT) and use the Williams−Stone−Misquitta (WSM) method[36−40] for determining distributed frequency-dependent polarizabilities needed for the dispersion model (Section 2). In order to keep the parameters physical, we have used a multistage fitting procedure to obtain the parameters for the short-range part of the potential (Section 3). The resulting parameter set then acts as a starting point for the generalization of the potential to larger PAHs (Section 4). In this stage, we have used SAPT(DFT) interaction energies calculated by Podeszwa and Szalewicz[18] for dimers of naphthalene ($C_{10}H_8$), anthracene ($C_{14}H_{10}$), and pyrene, ($C_{16}H_{10}$) in a variety of configurations. Finally, we conclude with an assessment of the accuracy of SAPT(DFT) energies and possible directions for future applications of the potential (Section 5).

2. Constructing the Benzene Dimer Potential

The basic strategy for constructing an analytic potential for molecules consisting of more than two atoms has been described in a recent review.[41] The potential is logically separated into long- and short-range parts (eq 5). The long-range part depends upon molecular properties, such as multipole moments, polarizabilities, and dispersion coefficients. Long-range polarization, or induction, is expected to be weak in molecules which do not possess strong multipole moments, such as PAHs. We have, therefore, omitted an explicit induction term in our model potentials. The short-range energies include the exchange-repulsion, the penetration energies (see below), and the second-order induction effects (which make a small but significant contribution). These energies all decay exponentially with increasing separation. We have fitted the parameters of the exponential terms via the density overlap model, using the procedure described in ref 34 and outlined below.

2.1. Molecular Geometry and Basis Sets. The geometry of benzene was obtained by in vacuo optimization using DFT with the B3LYP functional and the 6-31G* basis set with the Gaussian03[42] program. The molecule was assumed to be rigid. Calculated atomic coordinates are provided in the Supporting Information. The D_{6h} symmetry of benzene allows us to identify just two unique atom types: a carbon and a hydrogen. Symmetry was imposed during the calculation of the distributed properties and the subsequent fitting process for the Born−Mayer parameters.

Interaction energies and molecular properties have been calculated using the CamCASP[43] program from molecular wave functions obtained using the Dalton[44] program. The molecular wave functions were calculated with the asymptotically corrected PBE0[45] exchange−correlation functional and the Sadlej-pVTZ[46] basis. We have used the Tozer−Handy asymptotic correction[47,48] with a vertical ionization potential of 0.3397 a.u., obtained from ref 49. The linear-response DFT calculations needed for second-order SAPT(DFT) energies were performed using a hybrid adiabatic local density approximation (LDA) and coupled Hartree−Fock kernel.[27,38]

All calculations have used the Sadlej-pVTZ basis[46] set in two basis types: (i) a 'monomer-centered' (MC) basis, which includes basis functions on atomic sites only, and (ii) a 'monomer-centered-plus' (MC+) basis type, which additionally includes basis functions placed in the bonding region between the two molecules and on the atomic sites of the partner molecule. The MC type of basis was used for molecular properties and for the first-order interaction energies (the electrostatic and exchange-repulsion energies) and density-overlaps. This type of basis is not suitable for calculations of the second-order interaction energies which are slow to converge with basis set.[50] For these, we have used the MC+ type of basis with a 3s2p1d basis set for the mid-bond functions, placed at a position determined by a generalization of the weighting scheme described in ref 51.

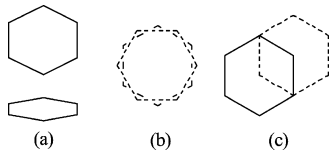

Figure 5. Benzene: (a) T-shape, symmetry C_{2v}; (b) Crossed, symmetry D_{6d}; (c) Slipped-parallel, symmetry C_{2v}.

The CamCASP program uses density fitting in the evaluation of interaction energies, molecular properties, and density overlaps. We used two kinds of auxiliary basis sets in our calculations: (i) the aug-cc-pVTZ auxiliary basis[52] has been used for the calculations of molecular properties and SAPT(DFT) energies with the MC+ basis type, and (ii) the smaller JK-TZVPP auxiliary basis[52] has been used for the calculation of the density overlap and the first-order SAPT(DFT) energies used in the first part of the fitting process.

2.2. SAPT(DFT) Dimer Energies. SAPT(DFT) interaction energies for the benzene dimer were calculated at a variety of dimer configurations so as to model the exchange−repulsion, penetration, and induction energies as well as to provide a set of reference dispersion energies for the assessment of our dispersion models. Based upon an earlier study,[27,34] we used the following formulation of the SAPT(DFT) interaction energy:

$$E_{\mathrm{int}}^{(2)} = E_{\mathrm{elst}}^{(1)}(\mathrm{KS}) + E_{\mathrm{exch}}^{(1)}(\mathrm{KS}) + E_{\mathrm{ind,tot}}^{(2)} + E_{\mathrm{disp,tot}}^{(2)} \quad (6)$$

where, $E_{\mathrm{elst}}^{(1)}(\mathrm{KS})$ and $E_{\mathrm{exch}}^{(1)}(\mathrm{KS})$ are the first-order electrostatic and exchange−repulsion energies, $E_{\mathrm{ind,tot}}^{(2)}$ is the total induction energy defined as sum of the polarization and exchange contributions,[27] $E_{\mathrm{ind,pol}}^{(2)} + E_{\mathrm{ind,exch}}^{(2)}$, and likewise, $E_{\mathrm{disp,tot}}^{(2)}$ is the total dispersion energy defined as $E_{\mathrm{disp,pol}}^{(2)} + E_{\mathrm{disp,exch}}^{(2)}$. The second-order terms are calculated using Kohn−Sham linear response theory. Terms of third- and higher-order in the interaction operator have been omitted, as these are not expected to be significant for systems without hydrogen bonds.[33,38]

It is important to select the dimers so as to get a uniform coverage of the space of physically important configurations. We have done this by keeping one of the molecules of the dimer fixed and centered at the origin and by translating and rotating the other using the following algorithm:[34]

• Using a Sobol quasi-random sequence, generate a random direction vector for the translation and, using Shoemake's uniform distribution algorithm,[53] generate a quaternion for rotation.

• Starting with both molecules centered at the origin, rotate one using the quaternion. Using the standard van der Waals radii,[54] determine the distance of van der Waals contact R_0 along the direction vector.

• Translate the rotated molecule along the direction vector by a few (1−5) randomly selected distances chosen to lie between $R_0 - \Delta R_{\mathrm{min}}$ and $R_0 + \Delta R_{\mathrm{max}}$. We have used ΔR_{min} and ΔR_{max} to be 1.5 and 1.2 a.u., respectively.

In addition to 500 benzene dimer configurations selected, using the algorithm described above, we have used an additional 27 energies that were calculated at specific dimer orientations shown in Figure 5. Here the slipped-parallel and crossed configurations represent stacked dimer configurations, and the interaction energy has been calculated at a set of interplanar spacings. Likewise, the T-shaped configuration energies have been calculated at a set of separations of the monomer centers of mass.

2.3. Molecular Properties. Ideally, when the effect of electron density overlap is negligible, we would describe the electrostatic interaction energy using a distributed multipole model.[55,56] Such models have had a lot of success in organic crystal structure prediction.[57] However, for the benzene dimer, we have found that a simple point charge model suffices, and the higher ranking multipole moments are not so important. This is probably because of the absence of strong directional moments, such as those seen in hydrogen-bonding complexes. Additionally, there are few simulation programs that can use distributed multipole models. Consequently, we have used a distributed point charge model to describe the electrostatic interaction at long range. This model was calculated with the Gaussian03[42] program using the PBE0 functional and Sadlej-pVTZ basis using the Merz−Singh−Kollman scheme,[58] which fits the molecular electrostatic potential to a set of atom-centered point charges. In principal, the electrostatic term, which varies as $1/R_{ab}$ should be damped at short range to avoid the divergence as $R_{ab} \rightarrow 0$. In practice however, the low power of R_{ab} means that this divergence is manifest only at very small separations, and hence, damping can be ignored in many cases.

Frequency-dependent polarizabilities are needed to calculate the dispersion coefficients with which we model the second-order dispersion energies at intermolecular separations where orbital overlap effects can be neglected. These polarizabilities need to be distributed in order to ensure rapid convergence with rank of the multipole expansion; in fact, even for a molecule the size of benzene, the single center multipole expansion will not converge for the physically important dimer configurations. Distributed frequency-dependent polarizabilities of ranks one, two, and three for carbon and hydrogen atoms have been obtained using the WSM method. This distribution method has been shown to result in models which exhibit very good convergence properties, while resulting in a physically meaningful partitioning of the molecular properties.

2.4. Dispersion Models. In the multipole expansion, the second-order dispersion energy between two molecules, A and B, is given by[59]

$$E_{\mathrm{disp,d}}^{(2)} = -\sum_{a\in \mathrm{A}}\sum_{b\in \mathrm{B}}\left(f_6(R_{ab})\frac{C_6^{ab}}{R_{ab}^6} + f_7(R_{ab})\frac{C_7^{ab}}{R_{ab}^7} + f_8(R_{ab})\frac{C_8^{ab}}{R_{ab}^8} + \ldots\right) \quad (7)$$

where the C_n^{ab} are the dispersion coefficients which are orientationally dependent for non-spherically symmetric sites. For interactions between spherically symmetric sites, the C_n^{ab} coefficients contain no angular dependence and terms that are odd in n are zero. The $f_n(R_{ab})$ are damping functions,[59] which are needed to remove the divergence of the expansion at small R_{ab}. Here, we use the Tang−Toennies damping functions:[60]

$$f_n(R_{ab}) = 1 - \exp(-\beta R_{ab}) \sum_{k=0}^{n} \frac{(\beta R_{ab})^k}{k!} \quad (8)$$

with an isotropic-damping parameter of $\beta = 2(2I)^{1/2}$, where I is the vertical ionization energy in a.u.[40] For benzene, using the value of I presented above, we obtain $\beta = 1.6485$ a.u.

The WSM method allows the calculation of a variety of dispersion models from the simple isotropic C_6 model to the very elaborate anisotropic C_{12} model[40] (a C_n model includes all terms from C_6 to C_n). We have calculated isotropic and anisotropic C_6, C_{10}, and C_{12} models using the CamCASP[43] program. In Figure 6, we present dispersion energies from these models calculated with the Orient program and displayed against SAPT(DFT) total dispersion energies, $E^{(2)}_{\mathrm{disp,tot}}$.

2.4.1. Refining the Isotropic C_6 Dispersion Model. In practice, elaborate anisotropic dispersion models are computationally demanding, and few simulation programs are able to use them. We have, therefore, created an *effective* damped isotropic C_6 model in the manner described in ref 40.

The scatter plot shown in Figure 6 shows that deviation of the model dispersion energies from the SAPT(DFT) energies is approximately linear for all the dispersion models. It, thus, becomes possible to introduce a scaling factor by which the $C_{6,\,\mathrm{iso}}$ 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 minimized

$$\Lambda = \sum_i w_i \left[E^i_{\mathrm{disp,tot}} + \xi \sum_{a\in \mathrm{A}, b\in \mathrm{B}} \frac{f_6(R_{ab}) C^{ab}_{6,\mathrm{iso}}}{R^6_{ab}} \right]^2 \quad (9)$$

where i labels the configurations, and $E_{\mathrm{disp,tot}}$ is the total SAPT(DFT) dispersion energy. The coefficient, ξ, is determined by a least-squares fit, and w_i 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 physically significant dispersion energies defined by the range from -20 to 0 kJ mol^{-1}. The root-mean-square (rms) error for the damped and scaled $C_{6,\,\mathrm{iso}}$ model over this range is only 0.47 kJ mol^{-1}. At small separations when total dispersion energies are lower than -20 kJ mol^{-1}, the dispersion energy is overestimated, but at such short ranges, repulsive interactions are expected to dominate. The dispersion is also overestimated at large separations, although it should be noted that this is true of any *effective* C_6 representation of dispersion using only atomic sites.[40]

3. Short-Range Energies

The short-range energy is defined as the difference in the total interaction energy (eq 6) and the energies calculated with the multipole expansions for the electrostatics and the dispersion. As mentioned above, since the long-range induction is expected to be weak in systems of PAH molecules, we have not included a multipole expansion for the induction energy. Therefore, we define the short-range energy is as:

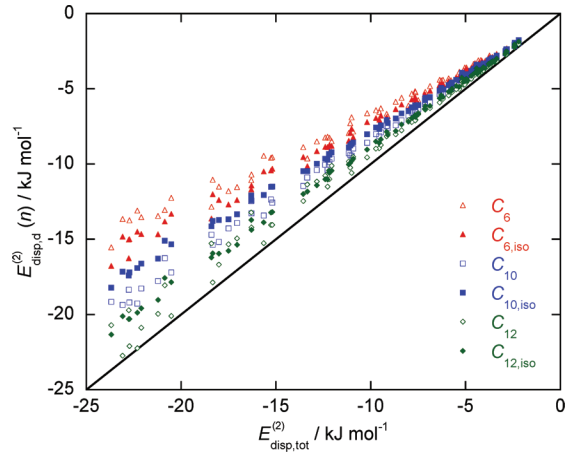

Figure 6. Dispersion energies for the benzene dimer. Scatter plot of dispersion energies calculated using the damped dispersion models represented by $E^{(2)}_{\mathrm{disp,d}}(n)$ against $E^{(2)}_{\mathrm{disp,tot}}$ calculated using SAPT(DFT). The dispersion models presented are anisotropic unless given the suffix 'iso', in which case they are isotropic.

$$E^{(2)}_{\mathrm{sr}} = (E^{(1)}_{\mathrm{elst}} - E^{(1)}_{\mathrm{elst}}(\mathrm{ESP})) + E^{(1)}_{\mathrm{exch}} + E^{(2)}_{\mathrm{ind,tot}} = E^{(1)}_{\mathrm{elst,pen}} + E^{(1)}_{\mathrm{exch}} + E^{(2)}_{\mathrm{ind,tot}} \quad (10)$$

where $E^{(1)}_{\mathrm{elst}}(\mathrm{ESP})$ is the electrostatic energy calculated using the point charge model, $E^{(2)}_{\mathrm{disp,d}}(C_{6,\,\mathrm{iso}})$ is the dispersion energy calculated using the *effective* damped $C_{6,\,\mathrm{iso}}$ model, and $E^{(1)}_{\mathrm{elst,pen}}$ is the electrostatic penetration energy defined implicitly above. The induction is included as a short-range energy because, in the absence of strong permanent multipoles, $E_{\mathrm{ind,tot}}$ is almost all due to orbital overlap effects. We could define a penetration-like contribution from the dispersion energy, $E^{(2)}_{\mathrm{disp,tot}} - E^{(2)}_{\mathrm{disp,d}}(C_{6,\,\mathrm{iso}})$, but this term is small and does not exhibit an exponential dependence with distance; consequently, we have omitted it here but have included it in the final stage of the fitting process.

Unlike the exchange−repulsion energy, $E^{(2)}_{\mathrm{sr}}$ is not always positive, as there will be configurations for which the penetration energies—which are generally negative—and the (negative) induction energies will be larger in magnitude than the corresponding exchange−repulsion energies. Nevertheless, this happens at very few dimer configurations, and when it does, $E^{(2)}_{\mathrm{sr}}$ tends to be small in magnitude. We have, therefore, omitted such configurations. Furthermore, since the penetration energy and exchange−repulsion energies both arise from the overlap of the molecular wave functions, they both exhibit an exponential dependence on intermolecular separation. If we assume the same distance dependence for all terms in eq 10, we can then fit the positive values of $E^{(2)}_{\mathrm{sr}}$ to the Born−Mayer term of eq 5:

$$\sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{B}} G\exp[-\alpha_{ab}(R_{ab} - \rho_{ab}(\Omega_{ab}))] \quad (11)$$

The hardness of the interaction is described by α_{ab}, and G is a constant energy unit taken to be 10^{-3} a.u. The shape

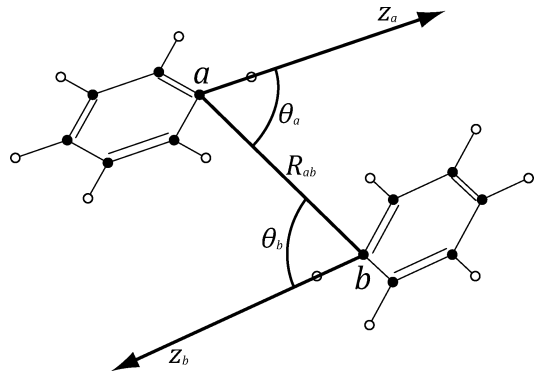

Figure 7. Schematic showing the axis system used to define short-range anisotropy. The local *z*-axes for both the carbon atoms and the attached hydrogen atoms point radially outward along the carbon−hydrogen bonds.

functions, $\rho_{ab}(\Omega_{ab})$, have been assumed to be the sum of the shape functions of the individual sites:[59]

$$\rho_{ab}(\Omega_{ab}) = \rho^a(\theta_a) + \rho^b(\theta_b) \tag{12}$$

where

$$\rho^a(\theta_a) = \rho^a_{00} + \rho^a_{10}\cos\theta_a + \frac{1}{2}\rho^a_{20}(3\cos^2\theta_a - 1) \tag{13}$$

and similarly for $\rho^b(\theta_b)$. Here the angle θ_a defines the angle between the site−site vector from a to b and the z-axis in the local axis system of site a. We have used an approximate axial symmetry at each atomic site with the z-axis in the local axis system pointing radially outward from the carbon to the attached hydrogen (Figure 7). In this work, the shape function for carbon included terms to rank two, and for hydrogen, where anisotropy is less important, only terms up to rank one were found to be necessary.

Rather than attempt a direct fit of eq 11 to the short-range energies, we have performed the fitting in several stages. One of the reasons for this is numerical stability. In general, due to the highly coupled nature of the parameters in eq 11, a direct fit tends to result in an unphysical parameter set and is, therefore, best avoided, particularly if the parameters are to be *transferred* to other, similar, systems. Another reason is one of computational efficiency; if a first approximation to the parameters can be obtained using low-level ab inito data obtained at a low computational cost and subsequently improved using a smaller set of higher quality ab inito data, then the overall computational cost of the fitting process is reduced.

3.1. Stage 1: Fit to $E^{(1)}_{sr} = E^{(1)}_{exch} + E^{(1)}_{elst,pen}$. In the first stage, we have fitted the Born−Mayer terms to the first-order contribution to the short-range energy:

$$E^{(1)}_{sr} = E^{(1)}_{exch} + E^{(1)}_{elst,pen} \tag{14}$$

Since the first-order energies constitute the major part of the total short-range energy, the resulting Born−Mayer parameters are a good starting point for the next stage of the fitting process. Additionally, because these terms are computationally inexpensive, we can calculate $E^{(1)}_{sr}$ at a fairly large number of dimer configurations with a modest amount of computational resource. We have computed $E^{(1)}_{sr}$ at 500 benzene configurations using the CamCASP[43] program. These calculations were performed using the MC Sadlej-pVTZ basis together with the JK-TZVPP auxiliary basis.

We have used the method described in ref 34 to fit eq 11 to $E^{(1)}_{sr}$. As pointed out above, the parameters in eq 11, which is a sum of exponential terms, tend to be highly coupled, and a direct fit often leads to an unphysical parameter set. It would be more appropriate to fit the exponential terms in eq 11 *individually*, for each pair of sites, but to achieve this, we need to partition the short-range energy $E^{(1)}_{sr}$ into contributions from pairs of atoms. While this partitioning cannot be rigorously defined, a reasonable break-up of the short-range energy can be obtained through the density overlap model[61] in a manner outlined below and described in more detail in refs 34 and 41.

The density overlap model postulates that the short-range energy is nearly proportional to the overlap between the molecular electron densities. The short-range energy is generally taken to be the exchange−repulsion energy, but here, we additionally include the first-order electrostatic penetration energy. Therefore, we have

$$E^{(1)}_{sr} = K_0 S^{\gamma}_{\rho} \tag{15}$$

where, K_0 and γ are constants, and if ρ^X_e is the electron density of molecule X, then the density overlap is defined as $S_\rho = \int \rho^A_e(\mathbf{r})\rho^B_e(\mathbf{r})d^3\mathbf{r}$. For the asymptotically correct densities we have used here, the exponent γ has been shown[62] to be exactly 1, so K_0 is the only free parameter in this model which can be determined by least-squares minimization.

Now, if we partition the electron density into atomic contributions, that is $\rho^A_e(\mathbf{r}) = \sum_{a\in A}\rho^a_e(\mathbf{r})$, then we can define a *distributed* density overlap $S^{ab}_\rho = \int \rho^a_e(\mathbf{r})\rho^b_e(\mathbf{r})d^3\mathbf{r}$, and, hence, a generalized form of the overlap model:[41,63,64]

$$E^{(1)}_{sr} = \sum_{a\in A, b\in B} E^{(1)}_{sr}(ab) = \sum_{a\in A, b\in B} K^{ab} S^{ab}_{\rho} \tag{16}$$

where K^{ab} are constants to be fitted, and the partitioned short-range energy $E^{(1)}_{sr}(ab)$ is implicitly defined through the above equation. The density partitioning is not unique and can be achieved in a variety of ways. We have used density fitting to achieve the partitioning, which is analogous to the Gaussian multipole method of Wheatley.[65] Details of weighting schemes and constraints used in the fitting process are described in ref 34.

Having obtained the atom−atom partitioned short-range energy $E^{(1)}_{sr}(ab)$, it is now relatively easy to fit the contributions of individual pairs of sites to a single Born−Mayer term from eq 11. These fits are well-defined and result in physically sensible values for the parameters.

The various stages of the above process were performed using the CamCASP[43] and Orient[19] programs, and the overall weighted rms error for the fitted energies was 0.82 kJ mol^{-1}. Symmetry was taken into account at all stages in the fitting process.

Table 1. Parameters of BAP in a.u.[a]

atom pair	$l_a\kappa_a$	$l_b\kappa_b$	ρ	α	C_6
CC	00	00	4.1780	1.8683	30.452
	10	00	0.2535		
	20	00	−2.0390		
CH	00	00	5.4242	1.7370	12.490
	00	10	−0.4663		
	10	00	0.1472		
	20	00	−0.1422		
HH	00	00	3.4400	1.5263	5.092
	10	00	0.3611		

[a] The pre-exponential factor, G, is 0.001 a.u., and the damping factor, β, is 1.6485 a.u. The C_6 coefficients quoted here include the scaling factor of 1.372 (see discussion in Section 2.4.1). The point charges used for carbon and hydrogen atoms are −0.1111 and 0.1111 a.u., respectively.

3.2. Stage 2: Fit to $E_{sr}^{(2)}$. We now refine the parameter set obtained in Stage 1, Section (3.1), against a smaller but higher-quality data set of short-range energies, this time using the second-order terms, as defined in eq 10. $E_{sr}^{(2)}$ was obtained using SAPT(DFT) energies calculated at the first 100 configurations used in Stage 1 and the further 27 specific dimers at the orientations shown in Figure 5. This time we used the much larger MC+ basis type.

The relaxation of the parameter set was performed using penalty functions of the form $(p_i - p_i^0)^2$, where p_i^0 is the anchor value obtained from Stage 1. In this way, the fit was refined while preventing the parameters from taking on non-physical values. The final choice of parameters to be relaxed and the weights given to the harmonic constraints were chosen with an element of empiricism, although values of α_{ab} and $\rho_{00,00}^{ab}$ were constrained more tightly than other parameters. The final fit had a weighted rms residual error of 0.75 kJ mol^{-1}.

3.3. Stage 3: Final Fit to Relax all Parameters. In the final step, we simultaneously refined the Born−Mayer parameters and the dispersion coefficients (with harmonic constraints imposed) to fit SAPT(DFT) energies calculated in the MC+ basis at the 127 geometries used in Stage 2, Section (3.2). The parameters in the electrostatic model were kept fixed during this step. The SAPT(DFT) energies were weighted so as to favor more negative energies to ensure the potential well was accurately fitted. The weighting scheme used has been adapted from that used in ref 66 and is given by

$$w_i = \Theta(E_{int}^i - E_0)\left(\frac{E_0}{E_{int}^i}\right)^2 + [1 - \Theta(E_{int}^i - E_0)]\exp[\eta(E_0 - E_{int}^i)] \quad (17)$$

where $\Theta(x)$ is the Heaviside step function. E_{int}^i are the SAPT(DFT) energies, and the parameters E_0 and η were set to 3 and 0.1 mol kJ^{-1}, respectively.

The parameters for the resulting benzene anisotropic potential (BAP) are given in Table 1. It should be noted that for the shape function coefficients, $\rho_{l_a\kappa_a, l_b\kappa_b}^{ab}$, of a given atom pair (i.e., C−C, C−H, or H−H) $\rho_{00,00}^{ab}$ is defined to be the sum of the rank 00 terms in eq 13, that is, $\rho_{00,00}^{ab} = \rho_{00}^a + \rho_{00}^b$, whereas the mixed-rank terms are defined as $\rho_{i0,00}^{ab} = \rho_{i0}^a$, and $\rho_{00,i0}^{ab} = \rho_{i0}^b$ for $i \in \{1, 2\}$.[59]

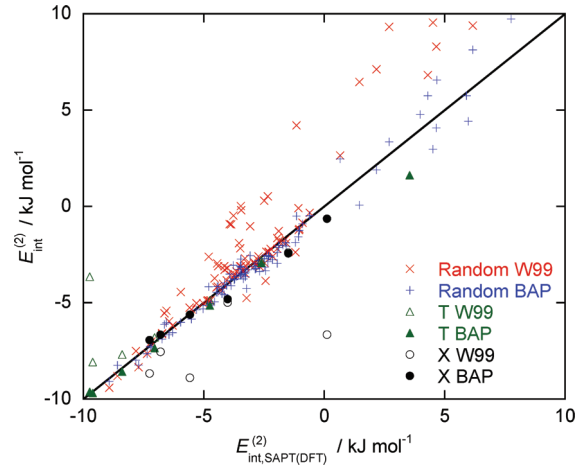

Figure 8. Comparison of the benzene anisotropic potential (BAP) with SAPT(DFT) energies calculated for 100 random benzene dimer orientations.

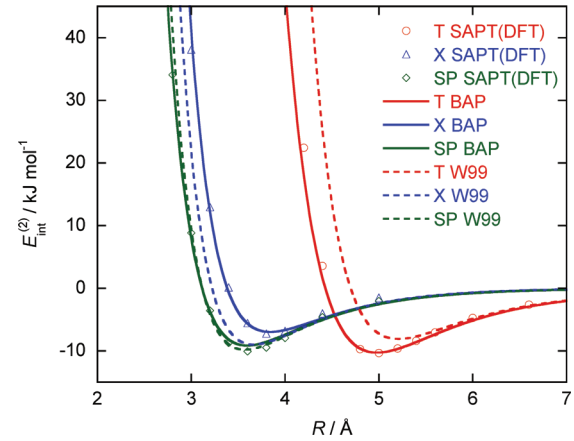

Figure 9. Comparison of BAP with SAPT(DFT) energies and W99 potential for T-shaped (T), crossed (X) and slipped-parallel (SP) dimer configurations.

The weighted rms residual energy $(E_{int}^i - E_{fit}^i)$ for the benzene anisotropic potential, when compared with the 127 SAPT(DFT) energies, was found to be 0.49 kJ mol^{-1}. Figure 8 shows the scatter of energies of the new potential compared to SAPT(DFT) energies for the random benzene dimer configurations and for some of the specific configurations chosen. For comparison, the W99 potential has been included, and the plot shows energies calculated with the new potential are noticeably less scattered. The scatter which remains for the new potential results is likely to be due to the damped C_6 isotropic dispersion model, which cannot be further improved without going to a more detailed model. Figure 9 shows a comparison of the new benzene potential with SAPT(DFT) energies and the W99 potential for the orientations shown in Figure 5. The new potential matches the SAPT(DFT) results in all configurations, especially the T-shaped configuration, where the W99 potential is notably poor.

The shape function shown in eq 12 imposes certain constraints on the parameters of the potential, e.g., $\rho_{10,00}^{HH} = \rho_{00,10}^{CH}$. However, these conditions were not imposed during the construction of the benzene anisotropic potential, and instead, shape functions for individual atoms were allowed

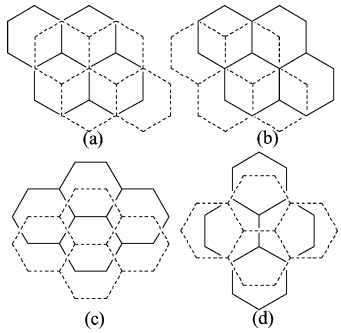

Figure 10. Pyrene: (a) Slipped-parallel L, symmetry C_{2h}; (b) Graphite-type, symmetry C_i; (c) Slipped-parallel S, symmetry C_{2h}; (d) Crossed, symmetry D_{2d}.

to vary depending on the specific atom pair considered. If the constraints are imposed, then the reduced flexibility of the functional form results in a poorer fit with a weighted rms residual of 0.96 kJ mol^{-1}.

4. Generalizing to Larger PAH Molecules

The shape function constraints given by eq 12 are probably inconsequential if the potential is restricted to a single system, but since they impose the idea of transferability,[59] they are needed if we wish to use the potential parameters on other, related, systems as well. So when generalizing the benzene potential to larger PAH molecules, we have imposed these constraints. This results in a potential with fewer parameters, but as shall be shown, it appears to perform remarkably well for the larger PAH dimers.

As can be seen from Figures 11−14, the BAP parameters transferred to the larger PAH molecules describe the interactions reasonably well. However, they are not entirely appropriate for a generalized transferable potential, as they do not satisfy the shape-function constraints. Consequently, we have used the benzene anisotropic potential parameters with these constraints imposed as the starting parametrization for the transferable anisotropic PAH potential. We have tuned these potential parameters against 111 SAPT(DFT) dimer energies, calculated by Podeszwa and Szalewicz,[18] for the naphthalene, anthracene and pyrene dimers at the orientations shown in Figures 1, 2, and 10. Molecular geometries were taken from ref 18, and ESP point charge models were calculated for each molecule with the Gaussian03[42] program in the same way as described for benzene. The geometries and partial atomic charges together with figures explaining the axes systems used in the larger PAH molecules are given in the Supporting Information. Also included in the Supporting Information are the files used to define the local axes systems in the Orient program.

The Orient program cannot simultaneously fit parameters to multiple types of molecular dimers, so an iterative scheme has been adopted. In this scheme, the initial benzene parametrization is used as the starting point for fitting the parameters for the naphthalene dimer; having obtained the new set of parameters, these now become the starting point for fitting to the anthracene dimer energies. This process is then continued, cycling through each set of dimer energies for each of the four PAH molecules. In order to converge to a parameter set, the harmonic constraints used in the fitting procedure were tightened after each iteration. Eventually the parameters are so tightly constrained that they could not be varied; this gave us the final parameter set. While this

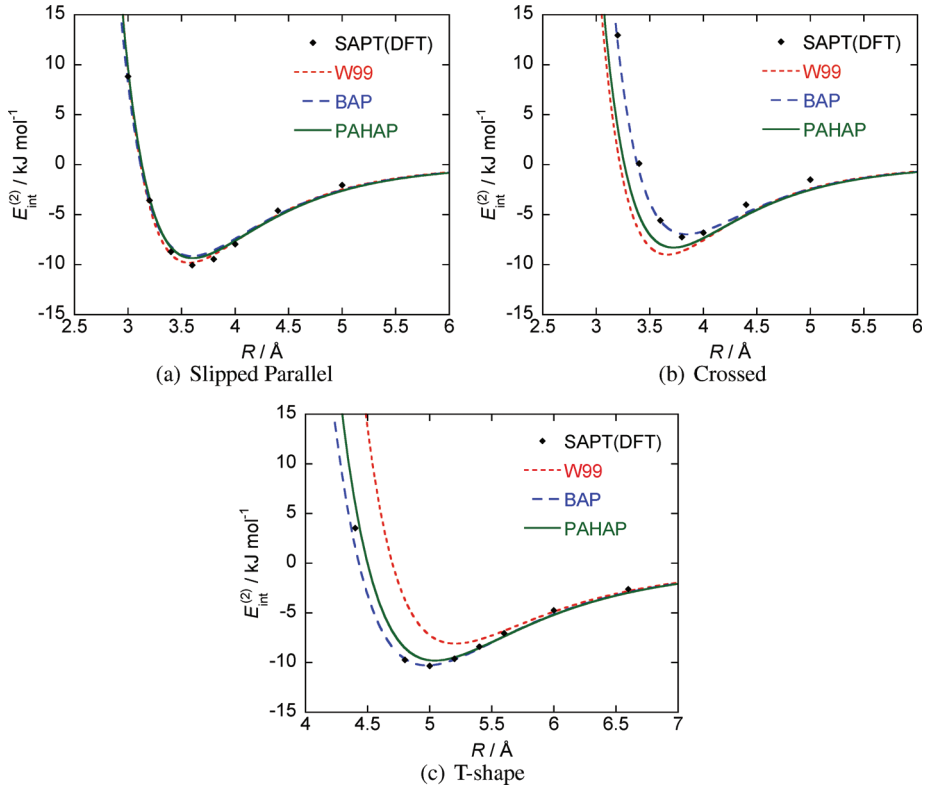

Figure 11. Comparison of the W99 potential, the BAP, and the PAH anisotropic potential (PAHAP) with SAPT(DFT) energies for benzene dimers. Model potential energies have been calculated using the Orient[19] program.

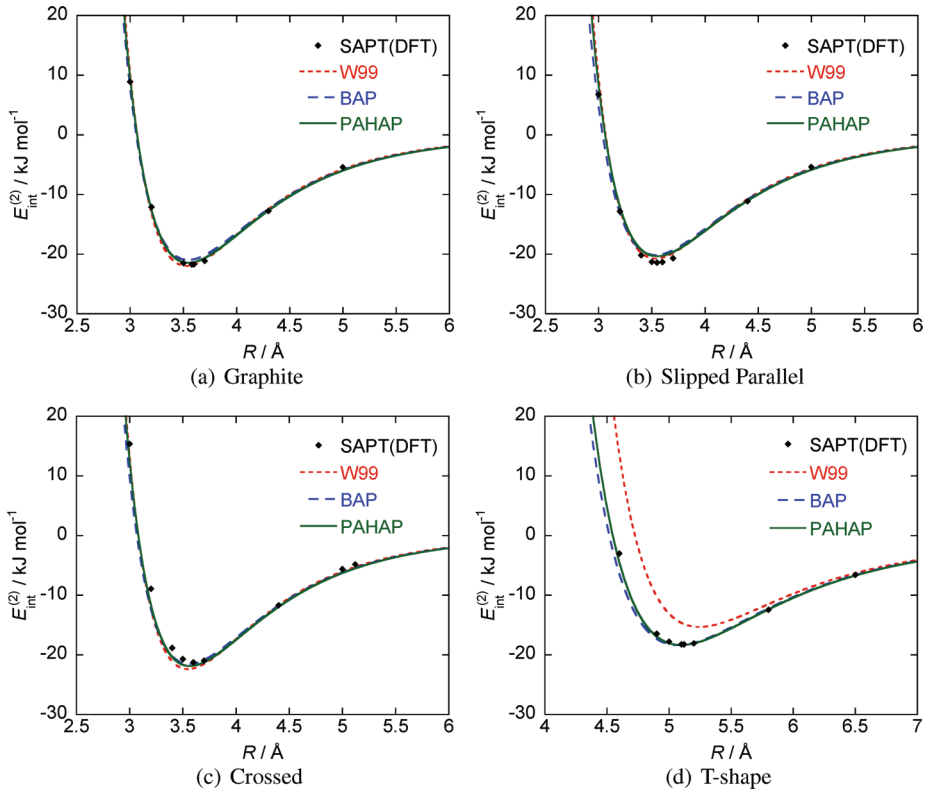

Figure 12. Comparison of the W99 potential, the BAP, and the PAHAP with SAPT(DFT) energies for naphthalene dimers. Model potential energies have been calculated using the Orient[19] program.

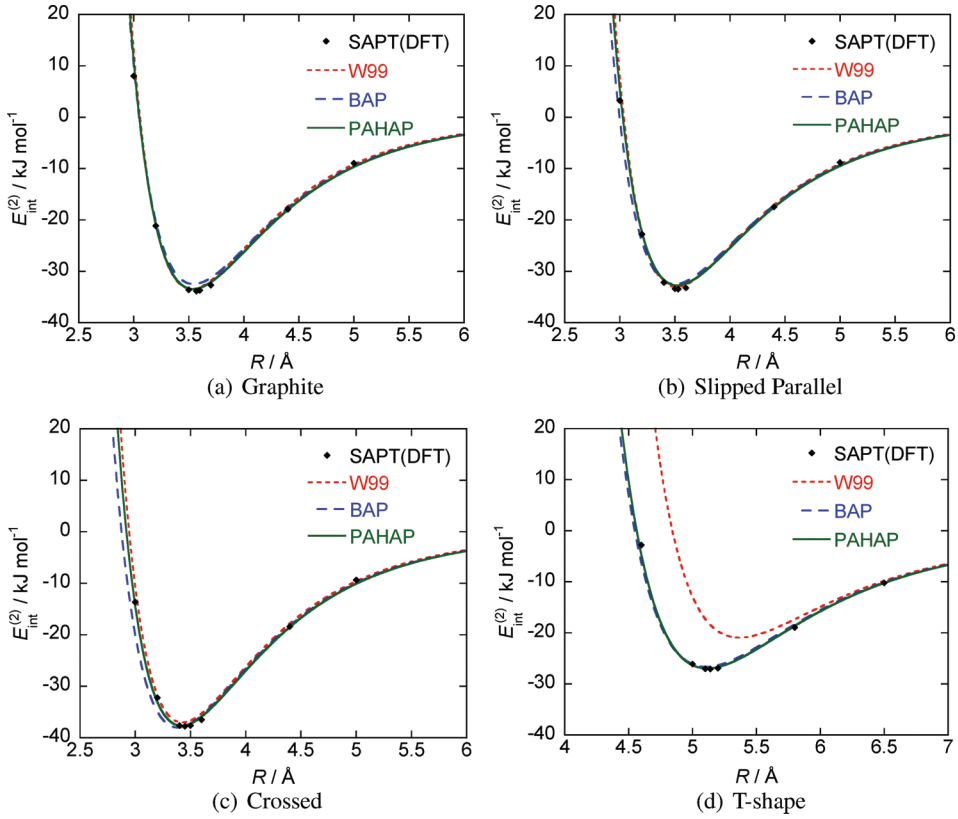

Figure 13. Comparison of the W99 potential, the BAP, and the PAHAP with SAPT(DFT) energies for anthracene dimers. Model potential energies have been calculated using the Orient[19] program.

procedure is by no means optimum, it has proved adequate and resulted in a generalized parameter set that is not only able to model the interactions of the larger PAH molecules but also the 127 benzene dimer geometries.

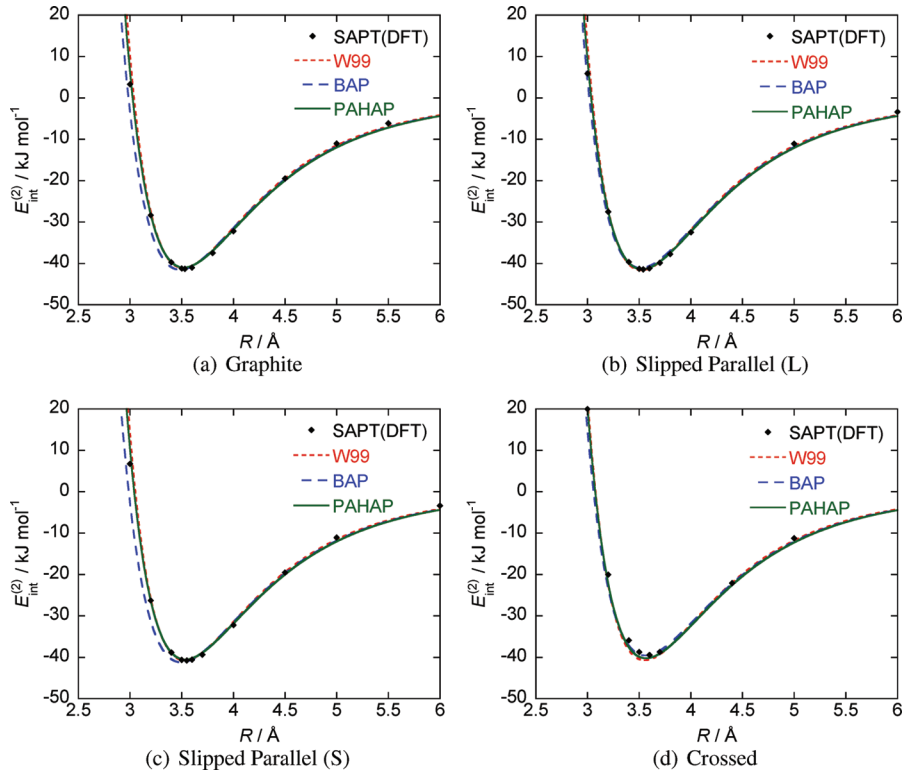

Figure 14. Comparison of the W99 potential, the BAP, and the PAHAP with SAPT(DFT) energies for pyrene dimers. Model potential energies have been calculated using the Orient[19] program.

Table 2. Parameters of PAH Anisotropic Atom−atom Potential in a.u.[a]

atom pair	$l_a\kappa_a$	$l_b\kappa_b$	ρ	α	C_6
C C	00	00	5.8147	1.8615	30.469
	10	00	0.0217		
	20	00	−0.2208		
C H	00	00	5.1505	1.7756	12.840
	00	10	−0.2718		
	10	00	0.0217		
	20	00	−0.2208		
H H	00	00	4.4862	1.4312	5.359
	10	00	−0.2718		

[a] The pre-exponential factor, G, is 0.001, and the damping factor, β, is 1.6485. The C_6 coefficients quoted here are effective coefficients and include an implicit scaling factor.

Table 2 shows the set of parameters obtained for the general PAH anisotropic potential (PAHAP) which, unlike the initial benzene parametrization, satisfy the constraints imposed by the shape function given in eq 12. A brief comparison of the new PAHAP shape functions with the equivalent W99 shape functions is given in the Supporting Information. From Figures 11−14, we see that the PAHAP parametrization is also a slight improvement over the benzene parametrization for the larger PAHs, without a significant loss of accuracy for the benzene dimer energies, although the description of the crossed benzene configuration is poorer. In particular, in contrast to the W99 potential, the PAHAP potential correctly models the PAH interaction energies at both the stacked as well as the T-shaped configurations. The overall weighted rms residual error over the 238 dimer configurations considered was found to be 0.73 kJ mol^{-1}, which is more than three times less than the error of 2.54 kJ mol^{-1} incurred by the W99 potential.

5. Discussion

Using ab initio calculations, we have developed a transferable anisotropic potential for polycyclic aromatic hydrocarbons that surpasses some of the best empirically derived isotropic potentials in accuracy. In particular, this potential accurately predicts intermolecular interactions for both stacked and non-stacked dimer configurations, such as the T-shape dimer. This is important, while stacked configurations are generally energetically more favorable for most PAH dimers, when modeling clusters of PAHs, as in the context of nascent soot particles, non-stacked configurations are also present.[13,67]

To assess the overall accuracy of our potential, we have to consider the accuracy of both the fit and the SAPT(DFT) energies. The accuracy of the former is shown by the weighted rms residual error, which was calculated to be 0.73 kJ mol^{-1} over the 238 dimer configurations considered. To determine the accuracy of the latter, comparison must be made to other ab initio results. Highly accurate benzene dimer energies have been obtained at CCSD(T) and QCISD(T) levels by Janowski and Pulay.[68] In this work, the largest calculation at QCISD(T)/aug-cc-pVQZ level involved 30 correlated orbitals and 1512 basis functions. Of the three benzene dimers configurations considered, only energy calculations for the T-shape dimer are directly comparable to our SAPT(DFT) calculations. At a separation of 4.989 Å, the QCISD(T) binding energy extrapolated to infinite basis is 11.23 kJ mol^{-1}, whereas the corresponding SAPT(DFT) binding energy calculated at a separation of 5.0 Å in our MC+ basis is 10.33 kJ mol^{-1}. This discrepancy is probably due to the difference in basis sets. The MC+ basis used in our SAPT(DFT) calculations comprises the Sadlej-pVTZ

basis for the monomer centered functions and the extra 3s2p1d midbond functions, but this combined basis is considerably smaller than that used in QCISD(T) calculations and is the likely cause for the underestimation of the binding energy. Using larger basis sets would reduce the error but would increase computational demands prohibitively. Thus, at our chosen level of theory, the error is unavoidable, and given the transferable nature of our PAH potential, we believe this error to be quite acceptable.

The new transferable PAH anisotropic potential represents a first step in our planned investigation of the intermolecular chemistry involved in the clustering of PAHs, which is thought to be an important step in the formation of nascent soot particles. While this potential can be used in its own right, it is hoped that it will also provide an accurate reference against which we can produce a general coarse-grained PAH potential, necessary for the study of large molecular clusters. The potential may also find applications in other fields were the effects of anisotropy could be important, such as organic crystal structure prediction.[69]

Acknowledgment. T.S.T. gratefully acknowledges financial support from the Engineering and Physical Sciences Research Council, Shell Research Ltd. and from Churchill College, Cambridge. The authors thank R. Podeszwa and K. Szalewicz for data provided relating to work reported in ref 18.

Supporting Information Available: Monomer coordinates with partial atomic point charges for benzene, naphthalene, anthracene and pyrene are provided. Axes files which define the local axis system used for each molecule are also provided for use with the Orient program.[19] This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Herdman, J. D.; Miller, J. H. *J. Phys. Chem. A* **2008**, *112*, 6249–6256.

(2) Chen, H. X.; Dobbins, R. A. *Combust. Sci. Technol.* **2000**, *159*, 109–128.

(3) Ishiguro, T.; Takatori, Y.; Akihama, K. *Combust. Flame* **1997**, *108*, 231–234.

(4) Vander Wal, R. L.; Yezerets, A.; Currier, N. W.; Kim, D. H.; Wang, C. M. *Carbon* **2007**, *45*, 70–77.

(5) McKinnon, J. T.; Howard, J. B. *Proc. Combust. Inst.* **1992**, *24*, 965–971.

(6) Miller, J. H. *Proc. Combust. Inst.* **1990**, *23*, 91–98.

(7) Miller, J. H. *Proc. Combust. Inst.* **2005**, *30*, 1381–1388.

(8) Miller, J. H.; Smyth, K. C.; Mallard, W. G. *Proc. Combust. Inst.* **1985**, *20*, 1139–1147.

(9) Schuetz, C. A.; Frenklach, M. *Proc. Combust. Inst.* **2002**, *29*, 2307–2314.

(10) Appel, J.; Bockhorn, H.; Wulkow, M. *Chemosphere* **2001**, *42*, 635–645.

(11) Happold, J.; Grotheer, H.; Aigner, M. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1247–1254.

(12) Miller, J. H.; Mallard, W. G.; Smyth, K. C. *J. Phys. Chem.* **1984**, *88*, 4963–4910.

(13) Rapacioli, M.; Calvo, F.; Spiegelman, F.; Joblin, C.; Wales, D. J. *J. Phys. Chem. A* **2005**, *109*, 2487–2497.

(14) van de Waal, B. W. *J. Chem. Phys.* **1983**, *79*, 3948–3961.

(15) Williams, D. E. *J. Mol. Struct.* **1999**, *485−486*, 321–347.

(16) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1–20.

(17) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1154–1166.

(18) Podeszwa, R.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2735–2746.

(19) Stone, A. J.; Dullweber, A.; Engkvist, O.; Fraschini, E.; Hodges, M. P.; Meredith, A. W.; Nutt, D. R.; Popelier, P. L. A.; Wales, D. J. *ORIENT: a program for studying interactions between molecules*, version 4.6; University of Cambridge: Cambridge, U.K., 2002; http://www-stone.ch.cam.ac.uk/programs.html. Accessed on October, 2009.

(20) Weilmünster, P.; Keller, A.; Homann, K. H. *Combust. Flame* **1999**, *116*, 62–83.

(21) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2004**, *120*, 647–659.

(22) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **2000**, *319*, 547–554.

(23) Gerenkamp, M.; Grimme, S. *Chem. Phys. Lett.* **2004**, *392*, 229–235.

(24) Misquitta, A. J.; Szalewicz, K. *Chem. Phys. Lett.* **2002**, *357*, 301–306.

(25) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2003**, *91*, 033201.

(26) Misquitta, A. J.; Szalewicz, K. *J. Chem. Phys.* **2005**, *122*, 214109.

(27) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005**, *123*, 214103.

(28) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464–470.

(29) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319–325.

(30) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778–784.

(31) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.

(32) DiStasio, R. A., Jr.; von Helden, G.; Steele, R. P.; Head-Gordon, M. *Chem. Phys. Lett.* **2007**, *437*, 277–283.

(33) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Phys. Chem. A* **2006**, *110*, 10345–10354.

(34) Misquitta, A. J.; Welch, G. W. A.; Stone, A. J.; Price, S. L. *Chem. Phys. Lett.* **2008**, *456*, 105–109.

(35) Podeszwa, R.; Bukowski, R.; Rice, B. M.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5561–5569.

(36) Williams, J. G; Stone, A. J. *J. Chem. Phys.* **2003**, *119*, 4620–4628.

(37) Misquitta, A. J.; Stone, A. J. *J. Chem. Phys.* **2006**, *124*, 024111.

(38) Misquitta, A. J.; Stone, A. J. *J. Chem. Theory Comput.* **2008**, *4*, 7–18.

(39) Misquitta, A. J.; Stone, A. J.; Price, S. L. *J. Chem. Theory Comput.* **2008**, *4*, 19–32.

(40) Misquitta, A. J.; Stone, A. J. *Mol. Phys.* **2008**, *106*, 1631–1643.

(41) Stone, A. J.; Misquitta, A. J. *Int. Rev. Phys. Chem.* **2007**, *26*, 193–222.

(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr. J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*, Gaussian, Inc.: Wallingford, CT, 2004.

(43) Misquitta, A. J.; Stone, A. J. *CamCASP: a program for studying intermolecular interactions and for the calculation of molecular properties in distributed form*, University of Cambridge: Cambridge, U.K., 2007; http://www-stone.ch.cam.ac.uk/programs.html. Accessed on 10/2009.

(44) *DALTON, a molecular electronic structure program, Release 2.0*, 2005, see http://www.kjemi.uio.no/software/dalton/dalton.html. Accessed on 10/2009.

(45) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(46) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995.

(47) Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 10180–10189.

(48) Tozer, D. J. *J. Chem. Phys.* **2000**, *112*, 3507–3515.

(49) Lias, S.; Liebman, J. Ion Energetics Data in NIST Chemistry WebBook, *NIST Standard Reference Database Number 69*, Linstrom P. J. Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, p. 20899; http://webbook.nist.gov. Accessed on October, 2009.

(50) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, B. *J. Chem. Phys.* **1995**, *103*, 7374–7391.

(51) Akin-Ojo, O.; Bukowski, R.; Szalewicz, K. *J. Chem. Phys.* **2003**, *119*, 8379–8396.

(52) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(53) Shoemake, K. Uniform random rotations. In *Graphics Gems III*; Academic Press Professional, Inc.: San Diego, CA, 1992; pp 124−132.

(54) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(55) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047–1064.

(56) Stone, A. J.; Alderton, M. *Mol. Phys.* **2002**, *100*, 221–233.

(57) Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J. Phys. Chem.* **1996**, *100*, 7352–7360.

(58) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.

(59) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1996; pp 156−184.

(60) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726–3741.

(61) Kim, Y. S.; Kim, S. K.; D., L. W. *Chem. Phys. Lett.* **1981**, *80*, 574.

(62) Misquitta, A. J.; Stone, A. J. Ab initio atom-atom potentials using CamCASP: pyridine as an example, in preparation.

(63) Nobeli, I.; Price, S. L. *J. Phys. Chem. A* **2000**, *103*, 6448–6457.

(64) Hodges, M. P.; Wheatley, R. J. *Chem. Phys. Lett.* **2000**, *326*, 263–268.

(65) Wheatley, R. J. *Mol. Phys.* **1993**, *79*, 597–610.

(66) Bukowski, R.; Sadlej, J.; Jeziorski, B.; Jankowski, P.; Szalewicz, K.; Kucharski, S. A.; Williams, H. L.; Rice, B. M. *J. Chem. Phys.* **1999**, *110*, 3785–3803.

(67) Totton, T. S.; Misquitta, A. J.; Chakrabarti, D.; Wales, D. J.; Kraft, M. Modelling the Internal Structure of Nascent Soot Particles, 2009, Accepted for publication in Combustion and Flame.

(68) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27–32.

(69) Price, S. L.; Price, L. S. Modelling intermolecular forces for organic crystal structure prediction. In *Intermolecular Forces and Clusters I: Structure and Bonding*, 2nd ed.; Mingos, D., Wales, D. J., Eds.; Springer-Verlag: Berlin, Heidelberg, Germany, 2005; Vol. 115, pp 81−123.

CT9004883