

**THEORETICAL INVESTIGATIONS OF p-p
INTERACTIONS AND THEIR ROLE IN MOLECULAR
RECOGNITION**

A Thesis
Presented to
The Academic Faculty

By
Mutasem Omar Sinnokrot

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Chemistry

Georgia Institute of Technology
July 2004

**THEORETICAL INVESTIGATIONS OF p-p
INTERACTIONS AND THEIR ROLE IN MOLECULAR
RECOGNITION**

Approved by:

Dr. C. David Sherrill, Advisor

Dr. Mostafa A. El-Sayed

Dr. Rigoberto Hernandez

Dr. James A. Mulholland

Dr. Robert L. Whetten

Date Approved: July 1st, 2004

ACKNOWLEDGEMENT

I wish to thank everyone who contributed directly or indirectly to making this thesis possible. In particular I am deeply grateful to the reason of my existence, my mother Aisha and my father Omar for all their tireless help and support throughout the years. Their selflessness and priceless support is beyond what I can explain in words.

I am also deeply thankful and indebted to my wife Dareen who has been very patient with me in the past two years, offering all the assistance and moral support, especially in the most crucial parts of my tenure as a Ph.D. student. Not only is she my wife, but also my best friend.

My brothers Mishab, Mohammed, Ramzy and Mahdy are to be thanked for their encouragement and for being there when I needed them.

At last, but not least, I wish to thank my Ph.D. advisor Professor C. David Sherrill for all the great help and support throughout the last four years. Not only is he my adviser, but a true friend whom I will always cherish and respect.

TABLE OF CONTENTS

ACKNOWLEDGEMENT		iii
LIST OF TABLES		vi
LIST OF FIGURES		vii
SUMMARY		x
CHAPTER 1	INTRODUCTION	1
1.1	Noncovalent Interactions and Molecular Recognition	1
1.2	Thesis Objectives	2
CHAPTER 2	BRIEF OVERVIEW OF ELECTRONIC STRUCTURE THEORY	4
2.1	The Schrödinger Equation	4
2.2	The Hartree-Fock Approximation	6
2.3	Basis Sets	12
2.4	Electron Correlation Methods	15
2.5	Møller-Plesset Perturbation Theory (MPPT)	15
2.6	The Coupled-Cluster Methods	18
2.7	Energy Decomposition: Symmetry-Adapted Perturbation Theory (SAPT)	21
CHAPTER 3	ESTIMATES OF THE AB INITIO LIMIT FOR p-p INTERACTIONS: THE BENZENE DIMER	24
3.1	Introduction	24
3.2	Theoretical Methods	29
3.3	Results and Discussion	32

3.4	Conclusions	43
CHAPTER 4	UNEXPECTED SUBSTITUENT EFFECTS IN FACE-TO-FACE p-STACKING INTERACTIONS	46
4.1	Introduction	46
4.2	Theoretical Methods	48
4.3	Results and Discussion	49
4.4	Conclusions	52
CHAPTER 5	SUBSTITUENT EFFECTS IN p-p INTERACTIONS: SANDWICH AND T-SHAPED CONFIGURATIONS	53
5.1	Introduction	53
5.2	Theoretical Methods	58
5.3	Results and Discussion	61
5.4	Conclusions	81
CHAPTER 6	HIGHLY ACCURATE COUPLED-CLUSTER POTENTIAL ENERGY CURVES FOR BENZENE DIMER	84
6.1	Introduction	85
6.2	Theoretical Methods	90
6.3	Results and discussion	92
6.4	Conclusions	102
CHAPTER 7	SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS	104
REFERENCES		110

LIST OF TABLES

Table 3.1	Benzene Dimer Geometries (R in Å)	36
Table 3.2	Binding Energies (kcal mol ⁻¹) for Different Configurations of the Benzene Dimer	38
Table 4.1	Interfragment Distances (Å) and Binding Energies (kcal mol ⁻¹) for Face-to-Face Dimers of Benzene with Substituted Benzenes	50
Table 5.1	Interaction Energies (in kcal mol ⁻¹) for Various Dimers	62
Table 5.2	Interaction Energies (in kcal mol ⁻¹) Relative to Benzene Dimer	65
Table 5.3	Contributions to the Interaction Energy (kcal mol ⁻¹) for Different Sandwich Dimer Configurations 2a-e	72
Table 5.4	Contributions to the Interaction Energy (kcal mol ⁻¹) for Different T-shaped Dimer Configurations 3a-e	77
Table 5.5	Contributions to the Interaction Energy (kcal mol ⁻¹) for Different T-shaped(2) Dimer Configurations 4a-e	80
Table 6.1	Benzene Dimer Geometries (R in Å)	92
Table 6.2	Binding Energies (kcal mol ⁻¹) for Different Configuration of Benzene Dimer	94

LIST OF FIGURES

Figure 3.1	Sandwich, T-shaped, and parallel-displaced configurations of the benzene dimer	25
Figure 3.2	Equilibrium geometry of benzene (bond distances in Å)	32
Figure 3.3	Potential energy curves for the sandwich configuration at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory	34
Figure 3.4	Potential energy curves for the T-shaped configuration at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory	34
Figure 3.5	Potential energy curves for the parallel-displaced configuration at the (CP-corrected) MP2/aug-cc-pVDZ level of theory	35
Figure 3.6	Hartree-Fock binding energies for each dimer structure as a function of basis set. All computations were performed at the same best estimate geometry for each configuration. All energies are negative (repulsive). Labels aXZ denote the aug-cc-pVXZ basis sets	40
Figure 3.7	MP2 electron correlation energy contributions to binding energies for each dimer structure as a function of basis set. The total MP2 binding energies are obtained by adding these values to the Hartree-Fock contributions in the previous figure. All computations were performed at the same best estimate geometry for each configuration. Labels aXZ denote the aug-cc-pVXZ basis sets	41
Figure 4.1	Sandwich configurations of benzene dimer and benzene-monosubstituted benzene dimers	48

Figure 4.2	Electrostatic potentials computed using Hartree-Fock and a 6-31G* basis set with a scale of -25 to +25 kcal mol ⁻¹ . Potentials using B3LYP/6-31G* appear very similar	51
Figure 5.1	The Sandwich and two T-shaped configurations of benzene dimer and benzene-monosubstituted benzene dimers	56
Figure 5.2	Electrostatic potentials computed using Hartree-Fock and a 6-31G* basis set with a scale of -25 to +25 kcal mol ⁻¹ . Potentials using B3LYP/6-31G* appear very similar	67
Figure 5.3	Plot of dispersion energy for sandwich dimers (SAPT2/aug-cc-pVDZ', exchange-corrected) vs. polarizability in the direction perpendicular to the aromatic plane (HF/aug-cc-pVDZ)	75
Figure 5.4	The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different sandwich dimers compared to benzene dimer	75
Figure 5.5	The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different T-shaped dimers compared to benzene dimer	78
Figure 5.6	Plot of dispersion energy for T-shaped(2) dimers (SAPT2/aug-cc-pVDZ', exchange-corrected) vs. polarizability in the direction perpendicular to the aromatic plane (HF/aug-cc-pVDZ)	81
Figure 5.7	The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different T-shaped(2) dimers compared to benzene dimer	82
Figure 6.1	Sandwich, T-shaped, and parallel-displaced configurations of the benzene dimer	89

Figure 6.2	MP2 and CCSD(T) potential energy curves for the sandwich configuration of the benzene dimer	93
Figure 6.3	MP2 and CCSD(T) potential energy curves for the T-shaped configuration of the benzene dimer	95
Figure 6.4	Asymptotic $1/r^6$ fit of the estimated CCSD(T)/aug-cc-pVQZ* potential energy curve for the sandwich configuration of benzene dimer	97
Figure 6.5	Asymptotic $1/r^6$ fit of the estimated CCSD(T)/aug-cc-pVQZ* potential energy curve for the T-shaped configuration of benzene dimer	98
Figure 6.6	Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) MP2/aug-cc-pVTZ level of theory	99
Figure 6.7	Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) MP2/aug-cc-pVQZ* level of theory	99
Figure 6.8	Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) CCSD(T)/aug-cc-pVDZ* level of theory	100
Figure 6.9	Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) estimated CCSD(T)/aug-cc-pVQZ* level of theory	101

SUMMARY

Noncovalent interactions are of pivotal importance in many areas of chemistry, biology, and materials science, and the intermolecular interactions involving aromatic rings in particular, are fundamental to molecular organization and recognition processes. The work detailed in this thesis involves the application of state-of-the-art ab initio electronic structure theory methods to elucidate the nature of π - π interactions.

The binding energies, and geometrical and orientational preferences of the simplest prototype of aromatic π - π interactions, the benzene dimer, are explored. We obtain the first converged values of the binding energies using highly accurate methods and large basis sets. Results from this study predict the T-shaped and parallel-displaced configurations of benzene dimer to be nearly isoenergetic.

The role of substituents in tuning π - π interaction is investigated. By studying dimers of benzene with various monosubstituted benzenes (in the sandwich and two T-shaped configurations), we surprisingly find that *all* of the substituted sandwich dimers considered bind more strongly than benzene dimer. We also find that these interactions can be tuned by a modest degree of substitution. Energy decomposition analysis using symmetry-adapted perturbation theory (SAPT) reveals that models based solely on electrostatic effects will have difficulty in reliably predicting substituent effects in π - π interactions.

CHAPTER 1

INTRODUCTION

1.1 Noncovalent Interactions and Molecular Recognition

Attractive interactions between aromatic π systems are one of the principal noncovalent forces governing supramolecular organization and recognition processes. These aromatic π - π interactions are ubiquitous in diverse areas of science and molecular engineering.¹⁻³ They are key interactions influencing the tertiary structure of proteins,^{4,5} the vertical base stacking in DNA,⁶⁻⁹ and the intercalation of different drugs into DNA.⁶ π - π interactions also play a major role in stabilizing host-guest complexes^{10,11} and in self-assembly based on synthetic molecules.¹²⁻¹⁴

Despite the fact that the importance of π - π interactions is widely recognized among scientists, their small value (2-3 kcal mol⁻¹) presents a challenge for both theorists and experimentalists to obtain a full understanding of their origins and geometrical preferences. In contrast to hydrogen bonds, which are well-understood, strong, and highly directional electrostatic interactions, π - π interactions are harder to model due to the multiple points of intermolecular contact and strong dependency on functional groups involved. Therefore, one has to take into account the electrostatic, dispersion, exchange-repulsion and induction forces which all play a major role in stabilizing these interactions.

1.2 Thesis Objectives

To achieve the goal of understanding π - π interaction and how they may be tuned, we first consider in this thesis the simplest prototype of aromatic π - π interactions, the benzene dimer. We explore different configurations of benzene dimer using state-of-the-art electronic structure methods. Explicitly correlated MP2-R12/A techniques,¹⁵⁻¹⁷ coupled with focal-point analysis,¹⁸ yield the first converged estimates for the binding energy of this system which are accurate to within a few tenths of one kilocalorie per mole.

We then explore the role of substituents in tuning π -stacking interactions by examining the dimer interactions of benzene with monosubstituted benzenes, with substituents OH, CH₃, F, and CN. Very little theoretical or experimental work has been done to examine substituent effects on binding energies and geometries of π - π systems. By enhancing or reducing the strength of π - π interactions compared to benzene dimer, substituents may alter the energy landscape. We also provide a critical analysis of the Hunter-Sanders rules² that are widely used to make qualitative predictions regarding π - π interactions.

In the latter parts of the thesis, we further explore the energetics of benzene dimer and monosubstituted dimers by decomposing the binding energy into its electrostatic, exchange-repulsion, dispersion, and induction components using symmetry-adapted perturbation theory (SAPT).^{19,20} Such analysis provides detailed information on the relative importance of these different energy components and their contribution to the overall stability of the dimers.

In addition to providing high-level estimates of the binding energies of different benzene dimer configurations, we generate highly accurate potential energy curves that display the variation of the binding energies as a function of intermonomer distances. The availability of these high quality potential energy curves will be critical to molecular mechanics force field developers in their endeavor to accurately model systems exhibiting π -stacking interactions.

CHAPTER 2

BRIEF OVERVIEW OF ELECTRONIC STRUCTURE THEORY

2.1 The Schrödinger Equation

The notation used in this chapter is based upon that of Ira N. Levine's "Quantum Chemistry", 4th edition (Prentice Hall, 1991),²¹ Frank Jensen's "Introduction to Computational Chemistry" (Wiley, 2003),²² and Atilla Szabo and Neil S. Ostlund's "Modern Quantum Chemistry" (Dover, 1996).²³

In this work, we utilize different electronic structure theory methods that attempt solve the electronic Schrödinger equation in its time-independent non-relativistic form. Also, since the ratio of the nuclear to electron mass is larger than 1800/1, that means that electrons move much faster than nuclei, and we can invoke the Born-Oppenheimer Approximation and assume that the nuclei are stationary; this will produce the *electronic Schrödinger equation*

$$\hat{H}_e(r; R)\Psi_e(r; R) = E_e(R)\Psi_e(r; R) \quad [2.1]$$

where \hat{H}_e is the electronic Hamiltonian operator that depends on the electronic coordinates (denoted here by r) and parametrically on the nuclear coordinates (denoted here by R), and Ψ_e is the electronic wavefunction with eigenvalues E_e .

The electronic Hamiltonian operator is composed of the electronic kinetic energy, the energies of nuclear-electron attractions, nuclear-nuclear repulsions, and electron-electron repulsions. In atomic units, it can be written as

$$\hat{H}_e = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} + \sum_{i>j} \frac{1}{r_{ij}} \quad [2.2]$$

This can be further simplified if we collect operators based on their electron indices. If we define the one-electron operators $h(i)$ as

$$h(i) = -\frac{1}{2} \nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \quad [2.3]$$

and V_{mn} as the nuclear-nuclear repulsion term $\sum_{A>B} \frac{Z_A Z_B}{R_{AB}}$, then the electronic Hamiltonian

can be simply written in terms of one and two-electron operators, as

$$\hat{H}_e = \sum_i h(i) + \sum_{i>j} \frac{1}{r_{ij}} + V_{mn} \quad [2.4]$$

V_{mn} is a constant term for fixed nuclei-nuclei distances and can be added at the end of the calculation without changing the eigenfunctions; it will only shift the energy eigenvalues.

From the solution of equation [2.1], we obtain the electronic energy $E_e(R)$. When $E_e(R)$ is solved for all nuclear coordinates we obtain the *potential energy surface*, which is the potential energy felt by the nuclei. From the electronic energy $E_e(R)$, and the electronic wavefunction $\Psi_e(r;R)$ and their derivatives, we can obtain many types of molecular properties such as (but not limited to) multipole moments, polarizabilities, molecular structures, and various electronic, vibrational, and rotational spectra.

2.2 The Hartree-Fock Approximation

2.2.1 The Slater Determinants

Hartree-Fock theory is an approximate theory where each electron feels an *average* coulomb repulsion due to the presence of the other electrons. Consider an N -electron wavefunction $\Psi(x_1, x_2, x_3 \dots x_N)$ that depends on coordinates $x_1, x_2, x_3 \dots x_N$ of electrons $1, 2, 3 \dots N$, where x_i represents both the spatial (r_i) and spin ($\mathbf{w} = \mathbf{a}$ or \mathbf{b}) coordinates of the i -th electron and is represented by $x_i = (r_i, \mathbf{w})$. In Hartree-Fock theory, this wavefunction is expressed as a *single* Slater determinant (named after John Slater), which is a generalization of the Hartree Product.

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \mathbf{c}_1(x_1) & \mathbf{c}_2(x_1) & \mathbf{c}_3(x_1) & \cdots & \mathbf{c}_N(x_1) \\ \mathbf{c}_1(x_2) & \mathbf{c}_2(x_2) & \mathbf{c}_3(x_2) & \cdots & \mathbf{c}_N(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{c}_1(x_N) & \mathbf{c}_2(x_N) & \mathbf{c}_3(x_N) & \cdots & \mathbf{c}_N(x_N) \end{vmatrix} \quad [2.5]$$

where $\mathbf{c}_i(x_j)$ means that electron x_j is associated with orbital \mathbf{c}_i .

The Slater determinant expression of the wavefunction conforms to the indistinguishability of electrons, where we can see that each electron is associated with all the orbitals in the wavefunction. It also satisfies the antisymmetry principle, which states that the wavefunction should be antisymmetric under the interchange of any two electrons. If we interchange the labels x_1 , and x_2 above, then two rows of the determinant are interchanged, resulting in a change of sign for the wavefunction $\Psi(x_1, x_2, x_3 \dots x_N)$.

2.2.2 The Hartree-Fock Energy Expression

To generate a solution to the electronic Schrödinger equation we employ the variational principle, which states that any approximate wavefunction (in this case the Hartree-Fock wavefunction) will produce an energy that is larger than true energy. For normalized wavefunctions, the energy is calculated as the expectation value of the Hamiltonian operator:

$$E_e = \langle \Psi | \hat{H}_e | \Psi \rangle \quad [2.6]$$

Without going into the details of the derivations (see F. Jensen “Introduction to Computational Chemistry”, Wiley 2003), the Hartree-Fock energy can be written in terms of one- and two-electron integrals and the nuclear repulsion energy, as

$$E_{HF} = \sum_{i=1}^N \langle \mathbf{c}_i | h(i) | \mathbf{c}_i \rangle + \frac{1}{2} \sum_{ij}^N (\mathbf{J}_{ij} - \mathbf{K}_{ij}) + V_{nn} \quad [2.7]$$

where the first summation is made of one-electron integrals

$$\langle \mathbf{c}_i | h(i) | \mathbf{c}_i \rangle = \int dx_1 \mathbf{c}_i^*(x_1) h(r_1) \mathbf{c}_i(x_1) \quad [2.8]$$

and the second summation is made of two-electron coulomb \mathbf{J} and exchange \mathbf{K} integrals

$$\mathbf{J}_{ij} = \int dx_1 dx_2 \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_i(x_1) \mathbf{c}_j(x_2) \quad [2.9]$$

$$\mathbf{K}_{ij} = \int dx_1 dx_2 \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_j(x_1) \mathbf{c}_i(x_2) \quad [2.10]$$

Note that the above energy expression [2.7] is over spinorbitals $\mathbf{c}_i(x_j)$. Similar expressions over spatial orbitals exist when the spin part is explicitly integrated out of the energy equation.

2.2.3 The Hartree-Fock Equations

In the previous subsection, we presented the Hartree-Fock energy expression in terms one- and two-electron integrals over the molecular spinorbitals. The Hartree-Fock method involves determining the orbitals that minimize the above energy expression (equation [2.7]). This is achieved through the Lagrange's method of undetermined multipliers. In this method, the variation in the Lagrange function, L , with respect to orbital variation $\mathbf{c}_i \rightarrow \mathbf{c}_i + d\mathbf{c}_i$ is set equal to zero as follows:

$$L = E_{HF} - \sum_{ij}^N \mathbf{I}_{ij} \left(\int \mathbf{c}_i^*(x_1) \mathbf{c}_j(x_1) dx_1 - d_{ij} \right) \quad [2.11]$$

$$dL = dE_{HF} - \sum_{ij}^N \mathbf{I}_{ij} \left(\int d\mathbf{c}_i^*(x_1) \mathbf{c}_j(x_1) dx_1 + \mathbf{c}_i^*(x_1) d\mathbf{c}_j(x_1) dx_1 \right) = 0 \quad [2.12]$$

where \mathbf{I}_{ij} 's are the set of Lagrange's undetermined multipliers. After omitting a few algebraic steps that can be found in most quantum chemistry textbooks, we arrive at the Hartree-Fock equation for the spinorbitals:

$$f(x_1) | \mathbf{c}_i(x_1) \rangle = \mathbf{e}_i | \mathbf{c}_i(x_1) \rangle \quad [2.13]$$

where electron 1 has been arbitrarily assigned to orbital \mathbf{c}_i . \mathbf{e}_i is the orbital energy of spinorbital \mathbf{y}_i . The *Fock operator* f in the above equation is defined by:

$$f(x_1) = h(x_1) + \sum_j^N \mathbf{J}_j(x_1) - \mathbf{K}_j(x_1) \quad [2.14]$$

where $h(x_1)$ is the *core-Hamiltonian operator* defined above. $\mathbf{J}_j(x_1)$ is the *coulomb operator*, and $\mathbf{K}_j(x_1)$ is the exchange operator, and are both defined as

$$\mathbf{J}_j(x_1) \mathbf{c}_i(x_1) = \left[\int dx_2 \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_j(x_2) \right] \mathbf{c}_i(x_1) \quad [2.15]$$

$$\mathbf{K}_j(x_1)\mathbf{c}_i(x_1) = \left[\int dx_2 \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_i(x_2) \right] \mathbf{c}_j(x_1) \quad [2.16]$$

2.2.4 The Closed-Shell Hartree-Fock Equations

We can eliminate the spin from the Hartree Fock equation [2.13] by writing spinorbitals in terms of their spin and spatial parts as follows:

$$\mathbf{c}_i(x_1) = \begin{cases} \mathbf{y}_j(r_1)\mathbf{a}(\mathbf{w}) \\ \mathbf{y}_j(r_1)\mathbf{b}(\mathbf{w}) \end{cases} \quad [2.17]$$

For closed-shell and restricted (doubly-occupied) set of orbitals, we can write the spatial Hartree-Fock equation as

$$f(r_1)|\mathbf{y}_i(r_1)\rangle = \mathbf{e}_i|\mathbf{y}_i(r_1)\rangle \quad [2.18]$$

The *Fock operator* f in the above equation is defined by:

$$f(r_1) = h(r_1) + \sum_j^{N/2} 2J_j(r_1) - K_j(r_1) \quad [2.19]$$

where the sum is over the $N/2$ occupied orbitals. The closed-shell *coulomb* and *exchange operators* are now defined in terms of spatial orbitals as

$$J_j(r_1)\mathbf{y}_i(r_1) = \left[\int dr_2 \mathbf{y}_j^*(r_2) \frac{1}{r_{12}} \mathbf{y}_j(r_2) \right] \mathbf{y}_i(r_1) \quad [2.20]$$

$$K_j(r_1)\mathbf{y}_i(r_1) = \left[\int dr_2 \mathbf{y}_j^*(r_2) \frac{1}{r_{12}} \mathbf{y}_i(r_2) \right] \mathbf{y}_j(r_1) \quad [2.21]$$

and the energy expression for closed-shell Hartree-Fock is now also written in terms of spatial orbitals as

$$E_{HF} = 2 \sum_i^{N/2} (\mathbf{y}_i | h | \mathbf{y}_i) + \sum_{ij}^{N/2} 2(\mathbf{y}_i \mathbf{y}_i | \mathbf{y}_j \mathbf{y}_j) - (\mathbf{y}_i \mathbf{y}_j | \mathbf{y}_j \mathbf{y}_i) \quad [2.22]$$

where the one- and two-electron integrals are written in terms of spatial orbitals as

$$(\mathbf{y}_i | h | \mathbf{y}_j) = \int d\mathbf{r} \mathbf{y}_i^*(r_1) h(r_1) \mathbf{y}_j(r_1) \quad [2.23]$$

$$(\mathbf{y}_i \mathbf{y}_j | \mathbf{y}_k \mathbf{y}_l) = \int d\mathbf{r}_1 d\mathbf{r}_2 \mathbf{y}_i^*(r_1) \mathbf{y}_j(r_1) \frac{1}{r_{12}} \mathbf{y}_k^*(r_2) \mathbf{y}_l(r_2) \quad [2.24]$$

2.2.5 The Roothaan Equations

In this subsection and the following subsections, our discussion will be limited to the restricted closed-shell Hartree-Fock formalism.

Since we don't know the exact form of the molecular orbitals for most atoms and molecules, we introduce a set of K known spatial basis functions and express each of the spatial orbitals as a linear combination of these basis functions. Following the notation of Szabo and Ostlund, we have

$$\mathbf{y}_i = \sum_{m=1}^K C_m \mathbf{f}_m \quad [2.25]$$

where C_m are a set of unknown coefficients. The problem of calculating the orbitals is transformed to the problem of calculating this set of coefficients. Substituting \mathbf{y}_i into equation [2.18] yields a set of K equations known as the *Roothaan equations*. These equations can be written in matrix notation as

$$\mathbf{FC} = \mathbf{SCe} \quad [2.26]$$

where \mathbf{F} , the *Fock matrix*, is a Hermitian $K \times K$ matrix, and has the elements

$$F_{mn} = \int d\mathbf{r}_1 \mathbf{f}_m^*(r_1) f(r_1) \mathbf{f}_n(r_1) \quad [2.27]$$

and \mathbf{C} , the coefficients matrix, is a $K \times K$ square matrix. The \mathbf{S} matrix, called the *overlap matrix*, is also a Hermitian $K \times K$ square matrix with elements given by

$$S_{\mathbf{m}} = \int dr_1 \mathbf{f}_{\mathbf{m}}^*(r_1) \mathbf{f}_{\mathbf{n}}(r_1) \quad [2.28]$$

Finally, \mathbf{e} is a $K \times K$ diagonal matrix consisting of orbital energies \mathbf{e}_i .

From equation [2.27] we can obtain an explicit form of the Fock matrix elements by writing the Fock operator explicitly in terms of one- and two-electron operators as in equation [2.19]. After working through some algebra, we obtain

$$F_{\mathbf{m}} = H_{\mathbf{m}}^{core} + \sum_{\mathbf{1s}} P_{\mathbf{1s}} \{ (\mathbf{m}\mathbf{m} | \mathbf{1s}) - \frac{1}{2} (\mathbf{m}\mathbf{l} | \mathbf{n}\mathbf{s}) \} \quad [2.29]$$

where $H_{\mathbf{m}}^{core}$ is the *core-Hamiltonian matrix* defined by

$$H_{\mathbf{m}}^{core} = \int dr_1 \mathbf{f}_{\mathbf{m}}^*(r_1) h(r_1) \mathbf{f}_{\mathbf{n}}(r_1) \quad [2.30]$$

The $P_{\mathbf{1s}}$'s in eq. [2.29] are called the *density matrix elements* and defined as follows:

$$P_{\mathbf{1s}} = 2 \sum_i^{N/2} c_{\mathbf{1i}}^* c_{\mathbf{s i}} \quad [2.31]$$

and the two-electron integrals are now written in terms of the known basis functions:

$$(\mathbf{m}\mathbf{m} | \mathbf{1s}) = \int dr_1 dr_2 \mathbf{f}_{\mathbf{m}}^*(r_1) \mathbf{f}_{\mathbf{n}}(r_1) \frac{1}{r_{12}} \mathbf{f}_{\mathbf{l}}^*(r_2) \mathbf{f}_{\mathbf{s}}(r_2) \quad [2.32]$$

From the above analysis of the Fock operator, we see that the Fock matrix elements $F_{\mathbf{m}}$ depend on the expansion coefficients $C_{\mathbf{m}}$ and hence the wavefunction itself. Therefore, the Roothaan equations must be solved iteratively starting with a set of expansion coefficients $C_{\mathbf{m}}$ and testing for convergence at each iteration step.

The two-electron integrals over basis sets are the time-consuming part in the evaluation of the Fock matrix elements. Each of the basis functions in equation [2.32] can be located on a different atom, producing four-center integrals. The existence of these

four-center integrals results in a formal scaling of K^4 for the Hartree-Fock method unless special techniques are employed.

2.3 Basis Sets

To perform different *ab initio* calculations, one has to start with the choice of a basis set to complement the theoretical method used. As we have seen in the previous section, a linear combination of different basis functions is used to represent the different molecular orbitals. In addition to the quality of the theoretical method, the quality of the basis set used will play a major role in determining the accuracy of the computation. The combination of the *ab initio* method and the basis set used will define the *level of theory* at which the calculations are performed.

2.3.1 Gaussian Basis Functions

In this work, we will be using Gaussian basis functions that are centered on the atoms of the molecule; these atomic orbitals (AOs) are called Gaussian-type orbitals (GTOs) and were first introduced by S. F. Boys in 1950.²⁴ A *primitive* Cartesian Gaussian centered at $\vec{O} = x_0\hat{i} + y_0\hat{j} + z_0\hat{k}$ is given by:

$$\mathbf{f}_i(\vec{r}) = N_i (x - x_0)^l (y - y_0)^m (z - z_0)^n e^{-\alpha_i |\vec{r} - \vec{O}|^2} \quad [2.33]$$

where N_i is the normalization constant. This Gaussian is characterized by three nonnegative integers; l , m , and n . The sum of these integer $L = l + m + n$ is referred to as the angular momentum. The exponent α_i determines the radial spread of the Gaussian functions. The larger the exponent, the shorter the radial spread. Functions with larger

exponents are called *tight functions*. *Diffuse functions*, on the other hand, have small exponents and fall off slowly with increasing r .

In practical implementations of basis sets, several GTO's are combined together to provide better approximation of different orbitals in atomic and nuclear environments. Such combination of GTOs is known as contracted Gaussian-type orbitals (CGTOs). A contracted Cartesian Gaussian centered at $\vec{O} = x_0\hat{i} + y_0\hat{j} + z_0\hat{k}$ is given by:

$$\mathbf{f}_i(\vec{r}) = N_i(x - x_0)^l(y - y_0)^m(z - z_0)^n \sum_{j=1}^{l_c} d_{ij} e^{-\mathbf{a}_{ij}|\vec{r} - \vec{O}|^2} \quad [2.34]$$

where l_c is the contraction length, d_{ij} and \mathbf{a}_{ij} are the contractions coefficients and the exponents, respectively.

2.3.2 Basis Set Notation

A *minimal basis set* is a basis that uses one contracted Gaussian function to represent each atomic orbital of each atom. Such basis set would contain one function for H and He, and five basis functions for atoms from Li to Ne (one for $1s$, one for $2s$, and three for the $2p$ -orbitals). To get more accurate results, one needs to improve on minimal basis sets.

A *double-zeta basis set* represents a major improvement over the minimal basis set. Two CGTOs correspond to each atomic orbital. For the previous example of Carbon, the $1s$, $2s$, and three $2p$ -orbitals will be each represented by two CGTO's. In a triple-zeta basis set, three CGTO's correspond to each atomic orbital. Similarly, we can define quadruple and larger- zeta basis sets.

A *split-valence double-zeta basis* set uses one contracted function for each inner AO and two contracted functions for each valence AO. For a calculation on Carbon, for example, there would be one CGTO for the $1s$ -orbital, two for the $2s$ -orbital, and two for each of the three $2p$ -orbitals. For a *split-valence triple-zeta basis* set, there would be one function for $1s$, and 3 functions for the $2s$ and each of the three $2p$ -orbitals.

In the molecular environment, orbitals become distorted by the presence of the other atoms in the molecule. This is called the *polarization* of orbitals. To account for this polarization and distortion of orbitals, one needs to add basis functions of higher angular momentum than any of the occupied atomic orbitals. For example, p -type basis functions are added to describe the polarization of s -orbitals, and d -type basis functions are added to describe polarization of p -orbitals. Sometimes, additional higher angular momenta basis functions are used.

For anions, the orbitals are more expanded in space compared to neutral molecules. To be able to accurately compute different anion properties, one needs to add *diffuse* functions to the basis set. As mentioned previously, these functions have small exponents, and hence fall off slowly with increased radial distance.

Diffuse functions are also needed to accurately compute different properties of weakly bound systems held together by noncovalent interactions. Since our work involves studying different prototypes of π - π systems held together by noncovalent forces, we make extensive use of diffuse functions in the basis sets we employ to obtain optimized geometries and binding energies for these weakly-bound systems.

2.4 Electron Correlation Methods

The Hartree-Fock theory is an approximate theory. It does not take into account the *instantaneous* electron-electron repulsion, and replaces it by an average potential contained in the J and K terms (see equations [2.9] and [2.10]). This means that each electron feels an *average* electrostatic interaction due to the presence of the other $N-1$ electrons. In summary, we can say that the Hartree-Fock method does not contain an explicit treatment of *electron correlation*. A great body of work in electronic structure theory deals with developing models which explicitly treat electron correlation to obtain more accurate wavefunctions and more accurate estimations of the electronic energy.

In the next sections, a brief overview of two electron correlation methods used in this work will be presented. We will discuss the second-order Møller-Plesset perturbation theory (MP2) and the coupled-cluster theory treatments of electron correlation. We will also introduce symmetry-adapted perturbation theory (SAPT), which is employed to analyze the interaction energy in terms of physically meaningful components such as electrostatic, induction, dispersion, and exchange energies.

2.5 Møller-Plesset Perturbation Theory (MPPT)

An important characteristic of MPPT is its size-extensivity. This means that the energy of a system AB computed at infinite separation will be equal to the sum of energies of the two fragments A and B computed individually. Nevertheless, energies computed using perturbation theory methods are not variational, which means that they do not constitute upper limits to the true energy of the system.

In MPPT, the unperturbed electronic Hamiltonian \hat{H}^0 is defined as the sum of one-electron Fock operators:

$$\hat{H}^0 = \sum_{i=1}^N f(x_i) \quad [2.35]$$

and the zeroth-order energy is consequently given by

$$E^{(0)} = \langle \Psi_{HF} | \hat{H}^0 | \Psi_{HF} \rangle \quad [2.36]$$

The perturbation in MPPT is given by the difference between the sum of the Fock operators and the true electronic Hamiltonian

$$\hat{H}' = \hat{H} - \hat{H}^0 = \hat{H} - \sum_{i=1}^N f(x_i) \quad [2.37]$$

The first-order energy correction $E^{(1)}$ is given by the expectation value of the perturbation using the Hartree-Fock wavefunction:

$$E^{(1)} = \langle \Psi_{HF} | \hat{H}' | \Psi_{HF} \rangle \quad [2.38]$$

Adding the zeroth-order energy $E^{(0)}$ to $E^{(1)}$ gives

$$\begin{aligned} E^{(0)} + E^{(1)} &= \langle \Psi_{HF} | \hat{H}^0 | \Psi_{HF} \rangle + \langle \Psi_{HF} | \hat{H}' | \Psi_{HF} \rangle \\ &= \langle \Psi_{HF} | \hat{H}^0 + \hat{H}' | \Psi_{HF} \rangle = \langle \Psi_{HF} | \hat{H} | \Psi_{HF} \rangle \end{aligned} \quad [2.39]$$

But $\langle \Psi_{HF} | \hat{H} | \Psi_{HF} \rangle$ is just the Hartree-Fock energy, so we can write the sum of the zeroth-order energy and the first-order energy correction as

$$E^{(0)} + E^{(1)} = E_{HF} \quad [2.40]$$

It is obvious from the above discussion that in order to improve on the Hartree-Fock energy we need to evaluate the second-order energy correction. This correction is known as the MP2 energy correction. By applying Brillouin's Theorem and the Condon-

Slater rules, it turns out that only doubly excited determinants contribute to the MP2 energy correction as follows

$$E^{(2)} = \sum_{ij}^{\text{occupied}} \sum_{ab}^{\text{virtual}} \frac{\langle \Psi_{HF} | \hat{H}' | \Psi_{ij}^{ab} \rangle \langle \Psi_{ij}^{ab} | \hat{H}' | \Psi_{HF} \rangle}{\mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_a - \mathbf{e}_b} \quad [2.41]$$

where Ψ_{ij}^{ab} represents a doubly-excited determinant that differs from the reference Hartree-Fock determinant Ψ_{HF} by the replacement of spinorbitals \mathbf{c}_i and \mathbf{c}_j with virtual spinorbitals \mathbf{c}_a and \mathbf{c}_b , respectively. The energies \mathbf{e}_i , \mathbf{e}_j , \mathbf{e}_a , and \mathbf{e}_b represent the orbital energies associated spinorbital \mathbf{c}_i , \mathbf{c}_j , \mathbf{c}_a , and \mathbf{c}_b , respectively.

In terms of two-electron integral over spinorbitals, the above expression can be written as:

$$E^{(2)} = \frac{1}{4} \sum_{ij}^{\text{occupied}} \sum_{ab}^{\text{virtual}} \frac{[\langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_a \mathbf{c}_b \rangle - \langle \mathbf{c}_a \mathbf{c}_b | \mathbf{c}_i \mathbf{c}_j \rangle]^2}{\mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_a - \mathbf{e}_b} \quad [2.42]$$

where

$$\begin{aligned} \langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_a \mathbf{c}_b \rangle &= \int \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_a(x_1) \mathbf{c}_b(x_2) dx_1 dx_2 \\ \langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_b \mathbf{c}_a \rangle &= \int \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_b(x_1) \mathbf{c}_a(x_2) dx_1 dx_2 \end{aligned} \quad [2.43]$$

The MP2 methods scales with basis set size as K^5 . This is because the integral transformation from the atomic orbital (AO) to the molecular orbital (MO) basis where the MP2 energies are computed, scales as K^5 . Due to their relatively moderate computational cost, MP2 methods are widely used in the calculations of the correlation energy.

2.6 The Coupled-Cluster Methods

The coupled-cluster method in its various approximations is highly accurate, and it is widely used to obtain a good account of electron correlation in molecular calculations. The average accuracy of the coupled-cluster method is about 0.004Å for the determination of the bond length, 0.03° for bond angles, and 2% for the determination of vibrational frequencies.

2.6.1 The Coupled-Cluster Wavefunction

In coupled-cluster theory the wavefunction is expressed as an exponential product. Higher order excitations from a reference are obtained as a product of lower order ones. If we consider the Hartree-Fock wavefunction as the reference wavefunction then the coupled-cluster wavefunction can be written as

$$\Psi_{cc} = e^{\hat{T}} \Psi_{HF} \quad [2.44]$$

where Ψ_{HF} is the ground-state Hartree-Fock wavefunction. The operator $e^{\hat{T}}$ is defined as

$$e^{\hat{T}} = 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{\hat{T}^k}{k!} \quad [2.45]$$

If we have a molecule with N electrons, then the above cluster operator \hat{T} is given by

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_N \quad [2.46]$$

The first operator on the right-hand side of the above equation is \hat{T}_1 ; this is an operator which generates the single excitations from the Hartree-Fock reference. The second operator \hat{T}_2 generates the double excitations. These two operators are defined by

$$\hat{T}_1 \Psi_{HF} = \sum_i^{\text{occupied}} \sum_a^{\text{virtual}} t_i^a \Psi_i^a \quad [2.47]$$

$$\hat{T}_2 \Psi_{HF} = \sum_{i < j}^{\text{occupied}} \sum_{a < b}^{\text{virtual}} t_{ij}^{ab} \Psi_{ij}^{ab} \quad [2.48]$$

The expansion coefficients t_i^a and t_{ij}^{ab} are usually called the t-amplitudes.

In the coupled-cluster singles and doubles approximation, denoted CCSD, both the \hat{T}_1 and \hat{T}_2 operators are used in constructing Ψ_{cc} , and the wavefunction is written as

$$\Psi_{CCSD} = e^{\hat{T}} \Psi_{HF}, \quad \hat{T} = \hat{T}_1 + \hat{T}_2 \quad [2.49]$$

By using a Taylor-series expansion of $e^{\hat{T}}$, we can rewrite the above equation as

$$\Psi_{CCSD} = \left\{ 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}(\hat{T}_1^2 + \hat{T}_1 \hat{T}_2 + \hat{T}_2 \hat{T}_1 + \hat{T}_2^2) + \frac{1}{6}(\hat{T}_1^3 + \hat{T}_1^2 \hat{T}_2 + \hat{T}_1 \hat{T}_2 \hat{T}_1 + \hat{T}_1 \hat{T}_2^2 + \hat{T}_2 \hat{T}_1^2 + \hat{T}_2 \hat{T}_1 \hat{T}_2 + \hat{T}_2^2 \hat{T}_1 + \hat{T}_2^3) + \dots \right\} \Psi_{HF} \quad [2.50]$$

We can see from the above equation that “disconnected” higher order excitations (like \hat{T}_1^2 and \hat{T}_2^2) in coupled-cluster theory can be obtained as a product of lower order connected excitations (\hat{T}_1 and \hat{T}_2). Also, since the cluster operators \hat{T}_1 and \hat{T}_2 commute with themselves and with each other, many of the above expressions in equation [2.50] can be simplified. For example, the product $\hat{T}_1 \hat{T}_2 \hat{T}_1$ can be simplified to $\hat{T}_2 \hat{T}_1^2$.

2.6.2 The Coupled-Cluster Energy Expression

Coupled-cluster theory (and many-body perturbation theory) use non-variational energy expressions. Using the coupled-cluster wavefunction $\Psi_{cc} = e^{\hat{T}} \Psi_{HF}$, the time-independent Schrödinger equation becomes

$$\hat{H}e^{\hat{T}}\Psi_{HF} = E_{cc}e^{\hat{T}}\Psi_{HF} \quad [2.51]$$

Multiplying by Ψ_{HF} and integrating gives

$$\langle \Psi_{HF} | \hat{H}e^{\hat{T}} | \Psi_{HF} \rangle = E_{cc} \langle \Psi_{HF} | e^{\hat{T}} \Psi_{HF} \rangle \quad [2.52]$$

Using intermediate normalization ($\langle \Psi_{HF} | \Psi_{cc} \rangle = 1$), equation [2.52] becomes

$$E_{cc} = \langle \Psi_{HF} | \hat{H}e^{\hat{T}} | \Psi_{HF} \rangle = \langle \Psi_{HF} | \hat{H}(1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 + \dots) | \Psi_{HF} \rangle \quad [2.53]$$

Since Hartree-Fock orbitals are used as a reference, Brillouin's theorem holds. Also, since the Hamiltonian is a one- and two-particle operator, terms like $\hat{T}_1\hat{T}_2$ and \hat{T}_2^3 will have zero matrix elements with the Hamiltonian in the above equation. After taking into account the above properties of the Hamiltonian, we get the following energy expression in terms of the two-electron integrals over molecular spinorbitals \mathbf{c}_i , \mathbf{c}_j , \mathbf{c}_a , and \mathbf{c}_b and the amplitudes²²

$$E_{cc} = E_{HF} + \sum_{i < j} \sum_{a < b}^{\text{occupied virtual}} (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) (\langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_a \mathbf{c}_b \rangle - \langle \mathbf{c}_a \mathbf{c}_b | \mathbf{c}_i \mathbf{c}_j \rangle) \quad [2.54]$$

where

$$\begin{aligned} \langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_a \mathbf{c}_b \rangle &= \int \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_a(x_1) \mathbf{c}_b(x_2) dx_1 dx_2 \\ \langle \mathbf{c}_i \mathbf{c}_j | \mathbf{c}_b \mathbf{c}_a \rangle &= \int \mathbf{c}_i^*(x_1) \mathbf{c}_j^*(x_2) \frac{1}{r_{12}} \mathbf{c}_b(x_1) \mathbf{c}_a(x_2) dx_1 dx_2 \end{aligned} \quad [2.55]$$

The next step is to obtain expressions for each of the amplitudes that enter into the energy equation. To facilitate that, the operator \hat{T} is truncated at a certain excitation level. The coupled-cluster singles and doubles (CCSD) method is obtained by using

$\hat{T} = \hat{T}_1 + \hat{T}_2$. Using $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$ gives the CCSDT method, which scales with basis set size as K^8 , as compared with K^6 for CCSD.

For the CCSD method, the amplitudes are generated by left projection of equation [2.51] with singly and then doubly excited determinants. This will produce exactly the right number of equations to solve for the amplitudes t_{ij}^{ab} , $t_i^a t_j^b$ and $t_i^b t_j^a$ in eq. [2.54]. The details of the derivations and the set of equations used to find the amplitudes are omitted and can be found in many advanced quantum chemistry books.^{21,22}

Since the CCSDT method is computationally very demanding, its use is limited to small systems. Therefore, to account for the effect of triple excitation in our current work (and hence obtain a better account of electron correlation), we use the coupled-cluster singles and doubles with perturbative triples CCSD(T) method.²⁵ In this method, the contribution of triple excitations to the energy is computed by perturbation theory and added to the CCSD energy. The CCSD(T) method scales with basis set size as K^7 , and produces highly-accurate results for many properties and parameters that include (but not limited to) bond lengths, bond angles, dipole moments, vibrational frequencies, and relative energies. This method is commonly referred to as “the gold standard of quantum chemistry”.

2.7 Energy Decomposition: Symmetry-Adapted Perturbation Theory (SAPT)

As mentioned above, symmetry-adapted perturbation theory (SAPT)^{19,20} is employed to analyze the interaction energy in terms of physically meaningful

components such as electrostatic, induction, dispersion, and exchange-repulsion energies. In introducing SAPT, we will be using the original notation of Jeziorski and coworkers.¹

In SAPT, the Hamiltonian of the dimer is decomposed into three parts as

$$\hat{H} = \hat{F} + \hat{V} + \hat{W} \quad [2.56]$$

where F is the Fock operator, representing the sum of the Fock operators for the separate monomers; W is the intramonomer correlation operator, accounting for the intramonomer correlation effects; and V is the intermolecular interaction operator. The SAPT interaction energy can be represented as

$$E_{\text{int}} = E_{\text{int}}^{\text{HF}} + E_{\text{int}}^{\text{CORR}} \quad [2.57]$$

where $E_{\text{int}}^{\text{HF}}$ represents lowest-order corrections that can be identified as describing interactions at the Hartree-Fock (HF) level. The $E_{\text{int}}^{\text{HF}}$ term can be represented as

$$E_{\text{int}}^{\text{HF}} = E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)} + E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + \mathbf{d}E_{\text{int,resp}}^{\text{HF}} \quad [2.58]$$

The superscripts (ab) denote orders in perturbation theory with respect to operators V and W , respectively. It can be seen from the above equation that the HF interaction energy includes first-order polarization and exchange, and second-order induction and exchange-induction contributions. The subscripts “resp” indicate that the induction and exchange-induction contributions include the coupled-perturbed HF response.²⁰ The $\mathbf{d}E_{\text{int}}^{\text{HF}}$ term contains the third- and higher-order HF induction and exchange induction contributions.

¹ Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorski, B.; Jeziorska, M.; Kucharski, S. A.; Misquitta, A. J.; Moszynski, R.; Patkowski, K.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wormer, P. E. S. *SAPT2002: An Ab Initio Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies. Sequential and Parallel Versions*, 2003.

In this thesis work, we employ the SAPT2 approach, in which the correlated portion of the interaction energy is nearly equivalent to the supermolecular MP2 correlation energy and can be represented as

$$E_{\text{int}}^{\text{CORR}} = E_{\text{elst,resp}}^{(12)} + E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)} + {}^t E_{\text{ind}}^{(22)} + {}^t E_{\text{exch-ind}}^{(22)} + E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)} \quad [2.59]$$

where ${}^t E_{\text{ind}}^{(22)}$ represents the part of $E_{\text{ind}}^{(22)}$ that is not included in $E_{\text{ind,resp}}^{(20)}$, and ${}^t E_{\text{exch-ind}}^{(22)}$ is

approximated as

$${}^t E_{\text{exch-ind}}^{(22)} \approx E_{\text{exch-ind,resp}}^{(20)} \frac{{}^t E_{\text{ind}}^{(22)}}{E_{\text{ind,resp}}^{(20)}} \quad [2.60]$$

Different studies on weakly bound dimers^{20,26-29} (like helium dimer, water dimer, Ar-H₂, and Ar-HF) that compared the SAPT approach with the corresponding supermolecular one have shown that SAPT interaction energies at the van der Waals minima differ from the supermolecular ones by a small 1-2%. All the SAPT computations reported in the current work were carried out using the SAPT2002 program.³⁰

CHAPTER 3

ESTIMATES OF THE AB INITIO LIMIT FOR π - π INTERACTIONS: THE BENZENE DIMER

3.1 Introduction

As mentioned in the first chapter of this thesis, π - π interactions are key noncovalent interaction influencing the structures of protein, vertical base stacking in DNA, host-guest complexes, and many solid state materials with aromatic rings.

In this chapter, we consider the simplest prototype of π - π interactions, the benzene dimer. Explicitly correlated (R12) quantum mechanical theories,¹⁵⁻¹⁷ coupled with focal-point analysis,^{18,31} yield binding energy estimates of unprecedented accuracy for this system. These results should be a key component in the development of a new generation of molecular mechanics force fields, especially polarizable force fields that are capable of accurately describing π - π interactions.

The small binding energy (~ 2 -3 kcal mol⁻¹) of gas-phase benzene dimer makes it a challenge for both experiment and theory. The dimer is stable only at low temperatures and low pressures and is typically prepared in supersonic jet expansions into the vacuum. Because clusters of various sizes are produced, it is necessary to use detection methods that are sensitive to their masses. Such challenges are a significant obstacle to a definitive experimental description. Moreover, the diverse experimental techniques employed to date have yielded seemingly contradictory results and are only consistent if there are at least two different low-energy potential energy minima or if the system is highly

fluxional with low barriers.³² The combined experimental and theoretical work to date suggests that the most favorable configurations are the perpendicular T-shaped and parallel-displaced (PD) geometries (see Figure 3.1), with the eclipsed sandwich (S) configuration somewhat higher in energy. Previous theoretical work³³ indicates that some

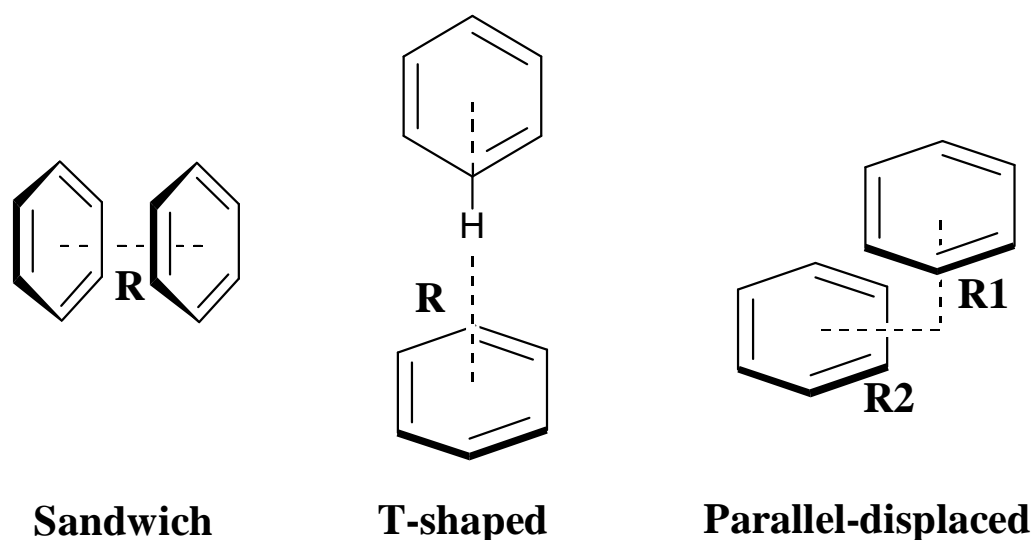


Figure 3.1. Sandwich, T-shaped, and parallel-displaced configurations of the benzene dimer.

minor variations of these configurations (e.g., the “edge-face” T-shaped configuration obtained by the top monomer in the T-shaped dimer of Figure 1.3 by 30° about the axis perpendicular to the page, or a rotated C_{6v} sandwich) are very similar in energy. Approximately perpendicular and offset-parallel configurations are frequently observed

in the crystal structure of simple aromatic compounds,^{5,34} and interacting pairs of aromatic side chains in proteins exhibit both orientations, perpendicular arrangements dominating.^{4,5} By contrast, directly overlapping rings, as in the S configuration, are rarely observed for these systems.^{4,5,34} Neutron diffraction experiments on solid benzene³⁵ find nearest neighbor orientations that are not quite T-shaped or parallel displaced.

Early molecular beam electric resonance studies by Klemperer and co-workers^{36,37} on the gas-phase benzene dimer provided evidence for a T-shaped configuration with C_{2v} symmetry. Subsequent studies of the rotational spectrum of gas-phase benzene dimer by Arunan and Gutowsky³⁸ using a microwave spectrometer gave a 4.96 Å separation between the benzene centers of mass. These experiments did not rule out the existence of the parallel-displaced or sandwich configurations, however, since they are only sensitive to molecules with dipole moments. Subsequent mass selective ionization-detected stimulated Raman Spectroscopy (IDSRS) studies of intermolecular vibrations of benzene dimer by Felker et al.^{39,40} were consistent with a dimer composed of monomers not related by a symmetry element (e.g. T-shaped). On the other hand, optical absorption spectra by Bernstein and co-workers⁴¹ support the existence of a parallel-displaced configuration with C_{2h} symmetry. Stimulated emission pumping (SEP) coupled with multiphoton ionization (MUPI) studies of the ground state vibrational spectra of benzene dimer by Schlag and co-workers⁴² support the two monomers being symmetry equivalent. Additional mass-selective hole-burning experiments by Schlag and et al.⁴³ were consistent with the existence of three different dimer configurations.

The binding energy of the dimer, obtained from the dissociation energy of the ion and the ionization potentials of the dimer and monomer, was measured as $D_0 = 1.6 \pm 0.2$ kcal mol⁻¹ by Krause et al.⁴⁴ and as 2.4 ± 0.4 kcal mol⁻¹ by Grover et al.⁴⁵

There have been a large number of theoretical studies of benzene dimer.^{32,33,46-50} The binding energy of the dimer is primarily due to London dispersion interactions,⁴⁸ which arise from favorable instantaneous multipole/induced multipole charge fluctuations. Also, electrostatic, induction, and exchange-repulsion are all major components of the binding energy for benzene dimer. Since Hartree-Fock molecular orbital theory describes each electron in the average field of all the other electrons present in the molecule, it is incapable of describing the instantaneous charge fluctuations giving rise to the stabilizing dispersion forces. Unfortunately, current implementations of Kohn-Sham density functional theory (DFT) rely on essentially local approximations for the density and are also incapable of accurately describing dispersion forces.^{49,51} Hence, wave functions based correlation methods are essential for a qualitatively accurate description of benzene dimer and other weakly-bound systems. Moreover, the need to describe the polarizability of the monomers accurately suggests that very large basis sets -including multiple polarization and diffuse functions- may be necessary.

The importance of using a large basis set has been shown by a number of theoretical studies.^{32,49,50} Before the date of finishing the current study, no truly large basis set has been used which can realistically approximate the complete basis set limit for benzene dimer. Due to their sometimes prohibitively high computational cost, the use of multiple diffuse functions in the basis sets has received little consideration. In this study, we consider the large correlation-consistent basis sets augmented by multiple

diffuse functions, through aug-cc-pVQZ (1512 basis functions), for second-order perturbation theory (MP2) computations. We also explore the effect of basis sets on the geometries by obtaining MP2 potential energy curves as a function of the distance between the centers of masses of monomers using a much larger basis function set (aug-cc-pVTZ, 828 basis functions) than previously used for benzene dimer geometry optimization.

Completer basis set (CBS) limits at the MP2 level have been estimated using the orbital invariant version of the explicitly correlated MP2-R12 method in the standard approximation A (designated as MP2-R12/A)¹⁵ with a custom basis set. Key studies by Jaffe and Smith,³² Hobza and co-workers,⁴⁶ and Tsuzuki and co-workers,^{49,50} have shown that MP2 overestimates the effect of electron correlation and hence overestimates the binding energies for benzene dimer and other weakly-bound systems. Moreover, three-body electron correlations, described by triple excitations relative to the reference configuration, are also significant. The inclusion of triple excitations in the description of the wavefunction provides a better account of electron correlation. Hence, coupled-cluster computations with perturbative triples⁵² [CCSD(T)] have also been performed and combined with the Hartree-Fock and MP2-R12/A values to estimate complete basis CCSD(T) binding energies for the sandwich, T-shaped, and parallel-displaced configurations for benzene dimer, which should be accurate to a few tenths of a kilocalorie per mole.

3.2 Theoretical Methods

Most computations on different benzene dimer configurations were performed using Dunning's augmented correlation-consistent polarized split-valence basis sets of contracted Gaussian functions,⁵³ specifically aug-cc-pVDZ (384 basis functions), aug-cc-pVTZ (828 basis functions), and aug-cc-pVQZ (1512 basis functions). The aug- prefix denotes that these basis sets have an extra set of diffuse functions (functions with low exponent) for each angular momentum appearing the basis. A special fully uncontracted 13s8p5d2f/9s3p1d C/H basis set was constructed for use in MP2-R12/A single-point calculations described below.

The optimum intermonomer distances for the planar sandwich, T-shaped, and parallel-displaced dimer configurations were computed including valence electron correlation via second-order MP2 theory in conjunction with the aug-cc-pVDZ and aug-cc-pVTZ basis sets. For the sandwich and T-shaped geometries, the distance between the centers of mass of the two benzene monomers (denoted by R in Figure 3.1) was systematically varied, while the monomer geometries were kept rigid and not allowed to change. For the parallel-displaced configuration, both the vertical and horizontal centers of mass (denoted by R1 and R2 in Figure 3.1) were systematically varied, again with rigid monomer geometries. Each dimer optimization used a monomer geometry that is fully optimized at the same level of theory. For example, the MP2/aug-cc-pVDZ dimer intermonomer optimizations used monomer geometries optimized at the MP2/aug-cc-pVDZ level of theory; these geometries are kept rigid in the dimer calculations. The counterpoise (CP) correction of Boys and Bernardi⁵⁴ was applied to account for the basis set superposition error (BSSE) which results from the use of finite basis sets.

Full geometry optimization at the MP2/cc-pVDZ level of theory supports the idea that monomers remain nearly rigid in the dimer, all C-C and C-H bond distances stay within 0.001 Å of their values in the monomer, except for 0.003 Å shortening of the C-H bond pointed at the other benzene ring in the T-shaped configuration. Angles did not change significantly in the full optimization. The above observations indicate that dimer binding energy calculations using rigid monomers will be very similar to the ones computed using fully optimized dimer geometries.

The optimal aug-cc-pVTZ MP2 intermonomer distances thus determined were coupled with the recommended benzene monomer geometry of Gauss and Stanton⁵⁵ (C-C = 1.3915 Å and C-H = 1.0800 Å) to yield our *best estimates* for the equilibrium geometry of each structure. These geometries were used to study the basis set dependence of Hartree-Fock and MP2 energies (see the Results and Discussion section for more details). They were also used to establish *ab initio* limits for the binding energies of each configuration.

The binding energy limit estimate was composed of three contributions (each counterpoise corrected): (1) the complete basis set limit at the Hartree-Fock level, approximated with the aug-cc-pVQZ basis; (2) the CBS limit for the MP2 valence correlation contribution, estimated by the MP2-R12/A method; (3) the effect of higher-order correlation, estimated as the difference between the CCSD(T) and MP2 valence correlation energies ($E_{binding}$) evaluated with the aug-cc-pVDZ basis as follows:

$$\Delta CCSD(T) = E_{binding}^{CCSD(T)/aug-cc-pVDZ} - E_{binding}^{MP2/aug-cc-pVDZ} \quad [3.1]$$

Although it was not possible to compute $\Delta CCSD(T)$ correction in equation [3.1] using a larger basis set, it should be relatively insensitive to basis set improvements, as

demonstrated by the success of additive schemes such as those found in focal-point analysis^{18,31} or the Gaussian-3 method.⁵⁶ Core orbitals were constrained to remain doubly occupied in all correlated computations. The effect of core correlation, estimated at the MP2 level using Dunning's core-valence aug-cc-pCVDZ basis,⁵⁷ was found to be negligible (less than 0.03 kcal mol⁻¹ for all configurations).

A brief comment is due on the energies obtained with the explicitly correlated MP2-R12/A method. The approximate resolution of identity utilized in the current form of the linear R12 theories puts certain requirements on the quality of the basis used in such computations. Specifically, the basis has to be complete enough in the one-particle sense that its use in the approximate resolution of the identity will not introduce significant errors. It seems that only through comparison with MP2-R12/A energies computed with larger basis sets may we rigorously evaluate the appropriateness of the custom basis set used here. Computations of such scope will only be possible with a massively parallel implementation of the method. However, the difference between the CP-corrected and uncorrected MP2-R12/A interaction energies may also be considered as estimate of the accuracy of our MP2-R12 computations of these energies. The computed differences (0.05, 0.23, and 0.13 kcal mol⁻¹ for the sandwich, T-shaped, and parallel-displaced configuration, respectively) indicate an accuracy of ~0.2 kcal mol⁻¹, which is the technical limit at the moment.

All the results reported in this chapter were obtained using the PSI 3.2,⁵⁸ and Q-Chem 2.0⁵⁹ programs except for the aug-cc-pVQZ MP2 computations, which were performed with Sandia's massively parallel quantum chemistry (MPQC) program⁶⁰⁻⁶² using 12-15 POWER3-II processors of an IBM supercomputer.

The MP2-R12/A computations were performed using the orbital invariant version of the method¹⁶ as implemented in the PSI suite.¹⁷ A new shared-memory parallel MP2-R12/A algorithm based on the direct MP2 transformation scheme of Head-Gordon et al.⁶³ was implemented for this work and made the current computations feasible. Each MP2-R12/A energy evaluation required approximately two weeks running on four processors of the IBM SP.

3.3 Results and Discussion

Dimer geometry optimization was performed at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory using rigid monomers. The monomer geometries were obtained at these same levels of theory and are displayed in Figure 3.2 along with the reference geometry of Gauss and Stanton.⁵⁵ The MP2/aug-cc-pVDZ geometry is in reasonably good agreement with the reference r_e geometry, with the bond length errors of

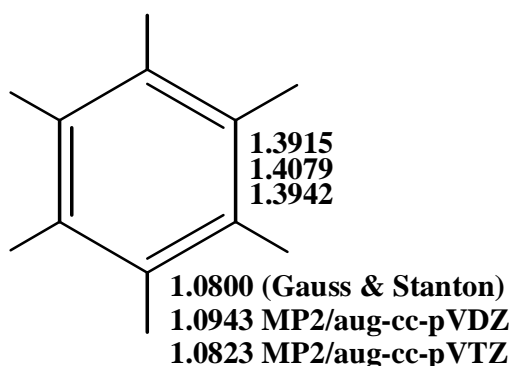


Figure 3.2. Equilibrium geometry of benzene (bond distances in Å).

about 1.2%. The larger aug-cc-pVTZ basis yields much better agreement, reducing errors to about 0.2%.

To our knowledge, no previous study has considered the effect of basis sets as large as aug-cc-pVTZ (828 basis functions) on the optimized geometries of benzene dimer. Potential energy curves for the sandwich, T-shaped, and parallel-displaced configurations were obtained using MP2 with the aug-cc-pVDZ and aug-cc-pVTZ basis sets. These curves, both uncorrected and counterpoise-corrected, are plotted in Figures 3.3-3.5. For clarity, only counterpoise-corrected MP2/aug-cc-pVDZ curves are displayed for the parallel-displaced dimer. One immediately observes a very large difference between the uncorrected and CP-corrected binding energies except near the dissociation limit where the difference is smaller. Near equilibrium for the T-shaped configuration, for example, this difference is 3-4 kcal mol⁻¹ for aug-cc-pVDZ and remains around 1.5 kcal mol⁻¹ for aug-cc-pVTZ. Note also that the minima of the potential energy curves for the uncorrected and CP-corrected curves are significantly different. Although the counterpoise procedure can overcorrect for BSSE, underestimating binding, in this system the CP-corrected values appear to converge faster to the complete basis set limit.

The counterpoise-corrected MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ curves are nearly parallel and give nearly the same equilibrium intermonomer distances. This suggests that, at the MP2 level, smaller basis sets such as aug-cc-pVDZ may be acceptable for the prediction of the equilibrium intermonomer geometries, so long as binding energies are counterpoise-corrected, and so long as the basis set contains adequate polarization and diffuse functions. As shown in Figure 3.5, the sandwich dimer represents a potential energy maximum (saddle point) along the displacement coordinate

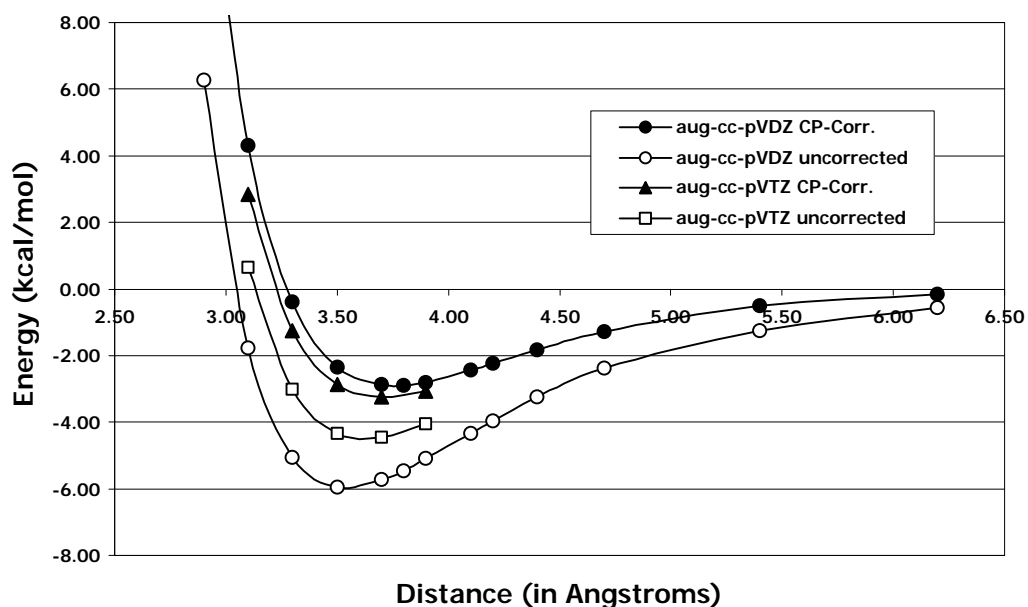


Figure 3.3. Potential energy curves for the sandwich configuration at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory.

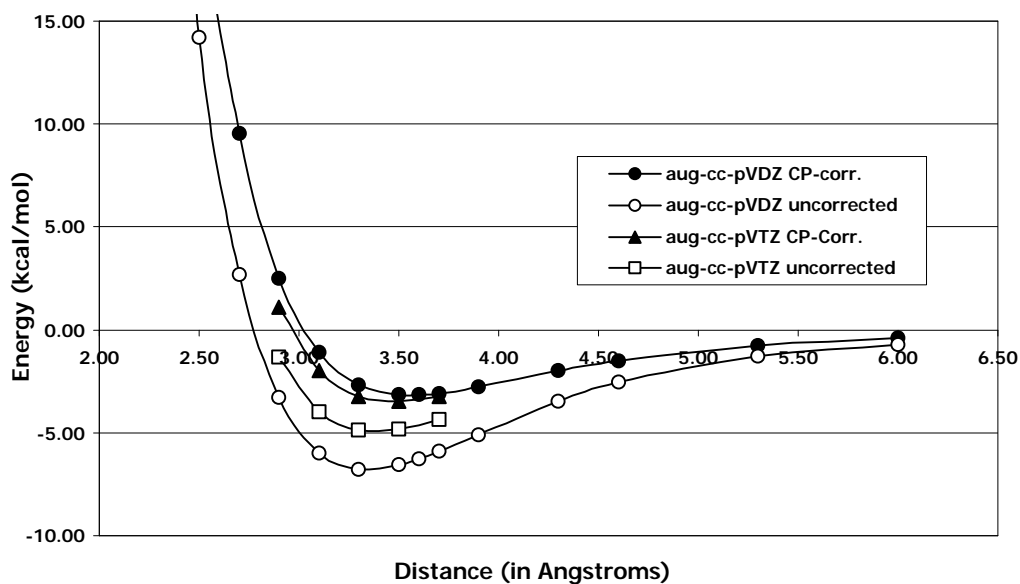


Figure 3.4. Potential energy curves for the T-shaped configuration at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory.

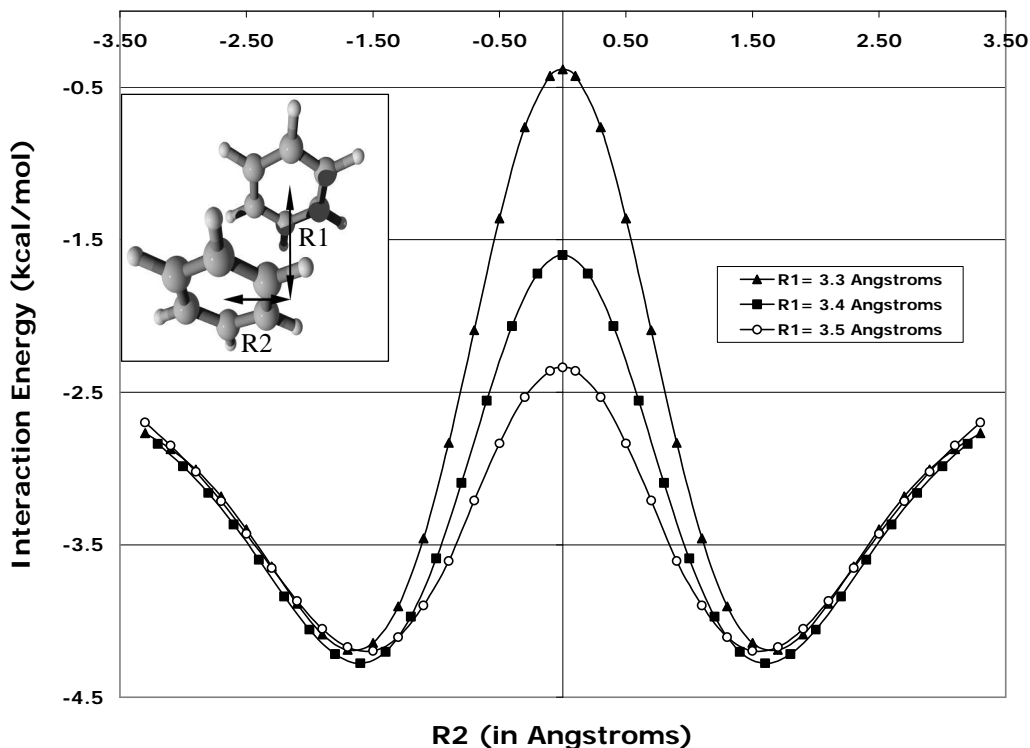


Figure 3.5. Potential energy curves for the parallel-displaced configuration at the (CP-corrected) MP2/aug-cc-pVDZ level of theory.

R2 which connects two equivalent PD configurations. This behavior agrees with the previous results of Jaffe and Smith.³² Whether the sandwich dimer represents a transition state or a higher-order saddle point cannot be determined with certainty on the basis of the present analysis.

The equilibrium dimer intermonomer distances are reported in Table 3.1. For the sandwich and parallel-displaced configurations, the MP2 optimized geometries are very similar for the aug-cc-pVDZ and aug-cc-pVTZ basis sets. Our value of 3.4 Å for the vertical separation between planes in the PD configuration agrees well with the observation³⁴ that in crystals many aromatic molecules form stacks with approximately

parallel molecular planes separated by 3.3-3.6 Å. For the T-shaped configuration, results with both basis sets are in good agreement with the microwave results of Arunan and Gutowsky,³⁸ who found a distance of 4.96 Å between the centers of mass for the gas-phase benzene dimer. Our values are also similar to the 5.05 Å mean distance between phenyl ring centroids for interacting aromatic side chains in proteins.⁴ In agreement with previous work,³² we found that rotating one monomer with respect to the other made essentially no difference to the interaction energy; at the aug-cc-pVDZ level, rotating one monomer by 30° about the axes R and R1 in Figure 3.1 produced energy changes of less than 0.1 kcal mol⁻¹. For the parallel-displaced benzene dimer configuration, the change was a very small 0.02 kcal mol⁻¹.

Table 3.1. Benzene Dimer Geometries (R in Å)^a

	Method	Basis	S	T	PD	
					R1	R2
Hobza et al. ^b	MP2	DZ+2P	3.9	5.0	3.5	1.6
Hobza et al. ^c	CCSD(T)	cc-pVDZ'	4.1	5.1	3.6	1.8
Jaffe and Smith ^d	MP2	6-311G(2d,2p)	4.1	5.1	3.6	1.8
This work	MP2	aug-cc-pVDZ	3.8	5.0 ^e	3.4	1.6
		aug-cc-pVTZ	3.7	4.9 ^f	3.4	1.6
Arunan and Gutowsky ^g	expt.			4.96		

^a All intermonomer parameters, in angstroms, obtained using rigid monomers. ^b Reference 33 using experimental monomer geometry. ^c Reference 46 using experimental monomer geometry. ^d Reference 32 using MP2/6-311G(2d,2p) monomer geometry. ^e Actual distance used was 5.0079 Å. ^f Actual distance used was 4.8942 Å. ^g Reference 38.

The present MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ geometries are in close agreement with the MP2/6-311G(2d,2p) results of Jaffe and Smith³² (see Table 3.1 above), suggesting again that CP-corrected MP2 geometries for benzene dimer are

relatively insensitive to improvements in the basis set beyond polarized double- ζ with diffuse functions. The CCSD(T) results of Hobza and co-workers⁴⁶ with a modified cc-pVDZ basis set are somewhat similar to MP2 results using a similar quality DZ+2P basis,³³ with the intermonomer distances being $\sim 0.1\text{-}0.2$ Å larger for the CCSD(T) results than the MP2 ones.

As it is clear from examining Figures 3.3 and 3.4, the binding energy is much harder to converge (with respect to basis set or theoretical method) than the optimized geometry. Several studies have investigated the effect of basis sets on the binding energy of benzene dimer. Tsuzuki et al.⁴⁷ found that the magnitude of the MP2 interaction energy of the sandwich configuration increased significantly going from the 6-31G* basis set to 6-311G(3d,3p) as each additional polarization function was added. Hobza et al.⁴⁶ and Jaffe and Smith³² have shown that diffuse functions also have a significant effect on the binding energy. In light of these studies and other studies on benzene dimer, it is of great interest to estimate the complete basis set limit for binding energies of the benzene dimer. In this work, we have examined the basis set dependence of the MP2 binding energy by comparing the previously mentioned aug-cc-pVDZ and aug-cc-pVTZ results to predictions with the even larger aug-cc-pVQZ basis at our best estimates of the geometry for each dimer configuration. Complete basis set estimates were obtained at these geometries using MP2-R12/A methods and the custom uncontracted Gaussian basis set described in the previous section. These binding energies (along with results from previous studies on benzene dimer) are presented in Table 3.2. At the MP2 level, the aug-cc-pVTZ basis stabilizes the dimer by $\sim 0.3\text{-}0.4$ kcal mol⁻¹ relative to the smaller aug-cc-pVDZ basis. This is a significant improvement and represents around 10% of the overall

MP2/aug-cc-pVTZ binding energies of 3.3-4.7 kcal mol⁻¹. Nevertheless, even the good MP2/aug-cc-pVTZ binding energies are still up to 0.4 kcal mol⁻¹ away from the estimated MP2 CBS limits (obtained using the MP2-R12/A methods), and using the aug-cc-pVQZ

Table 3.2. Binding Energies (kcal mol⁻¹) for Different Configurations of the Benzene Dimer^a

	Method	Basis	S	T	PD
Hobza et al. ^b	MP2	aug-cc-pVDZ	2.56	2.96	3.94
	CCSD(T)	aug-cc-pVDZ	1.12	2.17	2.02
Tsuzuki et al. ^c	MP2	6-311G**	1.30	2.12	
		aug(d)6-311G**	2.58		
	CCSD(T)	6-311G**	-0.02	1.40	
Jaffe and Smith ^d	MP2	aug(d)6-311G**	1.02		
		6-311+G(2d,p)	2.47	2.87	3.79
		6-311G(2df,p)	2.10	2.79	3.36
This work	MP2 ^e	aug-cc-pVDZ	2.90	3.16	4.28
	MP2 ^e	aug-cc-pVTZ	3.26	3.46	4.67
	MP2 ^f	aug-cc-pVQZ	3.37	3.54	4.79
	MP2-R12/A ^f	custom	3.64	3.63	4.95
	Δ CCSD(T)	aug-cc-pVDZ	-1.83	-0.89	-2.18
	estd CBS CCSD(T) D_e		1.81	2.74	2.78
	MP2 Δ ZPVE	cc-pVDZ	0.18	-0.35	-0.04
estd CBS CCSD(T) D_0		1.99	2.39	2.74	

^a All energies are counterpoise-corrected. ^b Reference 46 using MP2/DZ+2P dimer geometry with experimental monomer structure. ^c Reference 48 using MP2/6-31G* monomer geometries. ^d Reference 32 with MP2/6-311G(2d,2p) monomer and dimer geometries. ^e Geometry optimized, (monomer kept rigid) at each level of theory (see Table 3.1). ^f At the MP2/aug-cc-pVTZ optimized dimer geometry using experimentally deduced monomer geometries from ref. 55.

basis decreases this gap only to ~ 0.3 kcal mol⁻¹. As seen in Table 3.2, the MP2 interaction energies reported in the literature using smaller basis sets are quite far from the CBS limit and do not show convergence towards a certain value. The effect of diffuse functions was specifically examined by performing MP2 computations for the S configuration with the cc-pVDZ and cc-pVTZ basis sets (i.e. the diffuse functions were dropped). The resulting

binding energies (0.81 and 2.47 kcal mol⁻¹, respectively) show that adding a set of diffuse functions can be more important than going to the next larger basis in the cc-pVXZ series.

To elucidate more clearly on the effect of basis set on the MP2 interaction energies, Figures 3.6 and 3.7 present the Hartree-Fock and MP2 correlation energy contributions to the interaction energies as a function of basis set. The difference in height between the two bars for each basis represents the value of the counterpoise correction. By splitting the MP2 energies into their Hartree-Fock and correlation contributions, we see that the attraction arises purely from electron correlation at these geometries; the Hartree-Fock binding energies (Figure 3.6) are all negative (repulsive). The Hartree-Fock energies contain the dominant electrostatic and induction contributions, as well as short-range exchange repulsion. The T-shaped Hartree-Fock energies are the least repulsive of the three configurations because of the favorable quadrupole-quadrupole interactions. We can also see from Figure 3.6 that the basis set superposition error becomes very small for Hartree-Fock with the aug-cc-pVQZ basis, reflecting the high quality of such basis. The CP-corrected binding energies converge rapidly with respect to basis set.

The correlation component of the MP2 interaction energies (Figure 3.7) includes dispersion effects as well as correlation corrections to the electrostatic and induction contributions. These correlation contributions to the binding energies are all negative (attractive), and, in contrast to the Hartree-Fock components (Figure 3.6) they do not converge rapidly with respect to basis set; a significant CP-correction remains even for the enormous aug-cc-pVQZ basis. Moreover, the MP2/aug-cc-pVQZ computations needs

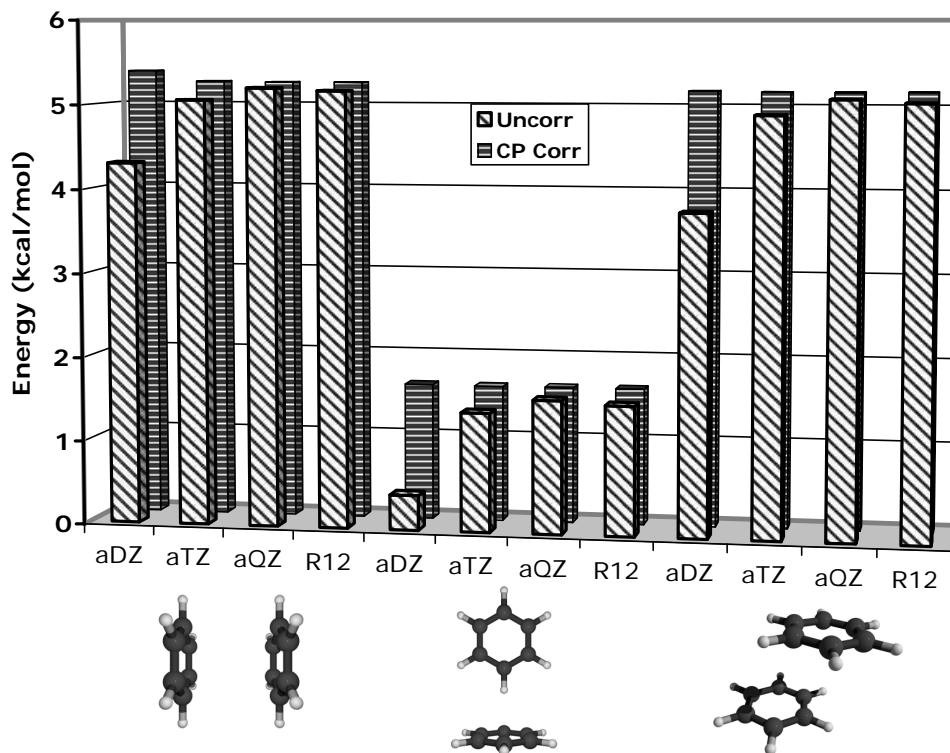


Figure 3.6. Hartree-Fock binding energies for each dimer structure as a function of basis set. All computations were performed at the same best estimate geometry for each configuration. All energies are negative (repulsive). Labels aXZ denote the aug-cc-pVXZ basis sets.

more computer time than the MP2-R12/A values they approach, strongly suggesting that R12 methods can be more affordable than the current CBS extrapolation methods. Correlation favors the S and PD configurations, which have larger dispersion energies than the T-shaped. Although one might expect the S configuration to have the largest dispersion interaction, the correlation contribution is actually more favorable for the PD geometry; this remains true even after improvements in the treatment of electron correlation (see the discussion below).

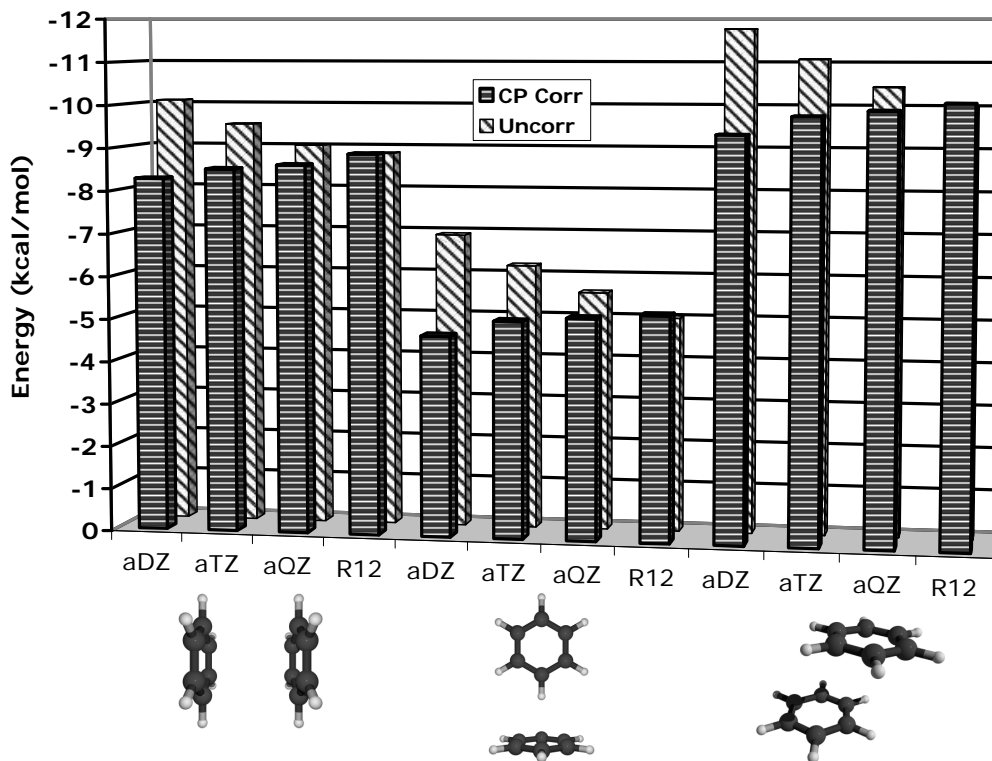


Figure 3.7. MP2 electron correlation energy contributions to binding energies for each dimer structure as a function of basis set. The total MP2 binding energies are obtained by adding these values to the Hartree-Fock contributions in the previous figure. All computations were performed at the same best estimate geometry for each configuration. Labels aXZ denote the aug-cc-pVXZ basis sets.

Hobza et al.⁴⁶ have reported that MP2 significantly overestimates the stabilization energy compared to CCSD(T), with overestimation of 30% for the T-shaped and 92% for the sandwich structure for the largest basis set they considered. Jaffe and Smith³² also reported that MP2 overestimated the electron correlation energy compared to MP4(SDQ) and MP4(SDTQ).

To better account for electron correlation, we performed CCSD(T) computations using the aug-cc-pVDZ basis set. The difference between the CCSD(T) and MP2 binding energies, denoted $\Delta\text{CCSD(T)}$ in Table 3.2, was assumed to be relatively insensitive to

basis set, and was added to the aug-cc-pVQZ Hartree-Fock and MP2-R12/A correlation energy results to estimate the complete basis set limit for CCSD(T). It is clear from the table that $\Delta\text{CCSD(T)}$ is very large, ranging from $-2.2 \text{ kcal mol}^{-1}$ for the parallel-displaced configuration to $-0.9 \text{ kcal mol}^{-1}$ for the T-shaped configuration. The final CBS CCSD(T) estimates of D_e should be within a few tenths of a kilocalorie per mole of the ab initio limit.

To gauge the size of the zero-point vibrational energy (ZPVE) corrections, vibrational frequencies were obtained for fully optimized geometries at the MP2/cc-pVDZ level of theory. At this level, imaginary frequencies are found for each configuration. The sandwich configuration has two imaginary frequencies of $51i \text{ cm}^{-1}$, each corresponding to a planar distortion of each benzene ring. The T-shaped configuration has a weak imaginary mode ($24i \text{ cm}^{-1}$) corresponding to the rocking of benzene about the line joining the ring centers of mass. The PD-configuration has an imaginary frequency of $79i \text{ cm}^{-1}$ corresponding to a rotation making the two rings nonparallel.

Given the sensitivity of the potential energy surfaces to the theoretical treatment, the MP2/cc-pVDZ level does not seem sufficient to confirm whether these stationary points are actually minima or saddle points. However, the ZPVE corrections should be reasonably well estimated. The ZPVE correction weakens the binding of the T-shaped and parallel-displaced configurations by 0.35 and $0.04 \text{ kcal mol}^{-1}$, respectively, while it strengthens the interaction of the sandwich by $0.18 \text{ kcal mol}^{-1}$. This result for the T-shaped dimer is consistent with an earlier, lower-level estimate⁶⁴ of $0.24 \text{ kcal mol}^{-1}$. The final CBS CCSD(T) estimates of D_e (D_0) predict that the T-shaped and parallel-displaced

configurations are isoenergetic within the expected error bars, with binding energies of 2.7 (2.4) and 2.8 (2.7) kcal mol⁻¹, respectively. The sandwich configuration is several tenths of a kilocalorie per mole higher than the other two configurations. Previous CCSD(T) computations with a modified aug-cc-pVDZ basis set found⁴⁶ the T-shaped and parallel-displaced configurations to be within 0.15 kcal mol⁻¹ of each other, with the T-shaped being more stable ($D_e = 2.17$ kcal mol⁻¹); the sandwich structure at this level was about 1 kcal mol⁻¹ less stable. We have shown that these results are qualitatively correct but basis set effects significantly increase the overall binding energy. In light of the present results, it seems clear that the most commonly cited experimental value⁴⁴ of $D_0 = 1.6 \pm 0.2$ kcal mol⁻¹ is too small. However, an older experimental study by Grover et al.⁴⁵ gives $D_0 = 2.4 \pm 0.4$ kcal mol⁻¹, which agrees well with our new theoretical estimates.

This 2.4-2.7 kcal mol⁻¹ attraction is appreciable and will significantly influence the interaction of phenyl rings in solution or other environments, in addition to other factors such as solvophobic effects. The rather flat potential energy surface, along with the S configuration being the least favorable, is entirely consistent with the observation that interacting pairs of phenylalanines in proteins are found in mostly T- and PD- like configurations, but they are scattered over a wide variety of conformational space with no strongly preferred single orientation.⁵ The benzene dimer itself is expected to be highly fluxional and without a rigid structure, like many van der Waals clusters.⁶⁵

3.4 Conclusions

In this work we have investigated the basis set and electron correlation effects for the simplest aromatic π -stacking system, the benzene dimer. This dimer is very

challenging for electronic structure theory and for experiments attempting to determine its optimum structure.

For constrained geometry optimization, the aug-cc-pVTZ basis set used here is much larger than basis sets employed in previous optimizations. Fortunately, we find the smaller aug-cc-pVDZ basis sufficient to obtain intermonomer distances very near those of aug-cc-pVTZ for the MP2 methods, so long as energies are counterpoise-corrected. The present theoretical geometries for the T-shaped configuration are in good agreement with the experimental data.

The counterpoise correction remains large even for the aug-cc-pVTZ basis, suggesting that even larger basis sets are required for definitive prediction of binding energies in benzene dimer. A new shared memory parallel algorithm has allowed us to perform MP2-R12/A computations for benzene dimer which estimate the complete basis set limit. As can be seen from the binding energies at the MP2 level, we have obtained converged values for the energy where the binding energy converge towards the complete basis set MP2 binding energies evaluated using the MP2-R12/A method. Combined with a correction for the difference between CCSD(T) and MP2 determined using the aug-cc-pVDZ basis with 384 basis functions, these results yield complete basis set (CBS) CCSD(T) estimates which should be within a few tenths of a kilocalorie per mole of the ab initio limit.

Our best estimates of the binding energy indicate that the T-shaped and parallel-displaced configurations are nearly isoenergetic, with D_e (D_0) values of 2.7 (2.4) and 2.8 (2.7) kcal mol⁻¹, respectively. The sandwich structure is not an optimum structure and has a somewhat higher D_e (D_0) of 1.8 (2.0) kcal mol⁻¹, respectively. These results indicate

that the experimental binding energy of Krause et al.⁴⁴ ($D_0 = 1.6 \pm 0.2$ kcal mol⁻¹) must be too small, but they support the older and less cited results of Grover et al. ($D_0 = 2.4 \pm 0.4$ kcal mol⁻¹). The preferred configurations and the rather flat potential energy surface are consistent with a variety of experimental observations of π - π interactions.

CHAPTER 4

UNEXPECTED SUBSTITUENT EFFECTS IN FACE-TO-FACE π -STACKING INTERACTIONS

In this work, state-of-the-art electronic structure methods have been applied to obtain the first high-quality theoretical results for substituent effects in π -stacking interactions. The sandwich configurations of benzene dimer, benzene-phenol, benzene-toluene, benzene-fluorobenzene, and benzene-benzonitrile have been studied using correlation consistent basis sets augmented by multiple diffuse functions, namely aug-cc-pVDZ and aug-cc-pVTZ at the second-order (MP2) level. Coupled-cluster singles and doubles with perturbative triples [CCSD(T)] computations were also performed on the dimers and combined with the above MP2 calculations to obtain an estimate of the CCSD(T)/aug-cc-pVTZ binding energies, which should be accurate to within several tenths of a kilocalorie per mole.

4.1 Introduction

As pointed out in the past chapter, the simplest prototype of π - π interactions, the benzene dimer, has been the subject of great experimental and theoretical interest.^{50,66-68} Using explicitly correlated MP2-R12/A wave functions,¹⁵ we estimated (in the past chapter) the complete basis set limit for second-order Møller-Plesset perturbation theory (MP2) gas-phase binding energies of benzene dimer.⁶⁷ After correction for higher-order correlation effects via coupled-cluster theory including perturbative triple substitutions

[CCSD(T)], the resulting binding energies should approach the ab initio limit within a few tenths of a kcal mol⁻¹. The results demonstrate that the perpendicular T-shaped configuration and the parallel-displaced (offset-stacked) configuration are nearly isoenergetic ($D_e = 2.7$ and 2.8 kcal mol⁻¹, respectively), while the face-to-face sandwich configuration is a saddle point on the potential energy surface and is significantly higher in energy (1.8 kcal mol⁻¹). However, substituents may substantially alter the energy landscape: for the toluene dimer, Kollman and co-workers have shown that the sandwich stacked configuration appear to be preferred over the T-shaped one, and this preference is observed both in the gas-phase and in aqueous solution.⁶⁹

Very little is known about substituent effects in π - π interactions.⁷⁰⁻⁷² Cozzi, Siegel, and co-workers have used NMR techniques to study substituent effects in 1,8-diarylnaphthalenes which force a nearly face-to-face arrangement. The barrier to rotation about the aryl-naphthyl bond, which should increase as the π - π interaction becomes more favorable, showed a linear relationship with the sum of the Hammett parameters σ_{para} of the substituents.^{70,71} The Hammett parameter was used as an indicator of the strength of electron-donating and electron-withdrawing groups. These results suggest that electron-withdrawing substituents stabilize the transition state by decreasing the repulsion between the π electrons on each aryl ring, while electron-donating substituents destabilize the transition state by increasing the repulsion between the two π systems, and that the electron donating or withdrawing character of the substituents is reasonably described by the σ parameter. Such an analysis is consistent with the Hunter-Sanders rules,² which make qualitative predictions of π - π interactions by focusing on electrostatics and considering the π electron cloud to have a negative charge and the σ framework to have a

positive charge. We will revisit the Hunter-Sanders rules and provide additional details about such rules in the next chapter of this thesis.

4.2 Theoretical Methods

Here we present preliminary results from the first ab initio study that focuses on the effect of different substituents in face-to-face (sandwich) π - π interactions. Dimers of benzene with substituted benzenes have been considered, with substituents OH, CH₃, F, and CN (see Figure 4.1). The centers of the two aromatic rings were aligned in an eclipsed configuration. Monomer geometries of benzene and monosubstituted benzenes were optimized using MP2 with the aug-cc-pVDZ and aug-cc-pVTZ basis sets, where the aug- prefix denotes one set of diffuse functions for every angular momentum in the basis.

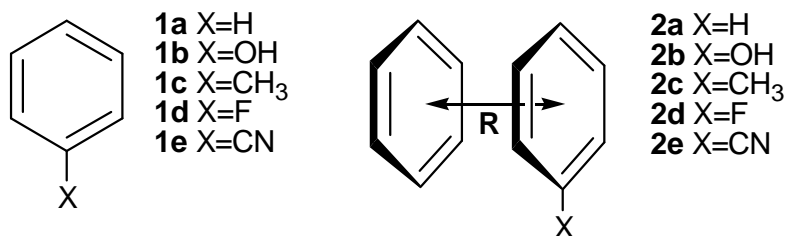


Figure 4.1. Sandwich configurations of benzene dimer and benzene-monosubstituted benzene dimers.

With frozen monomer geometries, the distance between the centers of the rings was optimized with the aug-cc-pVDZ and aug-cc-pVTZ basis sets and counterpoise-corrected MP2 binding energies. Results using the larger, aug-cc-pVTZ basis set are taken as more reliable, and these are corrected for higher-order correlation effects by CCSD(T) using an

aug-cc-pVDZ basis in which diffuse functions have been removed from hydrogen, and diffuse d functions have been removed from carbon (denoted aug-cc-pVDZ'). The prohibitively high computational cost of CCSD(T) computations with a full aug-cc-pVDZ basis precluded the use of such basis. In addition, some of the dimers studied have a low symmetry (C_1 and C_s) making the CCSD(T)/aug-cc-pVDZ computations not feasible.

The above-mentioned procedure, which approximates aug-cc-pVTZ CCSD(T), underestimates the binding energy for benzene dimer by approximately $0.2 \text{ kcal mol}^{-1}$ when compared to the more complete MP2-R12/A + Δ CCSD(T) approach,⁶⁷ which unfortunately is currently impractical for the lower-symmetry substituted dimers. We anticipate that remaining errors due to incompleteness of the basis set or correlation treatment are small and will approximately cancel when considering differences among the dimers. All the results reported in the present study were obtained using the PSI 3.2⁵⁸ and MOLPRO⁷³ programs.

4.3 Results and Discussion

Theoretical results are summarized in Table 4.1. The optimized distance between monomers, R , is relatively insensitive to the basis set, but using a larger basis set makes binding significantly more favorable. The Δ CCSD(T) corrections are very large, and they account for the known overbinding of the van der Waals complexes by the MP2 method. The most striking result from Table 4.1 is that *all* of the substituted dimers bind more strongly than benzene dimer, even though these substituents are typically characterized as

Table 4.1. Interfragment Distances (\AA) and Binding Energies (kcal mol^{-1}) for Face-to-Face Dimers of Benzene with Substituted Benzenes^a

Dimer	Method	R ^b	ΔE_{int}
Benzene-benzene	MP2/aug-cc-pVDZ ^c	3.80	-2.90
	MP2/aug-cc-pVTZ ^c	3.70	-3.26
	MP2/aug-cc-pVQZ ^{c,d}		-3.37
	MP2-R12/A ^{c,d}		-3.64
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^e		1.65
	estd. $\text{CCSD(T)}/\text{aug-cc-pVTZ}$		-1.61
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^{c,d}		1.83
Benzene-phenol	estd. $\text{CCSD(T)}/\text{CBS}$ ^{c,d}		-1.81
	MP2/aug-cc-pVDZ ^e	3.70	-3.40
	MP2/aug-cc-pVTZ ^e	3.60	-3.75
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^f		1.90
Benzene-toluene	estd. $\text{CCSD(T)}/\text{aug-cc-pVTZ}$ ^f		-1.85
	MP2/aug-cc-pVDZ ^e	3.70	-3.58
	MP2/aug-cc-pVTZ ^e	3.65	-3.94
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^f		1.90
Benzene-fluorobenzene	estd. $\text{CCSD(T)}/\text{aug-cc-pVTZ}$ ^f		-2.04
	MP2/aug-cc-pVDZ ^e	3.70	-3.50
	MP2/aug-cc-pVTZ ^e	3.70	-3.81
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^f		1.61
Benzene-benzonitrile	estd. $\text{CCSD(T)}/\text{aug-cc-pVTZ}$ ^f		-2.20
	MP2/aug-cc-pVDZ ^e	3.70	-4.49
	MP2/aug-cc-pVTZ ^e	3.60	-4.86
	$\Delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ ^f		2.07
	estd. $\text{CCSD(T)}/\text{aug-cc-pVTZ}$ ^f		-2.79

^a All computations include the counterpoise correction. ^b Distance between centers of rings optimized with frozen monomers. ^c Reference 67. ^d Using best estimate of monomer geometry and MP2/aug-cc-pVTZ inter-fragment distance. ^e Geometry optimized (monomers kept rigid) at each level of theory. ^f Using MP2/aug-cc-pVTZ monomer and inter-fragment geometry.

ranging from strongly electron donating (OH) to strongly electron withdrawing (CN).

This contrasts with the previously mentioned results of Cozzi, Siegel, and co-workers for biaryl-naphthalenes, which correlated with σ parameters.

It is tempting to ascribe the increased attraction for all substituents considered to dipole-induced dipole interactions, which are absent in benzene dimer. The differences in binding energies, however, do not correlate very well with the dipole moments of the

substituted monomers (toluene, 0.375; phenol, 1.224; fluorobenzene, 1.60; and benzonitrile, 4.18 D).⁷⁴ An alternative approach to understanding the electrostatic contribution to binding is afforded by the electrostatic potentials of the monomers **1a-e** (Figure 4.2), obtained using the SPARTAN program with a density isosurface of 0.002 electrons au⁻³.⁷⁵

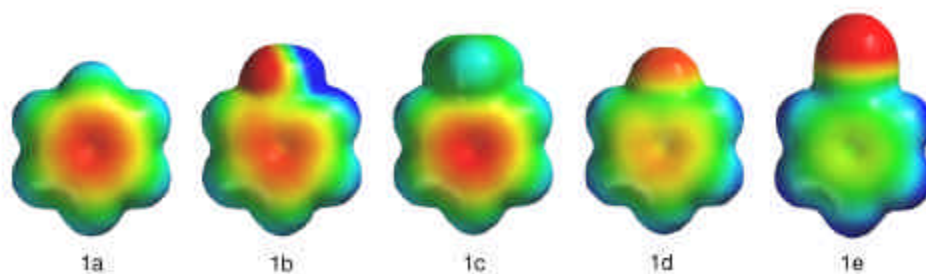


Figure 4.2. Electrostatic potentials computed using Hartree-Fock and a 6-31G* basis set with a scale of -25 to +25 kcal mol⁻¹. Potentials using B3LYP/6-31G* appear very similar.

According to the Hunter-Sanders rules, the most important consideration for a face-to-face sandwich configuration would be the negative π -electron charge at the ring center. Figure 4.2 indicates that this charge is greatest for benzene, toluene, and phenol, and followed by fluorobenzene and benzonitrile. The observed order, **2a** < **2b** < **2c** < **2d** < **2e** shows benzene to bind somewhat more weakly and toluene somewhat more strongly compared to phenol. The binding sequence **2a** < **2b** < **2c** is consistent with the order of the polarizabilities,⁷⁶ and the expected order of increasing dispersion interactions. Nevertheless, more robust conclusions cannot be drawn without actually decomposing the energy into its unique components and comparing the contributions of these different

components to the overall binding energy, and to the observed differences in binding energies for the substituted dimers.

4.4 Conclusions

In this work we have seen that the electrostatic potentials do not necessarily correspond to the classical electron donating or withdrawing effects observed in electrophilic aromatic substitution reactions. OH is a strongly activating (electron-donating) substituent in electrophilic aromatic substitutions, yet phenol displays no more π -electron charge in the ring than benzene. This suggests that the electron donating effect of OH is seen in the transition state and not the ground state of the substitution reaction. This observation has been made by Dougherty and co-workers in studies of cation- π binding.⁷⁷ For the different substituents examined in this work, we have seen that *all* substituted dimers bind more strongly than benzene dimer, with benzene-benzonitrile binding the most strongly.

We expect that an increased understanding of substituent effects in π - π interactions will aid molecular design efforts. In the next chapter, we report on additional dimer configurations, and provide a more detailed analysis of the binding energy for benzene dimer and different benzene-monosubstituted benzene dimers using energy decomposition analysis.

CHAPTER 5

SUBSTITUENT EFFECTS IN π - π INTERACTIONS: SANDWICH AND T-SHAPED CONFIGURATIONS

In this work, sandwich and T-shaped configurations of benzene dimer, benzene-phenol, benzene-toluene, benzene-fluorobenzene and benzene-benzonitrile are studied by coupled-cluster theory to elucidate how substituents tune π - π interactions. We find that all substituted sandwich dimers bind more strongly than benzene dimer, whereas dimers in two different T-shaped configurations bind more or less favorably depending on the substituent. Energy decomposition analysis using symmetry-adapted perturbation theory (SAPT) indicates that electrostatic, dispersion, induction, and exchange-repulsion contributions are all significant to the overall binding energies, and all but induction are important in determining relative energies. Models of π - π interactions based solely on electrostatics, such as the Hunter-Sanders rules, do not seem capable of explaining the energetic ordering of the dimers considered. Such models do not take into account the roles of dispersion, inductions, and exchange-repulsion in tuning these π - π systems.

5.1 Introduction

Noncovalent interactions are of pivotal importance in many areas of chemistry, biology, and material science, and π - π interactions in particular are fundamental to many supramolecular organization and recognition processes.⁶⁶ These interactions play a key role in phenomena as diverse as base-base interactions of DNA,⁷ side-chain interactions

in proteins,⁴ host-guest complexation,¹⁰ self-assembly based on synthetic molecules,^{12,13} and intercalation of certain drugs into DNA.⁶ As mentioned in the previous chapters, despite a wide body of theoretical and experimental studies addressing the importance of π - π interactions, a clear picture of their strength and geometrical preferences presents a challenge for both experiment and theory due to the weakness of the interactions and to the shallowness of the potential energy surfaces. However, any advances in rational supramolecular design will require a detailed understanding of these interactions and how substituent effects may tune them.

In previous work,⁶⁷ we provided the first definitive study of the simplest prototype of aromatic π - π interactions, the benzene dimer. Using explicitly correlated MP2-R12/A¹⁵ wavefunctions, we estimated the complete basis set limit gas-phase binding energies at the second-order Møller-Plesset perturbation theory (MP2) level. After accounting for higher-order electron correlation via coupled-cluster theory with singles, doubles, and perturbative triples [CCSD(T)],⁵² the resulting binding energies should be within a few tenths of one kcal mol⁻¹ of the ab initio limit. The estimated complete basis set CCSD(T) estimates of D_e (D_0) predict⁶⁷ that the T-shaped and parallel-displaced configurations are nearly isoenergetic, with binding energies of 2.7 (2.4) and 2.8 (2.7) kcal mol⁻¹, respectively. The face-to-face sandwich configuration is about 1 kcal mol⁻¹ less stable. These results show that the commonly cited experimental binding energy of Krause and co-workers ($D_0 = 1.6 \pm 0.2$ kcal mol⁻¹) is too low by about one kcal mol⁻¹. That interacting pairs of phenylalanines in proteins are generally found in orientations similar to the T-shaped or parallel-displaced configurations^{5,71,78} is totally

consistent with our theoretical predictions that these two configurations are nearly isoenergetic.

Substituents, however, may alter the energy landscape. For toluene dimer in both aqueous solution and the gas phase, two stacked configurations are predicted to be more stable than the T-shaped configuration.⁶⁹ Very little is known about substituent effects in π - π interactions, either theoretically or experimentally. A few studies have used NMR techniques to examine π - π interactions in substituted aromatics. As mentioned previously in the discussion of substituted sandwich dimers, Cozzi, Siegel, and co-workers^{70,71,79,80} have measured barrier heights to rotation in substituted 1,8-diarylnaphthalenes featuring a face-to-face (sandwich) arrangement. Rashkin and Waters recently reported experiments on substituent effects in a model system with an offset-stacked (parallel-displaced) configuration.⁷² Wilcox and co-workers devised a “molecular torsion balance” model system to examine substituent effects on perpendicular (T-shaped) π - π interactions.^{81,82} Hunter and co-workers have also examined T-shaped configurations using chemical double-mutant cycles and molecular zipper complexes.^{83,84} Because none of these experiments were performed in the gas phase, it is difficult to decouple the intrinsic binding energy from contributions due to the solvent or environment, which will change from system to system.⁸⁵ Additionally, due to secondary intramolecular interactions or steric constraints, the model system itself may complicate the interpretation of results.⁸⁶ Although theory can examine the bare interactions directly, experience with the benzene dimer⁶⁷ indicates that this is challenging, because coupled-cluster theory and augmented triple-zeta or better basis sets are required for reliable total binding energies. Systematic

theoretical studies of substituent effects in π - π interactions seem to be entirely absent, apart from a double-zeta MP2 study of T-shaped configurations by Hong and Kim.⁸⁷

In a recently published letter,⁸⁸ we presented preliminary results from the first ab initio study of substituent effects in face-to-face (sandwich) π - π interactions. Dimers of benzene with monosubstituted benzenes were considered, with substituents OH, CH₃, F, and CN (2 in Figure 5.1). The most surprising result was that *all* substituted dimers bind more strongly than benzene dimer, regardless whether the substituents are considered electron donating (OH, CH₃) or electron withdrawing (CN, F), in apparent contradiction to the Hunter-Sanders model².

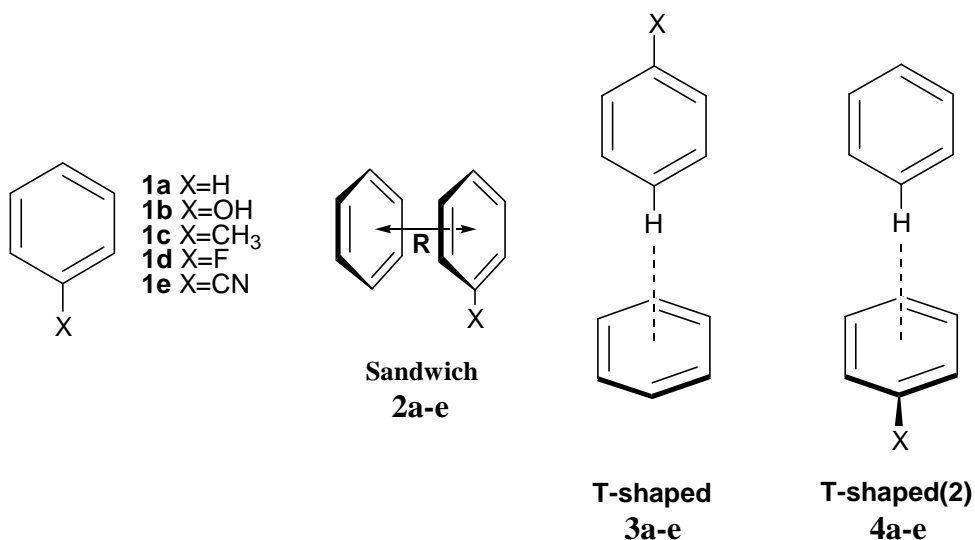


Figure 5.1. The Sandwich and two T-shaped configurations of benzene dimer and benzene-monosubstituted benzene dimers.

While acknowledging that dispersion has a major effect on the magnitude of π - π interactions, the Hunter-Sanders model makes qualitative predictions assuming that

geometries and substituent effects are determined by electrostatic interactions, with the π electron clouds being negatively charged and the σ framework being positively charged. Other effects, such as induction and short-range repulsion, are ignored. According to this model, electron-donating groups such as OH should increase the negative charge of the electron clouds on the substituted monomer, leading to increased repulsion in the sandwich dimer. This expectation does not fit our high-level theoretical results. A preliminary analysis of our data for the sandwich dimers suggested that it is not possible to understand the trends in binding based on electrostatic effects alone.⁸⁸

In this work we explore how substituents affect the binding of the T-shaped configurations (**3** and **4** in Figure 5.1). When substituting the upper ring in the T-shaped configuration (**3**), the substituent is placed only in the position para to the other benzene to minimize complications from direct interactions between the substituent and the other benzene. Likewise, when the lower ring is substituted (**4**), the substituent is placed as far away from the edges of the upper benzene as possible. We also extend our preliminary report on the sandwich dimers (**2**) by incorporating CCSD(T) corrections into the geometry optimization. This has a non-negligible effect on the total binding energies. Using an additive correction scheme, we estimate interfragment distances and binding energies at the highly reliable CCSD(T)/aug-cc-pVTZ level of theory. Lower levels of theory are also considered to investigate which ones are capable of accurately reproducing changes in binding due to substituents. Surprisingly, even though total binding energies are very sensitive to basis set and electron correlation, the relative energies for different substituted dimers are not. To further investigate the relative importance of electrostatic, dispersion, induction, and exchange-repulsion energies we

have decomposed binding energies into these components using symmetry-adapted perturbation theory (SAPT).^{19,20,28} This analysis confirms that electrostatic interactions alone are not sufficient to predict the correct energetic ordering of all the dimers.

5.2 Theoretical Methods

5.2.1 Supermolecular Approach

Most computations were performed using Dunning's augmented correlation-consistent polarized valence basis sets,⁵³ specifically the double- and triple-zeta basis sets aug-cc-pVDZ and aug-cc-pVTZ. The aug- prefix denotes the presence of one set of diffuse functions for each angular momentum in the basis; this adds a considerable number of diffuse functions to the standard, unaugmented cc-pVXZ basis sets. In a previous study of benzene dimer,⁶⁷ we found that augmenting the cc-pVDZ basis with diffuse functions was more important to the binding energy than increasing the basis set to cc-pVTZ. In that study we also explored basis set convergence by using basis sets as large as quadruple-zeta aug-cc-pVQZ (1512 basis functions), and complete basis set estimates were obtained using the explicitly correlated MP2-R12/A method.¹⁵ Unfortunately, such sophisticated computations are not yet feasible for all of the lower-symmetry substituted dimers in the present study due to their prohibitively high computational cost. Our previous benchmark results are reproduced here for comparison to the levels of theory presently employed.

In the current study, monomers were fully optimized at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory. For toluene, there is almost free rotation of the methyl group, so we chose a one H up, 2 H down, C_s configuration. Dimer geometries

were determined by optimizing the distance between the centers of the rings while keeping the monomers rigid. Changes in relative orientations between the two aromatic rings were not considered in this study. For benzene dimer, rotation of one ring about the axis joining the centers of mass of the two rings has very little effect on the energy.⁶⁷ For the present substituted benzenes, changes in the relative orientation of the two rings will lead to larger energy differences, depending on the substituent. For example, preliminary data suggest that rotating the upper benzene in **4** so that the upper benzene is coplanar with the C-X bond of the substituent leads to direct interactions between the substituent and the meta hydrogen of the upper benzene which are worth a few tenths of one kcal mol⁻¹. We hope to explore these additional complications in future work.

MP2 interaction energies using the larger, aug-cc-pVTZ basis set were approximately corrected for higher-order electron correlation effects by adding the difference between the CCSD(T) and MP2 energies as computed using a modified aug-cc-pVDZ basis, denoted aug-cc-pVDZ', which lacks diffuse functions on hydrogen and diffuse *d* functions on carbon. This provides an estimate of the CCSD(T)/aug-cc-pVTZ binding energies and interfragment distances. The counterpoise (CP) correction of Boys and Bernardi⁵⁴ was applied in all calculations to account for basis set superposition error because our previous work demonstrates that CP-corrected MP2 energies converge more quickly to the complete basis set limit for π - π interactions.⁶⁷ Core orbitals were constrained to remain doubly occupied in all correlated wavefunctions. All supermolecular results in the present study were obtained using the PSI 3.2⁵⁸ and MOLPRO⁷³ programs.

5.2.2 Symmetry-Adapted Perturbation Theory (SAPT) Approach

As mentioned in chapter 2, symmetry-adapted perturbation theory (SAPT)^{19,20} is used to analyze the interaction energy in terms of physically meaningful components such as electrostatic, induction, dispersion, and exchange energies. Using the original notation of Jeziorski and coworkers,³⁰ the SAPT interaction energy can be written in terms of its Hartree-Fock and correlation components as

$$E_{\text{int}} = E_{\text{int}}^{\text{HF}} + E_{\text{int}}^{\text{CORR}} \quad [5.1]$$

where $E_{\text{int}}^{\text{HF}}$ represents lowest-order corrections that be can identified as describing interactions at the Hartree-Fock level. $E_{\text{int}}^{\text{HF}}$ is represent by

$$E_{\text{int}}^{\text{HF}} = E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)} + E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + \mathbf{d}E_{\text{int,resp}}^{\text{HF}} \quad [5.2]$$

The superscripts (ab) denote orders in perturbation theory with respect to operators V and W , respectively.

In this work, we have employed the SAPT2 approach, in which the correlated portion of the interaction energy is nearly equivalent to the supermolecular MP2 correlation energy and can be represented as

$$E_{\text{int}}^{\text{CORR}} = E_{\text{elst,resp}}^{(12)} + E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)} + {}^tE_{\text{ind}}^{(22)} + {}^tE_{\text{exch-ind}}^{(22)} + E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)} \quad [5.3]$$

where, as indicated in a previous chapter, ${}^tE_{\text{ind}}^{(22)}$ represents the part of $E_{\text{ind}}^{(22)}$ that is not included in $E_{\text{ind,resp}}^{(20)}$, and ${}^tE_{\text{exch-ind}}^{(22)}$ is approximated as

$${}^tE_{\text{exch-ind}}^{(22)} \approx E_{\text{exch-ind,resp}}^{(20)} \frac{{}^tE_{\text{ind}}^{(22)}}{E_{\text{ind,resp}}^{(20)}} \quad [5.4]$$

All SAPT calculations reported here have been carried out using the above-mentioned aug-cc-pVDZ' basis set with the MP2/aug-cc-pVTZ optimized monomer

geometries. For the dimers considered in this study, the aug-cc-pVDZ' basis set ranges in size from 276 to 307 basis functions; the very high computational cost of the SAPT procedure precludes the use of a larger basis set. SAPT computations were performed using the SAPT2002 program.³⁰

5.3 Results and Discussion

For more clarity, in this section we will first present results from our supermolecular calculations, and then present results from our symmetry-adapted perturbation theory (SAPT) calculations. This, we believe, will provide a better insight into the utility of these two different approaches in energy analysis.

5.3.1 Supermolecular Approach

Theoretical results for binding energies and optimum intermonomer distances are summarized in Table 5.1. The estimated CCSD(T)/aug-cc-pVTZ results show that all of the substituted sandwich dimers are bound more strongly than benzene dimer, confirming our earlier report on the sandwiches.⁸⁸ While the OH, CH₃, and F substituents increase binding in the sandwich by 0.4-0.5 kcal mol⁻¹ at the best level of theory, CN has a much larger effect of increasing the binding by about 1.3 kcal mol⁻¹. The substituted T-shaped dimers **3b-e** and **4b-e**, by contrast, show both increases and decreases in binding relative to benzene dimer, depending on the substituent. Changes in binding due to substitution are smaller for dimers **3** and **4** than for the sandwiches **2**, but once again CN has by far the largest effect of all substituents examined.

Table 5.1. Interaction Energies (in kcal mol⁻¹) for Various Dimers^a

X	Method	Sandwich		T-shaped		T-shaped(2)	
		R ^b	ΔE_{int}	R ^b	ΔE_{int}	R ^b	ΔE_{int}
H	MP2/aug-cc-pVDZ ^c	3.80	-2.90	5.01	-3.16	5.01	-3.16
	MP2/aug-cc-pVTZ ^c	3.70	-3.26	4.89	-3.46	4.89	-3.46
	MP2/aug-cc-pVQZ ^d	-----	-3.37	-----	-3.54	-----	-3.54
	MP2-R12/A ^d	-----	-3.64	-----	-3.63	-----	-3.63
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{e, f}}$	-----	1.26	-----	0.76	-----	0.76
	estd. CCSD(T)/aug-cc-pVTZ ^f	3.90	-1.80	4.99	-2.62	4.99	-2.62
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{d}}$	-----	1.83	-----	0.89	-----	0.89
	estd. CCSD(T)/CBS ^d	-----	-1.81	-----	-2.74	-----	-2.74
OH	MP2/aug-cc-pVDZ ^c	3.70	-3.40	5.00	-3.14	4.95	-3.23
	MP2/aug-cc-pVTZ ^c	3.60	-3.75	4.90	-3.42	4.90	-3.52
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{f}}$	-----	1.44	-----	0.77	-----	0.75
	estd. CCSD(T)/aug-cc-pVTZ ^f	3.80	-2.17	5.00	-2.58	5.00	-2.67
CH ₃	MP2/aug-cc-pVDZ ^c	3.70	-3.58	5.00	-3.11	4.90	-3.60
	MP2/aug-cc-pVTZ ^c	3.65	-3.96	4.90	-3.39	4.80	-3.89
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{f}}$	-----	1.55	-----	0.78	-----	0.81
	estd. CCSD(T)/aug-cc-pVTZ ^f	3.80	-2.27	5.00	-2.55	5.00	-2.95
F	MP2/aug-cc-pVDZ ^c	3.70	-3.50	4.95	-3.35	5.00	-2.87
	MP2/aug-cc-pVTZ ^c	3.70	-3.81	4.90	-3.61	4.90	-3.17
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{f}}$	-----	1.40	-----	0.74	-----	0.73
	estd. CCSD(T)/aug-cc-pVTZ ^f	3.80	-2.29	5.00	-2.77	5.00	-2.38
CN	MP2/aug-cc-pVDZ ^c	3.70	-4.49	4.90	-3.79	5.00	-2.82
	MP2/aug-cc-pVTZ ^c	3.60	-4.86	4.80	-4.11	4.90	-3.08
	$\Delta\text{CCSD(T)/aug-cc-pVDZ}^{\text{f}}$	-----	1.58	-----	0.84	-----	0.81
	estd. CCSD(T)/aug-cc-pVTZ ^f	3.80	-3.05	4.90	-3.25	5.00	-2.20

^a All computations reflect counterpoise correction. ^b Distance from center of benzene ring to center of aromatic ring containing the substituent. ^c Optimized geometry (monomer kept rigid) at each level of theory. ^d Using the best estimates of monomer geometry (C-C = 1.3915, C-H = 1.0800Å) from reference 55, and intermolecular distance optimized using counterpoise-corrected MP2/aug-cc-pVTZ. ^e aug-cc-pVDZ' represents a cc-pVDZ basis on hydrogen and an aug-cc-pVDZ basis minus diffuse d functions on other atoms. ^f Using monomer geometry optimized with MP2/aug-cc-pVTZ and intermolecular distance optimized using estimated counterpoise-corrected CCSD(T)/aug-cc-pVTZ.

Because the different substituents have a larger stabilizing effect on the sandwich configurations, the energy difference between the sandwich and T-shaped configurations becomes smaller for the substituted dimers than for benzene dimer. For the cyano substituent, the sandwich **2e** actually becomes $0.9 \text{ kcal mol}^{-1}$ more stable than the T-shaped dimer **4e**, demonstrating that the preferred orientation in a π - π interaction can be changed by only a modest degree of substitution.

Concerning convergence of the theoretical predictions, we observe that the optimized distance between monomers, R , is relatively insensitive to the improvement of the basis set at the MP2 level (so long as the counterpoise correction is employed), but using the larger basis set makes the binding significantly more favorable (~ 0.3 - $0.4 \text{ kcal mol}^{-1}$) for all dimer configurations. The estimated CCSD(T)/aug-cc-pVTZ optimized intermonomer distances are ~ 0.1 - 0.2 \AA larger than the MP2 predictions. This means that the estimated CCSD(T)/aug-cc-pVTZ binding energies will differ from those reported in our previous work⁸⁸ for the sandwich configurations, where we used MP2/aug-cc-pVTZ interfragment geometries; using the coupled-cluster geometries makes binding more favorable by about $0.2 \text{ kcal mol}^{-1}$, or around 5-10% of the total binding energy. The Δ CCSD(T) corrections are significant for all dimers (see Table 5.1), and they account for the overestimation of the binding energy by the MP2 method.^{32,46} This correction is largest for the sandwich configurations, ranging from $1.4 \text{ kcal mol}^{-1}$ for benzene-fluorobenzene dimer to $1.8 \text{ kcal mol}^{-1}$ for benzene dimer. For the T-shaped and T-shaped(2) dimer configurations, Δ CCSD(T) is 0.7 - $0.9 \text{ kcal mol}^{-1}$. The large magnitude of Δ CCSD(T) arises both from the coupling of electron pairs in CCSD (which is neglected in MP2) and from the importance of triple substitutions in CCSD(T). A recent study by

Hopkins and Tschumper⁸⁹ shows that both of these effects are very important in weakly bound dimers, and furthermore that the effect of connected quadruple substitutions is small but possibly non-negligible.

Given the sensitivity of the binding energies to the basis set and theoretical method, it might appear that one would require the very highest level of theory to accurately predict changes in binding energies due to substitution. Fortunately, however, Table 5.2 demonstrates that the binding energies relative to benzene dimer are accurately predicted at any of the levels of theory considered here, with variations of less than 0.1 kcal mol⁻¹ in most cases. This suggests that even though the absolute binding energies are very difficult to compute reliably, lower levels of theory should be sufficient to predict relative changes due to substitution in future studies of larger molecules.

Sandwich Dimers

As noted above, Table 5.1 indicates that all of the substituted sandwich dimers bind more strongly than benzene dimer. This is a surprising result if we note that these substituents are typically characterized as ranging from strongly electron donating (OH) to strongly electron withdrawing (CN). Our results appear to be inconsistent with the experimental study of Cozzi, Siegel, and co-workers,⁷⁹ which showed a linear relationship between the interaction energies of substituted phenyl groups and the sum of the Hammett parameters σ_{para} of the substituents. In that work, barriers to rotation ΔG^\ddagger about the aryl-naphthyl bond were determined using NMR for a few substituted 1,8-diarylnaphthalenes in which the two phenyl groups are forced into a nearly face-to-face stacked geometry. During rotation about the aryl-naphthyl bond, the stacked interaction is

lost, and the authors therefore assumed that the barriers to rotation are determined by the strength of the π - π interaction in the stacked geometry.

Table 5.2. Interaction Energies (in kcal mol⁻¹) Relative to Benzene Dimer^a

Sandwich Dimers 2a-e	X=H	OH	CH ₃	F	CN
MP2/aug-cc-pVDZ' ^{b,c}	0.00	-0.40	-0.54	-0.51	-1.40
MP2/aug-cc-pVDZ ^d	0.00	-0.50	-0.68	-0.60	-1.59
MP2/aug-cc-pVTZ ^c	0.00	-0.49	-0.70	-0.55	-1.60
Est. CCSD(T)/aug-cc-pVTZ ^c	0.00	-0.37	-0.47	-0.49	-1.25
SAPT2/aug-cc-pVDZ ^e	0.00	-0.49	-0.61	-0.61	-1.56
T-shaped Dimers 3a-e					
MP2/aug-cc-pVDZ' ^{b,c}	0.00	0.02	0.05	-0.17	-0.58
MP2/aug-cc-pVDZ ^d	0.00	0.02	0.05	-0.19	-0.63
MP2/aug-cc-pVTZ ^c	0.00	0.04	0.07	-0.15	-0.65
Est. CCSD(T)/aug-cc-pVTZ ^c	0.00	0.04	0.07	-0.15	-0.63
SAPT2/aug-cc-pVDZ ^e	0.00	0.03	0.07	-0.21	-0.71
T-shaped(2) Dimers 4a-e					
MP2/aug-cc-pVDZ' ^{b,c}	0.00	-0.04	-0.38	0.27	0.29
MP2/aug-cc-pVDZ ^d	0.00	-0.07	-0.44	0.29	0.34
MP2/aug-cc-pVTZ ^c	0.00	-0.06	-0.44	0.29	0.39
Est. CCSD(T)/aug-cc-pVTZ ^c	0.00	-0.05	-0.33	0.24	0.42
SAPT2/aug-cc-pVDZ ^e	0.00	-0.04	-0.44	0.30	0.33

^a All computations reflect counterpoise correction. ^b aug-cc-pVDZ' represents a cc-pVDZ basis on hydrogen and an aug-cc-pVDZ basis minus diffuse d functions on other atoms. ^c Intermonomer distance optimized at the given level of theory with rigid MP2/aug-cc-pVTZ monomer geometries. ^d Monomer geometry and intermonomer distance optimized at the MP2/aug-cc-pVDZ level of theory. ^e Using MP2/aug-cc-pVTZ optimized monomer geometry and intermolecular distances of 3.7Å for **2a-e**, and 4.9Å for **3a-e** and **4a-e**.

We note, however, that even in the transition state to rotation, there is still an interaction between the two phenyl groups (although their orientation is now different) which may also be affected by substituents; therefore, the differences in barrier heights for the different substituents may not be determined solely by the differences in the

attraction at the stacked geometry. The present theoretical results, which measure the intrinsic interaction energy directly, do not show a good correlation between interaction energy and Hammett σ parameters.

In earlier work on substituted sandwich dimers,⁸⁸ we attempted to analyze substituent effects in terms of the Hunter-Sanders rules, which state that although dispersion is critical to making π - π interactions favorable, it is electrostatics which determine changes in binding due to geometry or substituent effects. For a face-to-face sandwich dimer configuration, the most important consideration would be the negatively charged π -electron cloud above the ring center. Electron withdrawing substituents should reduce the negative π charge and lead to decreased π - π electrostatic repulsion, and vice versa for electron donating substituents. Such a picture seems consistent with the experimental results of Cozzi et al.,^{70,71} but it is not consistent with our theoretical binding energies, if one assumes that the Hammett parameters are indicative of the degree of π -electron density. The Hammett σ parameters, however, were determined from the equilibrium constants for the dissociation of substituted benzoic acids,⁹⁰ and there is no reason to assume that they necessarily correlate with the π -electron density in the reactants for those dissociations. Indeed, recent work on cation- π interactions by Dougherty and co-workers shows that the hydroxyl group, which is a strongly activating, electron donating substituent in the context of electrophilic aromatic substitution, has nearly the same electrostatic potential above the center of the ring as unsubstituted benzene.⁷⁷

Because the Hunter-Sanders rules propose that electrostatics are the most important consideration, we have computed electrostatic potential maps of the monomers

(Figure 5.2), rather than relying on Hammett σ parameters as an indirect measure of electrostatics. Figure 5.2 indicates that benzonitrile has the least negative π cloud, followed by fluorobenzene. However, the electrostatic potentials of the π clouds are similar for benzene, toluene, and phenol. Even though OH is electron donating in some other contexts, it has little effect on the electrostatic potential on top of the ring, in agreement with the findings of Dougherty and co-workers.⁷⁷

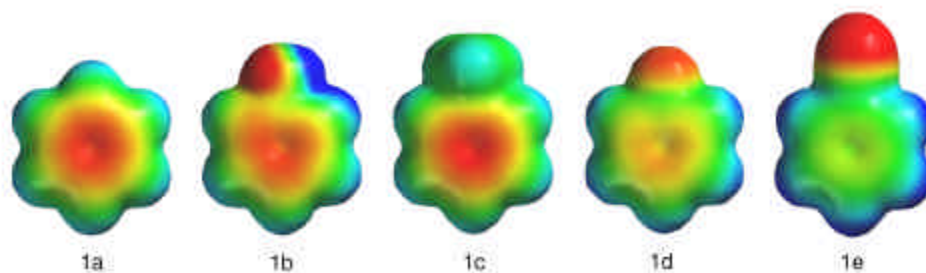


Figure 5.2. Electrostatic potentials computed using Hartree-Fock and a 6-31G* basis set with a scale of -25 to +25 kcal mol⁻¹. Potentials using B3LYP/6-31G* appear very similar.

Unfortunately, even if we ignore the Hammett σ parameters and consider the π charge as it appears from the electrostatic potentials, the Hunter-Sanders rules still do not give us qualitatively correct conclusions. Based on the electrostatic potentials computed at the HF/6-31G* level of theory (see Figure 5.2), we would expect benzene, toluene, and phenol to have nearly the same binding energies to benzene, which is not the case here (see Table 5.1 for the binding energies of different dimers). Instead, the difference in binding energies between toluene-benzene and benzene-benzene, which we would expect to be very small, is significantly larger than that between fluorobenzene-benzene and

toluene-benzene, which we would expect to be larger. Thus, even though Hunter-Sanders rules are useful in many instances for qualitative predictions of binding energies in π - π interactions, clearly they do not always predict the right trends for substituents because they lack other effects such as dispersion, induction, and exchange-repulsion. Concerning the three monomers (benzene, toluene, phenol) which ought to have similar electrostatic interactions with benzene based on the electrostatic potentials, their binding energies increase in the same order as their polarizabilities, suggesting that dispersion is important in determining the differences between substituted dimers. A more detailed analysis of the binding trends using SAPT analysis is described below.

T-shaped Dimers

Binding energies for the T-shaped dimers are also summarized in Table 5.1. Benzene-benzonitrile and benzene-fluorobenzene both bind more strongly than benzene dimer, by 0.63 and 0.15 kcal mol⁻¹, respectively, at the most reliable level of theory. Conversely, benzene-phenol and benzene-toluene are slightly less bound compared to benzene-dimer, by 0.04 and 0.07 kcal mol⁻¹, respectively. These changes are at least partially attributable to the electron donating or electron withdrawing effects of the substituent.

Electron withdrawing groups will decrease the exchange-repulsion term and increase the favorable electrostatic interaction between the partial-positive para hydrogen and the negatively charged π -electron cloud of the unsubstituted benzene ring below it; the opposite will happen for electron donating groups. Natural population analysis charges computed for the substituted monomers (B3LYP/cc-pVDZ) indicate relatively

small changes (= 0.004 a.u.) in the para hydrogen charge except in benzonitrile (0.008 a.u.), and the SAPT analysis below demonstrates that the largest changes generally come in the exchange-repulsion term, not the electrostatic term.

T-shaped(2) Dimers

As shown in Figure 5.1, in the T-shaped(2) dimers, a hydrogen from benzene points downward at the center of the ring of the substituted monomer. In this case, we expect changes in binding energies to correlate with the π -donating or withdrawing capacity of the substituents. A π -donating substituent should increase the negative charge of the π cloud on the substituted benzene, leading to a more favorable electrostatic interaction with the partial positive charge on the hydrogen pointing down at it. The electrostatic potential maps plotted in Figure 5.2 suggest that binding should be similar for benzene, phenol, and toluene, smaller for fluorobenzene, and smallest for benzonitrile. Indeed, the decreases in binding energies compared to benzene dimer for benzene-fluorobenzene and benzene-benzonitrile are 0.24 kcal mol⁻¹ and 0.42 kcal mol⁻¹, respectively, compared to benzene dimer. The binding energy of benzene-phenol is very similar to that of benzene dimer (0.05 kcal mol⁻¹ more stable), as expected. However, the binding energy of toluene is significantly increased, by 0.33 kcal mol⁻¹. Our SAPT analysis (below) indicates that this is due to greater dispersion energy for benzene-toluene than for benzene-phenol or benzene dimer.

According to the preceding analysis of supermolecular binding energies, it is clear that consideration of electrostatic effects alone (as advanced by the Hunter-Sanders rules) is not sufficient to fully explain the trends in the binding energies of the substituted

dimers. To better understand the observed trends, we now turn to SAPT to decompose the binding energy into its electrostatic, dispersion, induction, and exchange-repulsion components.

5.3.2 SAPT Approach

All SAPT computations reported here were performed using MP2/aug-cc-pVTZ optimized monomer geometries. The individual energy components of the SAPT analysis were found to be very sensitive to the interfragment distance; for example, when T-shaped benzene-benzonitrile **3e** is computed at distances of 4.8 and 4.9 Å, the exchange-repulsion term changes by 35%, and the $E_{disp}^{(20)}$ dispersion term changes by 16%. Such changes were often larger than the variations due to substituent effects. For that reason, we performed all SAPT computations at the same intermonomer distances: 3.7 Å for the sandwiches **2**, and 4.9 Å for the T-shaped configurations **3** and **4**. All SAPT computations were carried out using the modified aug-cc-pVDZ' basis. Critical to our SAPT analysis is the ability of such a modest basis set to faithfully reproduce the higher-level supermolecular (SM) ordering of the binding energies for the dimers studied. Table 5.2 indicates that the shifts in binding energies due to substituents are reliably predicted by SAPT2/aug-cc-pVDZ' (despite our using the same interfragment distance for all dimers of a given configuration, which is not the equilibrium distance in all cases).

As we did for the supermolecular part of the discussion, we will first discuss the sandwich dimers results, and then present the results for the two T-shaped dimer configurations.

Sandwich Dimers

The SAPT binding energies for the sandwich dimers are summarized in Table 5.3. The electrostatic component of the binding energy, represented here by the sum of $E_{elst}^{(10)}$ and $E_{elst,r}^{(12)}$, is always stabilizing. This is surprising from the point of view of the Hunter-Sanders model, which would imagine two negatively charged π clouds directly on top of each other for the benzene dimer sandwich. However, such a picture ignores the fact that the electron clouds interpenetrate, and the electrostatic penetration term is usually attractive. The electrostatic energies for the T-shaped dimers **3** and **4** (Tables 5.4 and 5.5) are much more attractive than for the sandwiches **2**, in agreement with the expected dominance of attractive σ - π interactions in the T-shaped configuration. The exchange-repulsion terms are substantially larger than the attractive electrostatic terms in the sandwiches, so that the sum of the electrostatic and exchange terms is overall repulsive.

Because CN is the most strongly electron withdrawing substituent, it should reduce the π density the most, decreasing unfavorable π - π repulsion. Indeed, we observe the most favorable electrostatic energy for benzene-benzonitrile (0.86 kcal mol⁻¹ more stable than benzene dimer). The next most favorable electrostatic energy is for benzene-fluorobenzene, which is consistent with F being the next most effective electron withdrawing substituent, as indicated by the electrostatic potentials in Figure 5.2. According to that figure, the electrostatic energies should be nearly the same for the sandwiches of benzene with benzene, phenol, and toluene, and this is what we observe in Table 3. Hence, the trends in electrostatic energies with respect to substitution seem consistent with expectations based on electrostatic potential maps of the monomers.

Table 5.3. Contributions to the Interaction Energy (kcal mol⁻¹) for Different Sandwich Dimer Configurations **2a-e**^a

	X=H	OH	CH ₃	F	CN
$E_{\text{int}}^{\text{HF}}$	5.330	4.947	5.352	4.534	3.914
$E_{\text{elst}}^{(10)}$	-0.520	-0.689	-0.544	-1.073	-1.757
$E_{\text{exch}}^{(10)}$	6.185	5.984	6.299	5.900	5.968
$E_{\text{ind},r}^{(20)}$	-2.196	-2.252	-2.291	-2.153	-2.150
$E_{\text{exch-ind},r}^{(20)}$	2.002	2.058	2.066	2.010	2.005
$dE_{\text{int},r}^{\text{HF}}$	-0.141	-0.153	-0.179	-0.150	-0.151
$E_{\text{elst},r}^{(12)}$	-0.454	-0.388	-0.483	-0.282	-0.075
$E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)}$	-0.151	-0.132	-0.089	-0.169	-0.190
${}^t E_{\text{ind}}^{(22)}$	0.054	0.035	0.043	0.038	0.049
${}^t E_{\text{exch-ind}}^{(22)}$	-0.050	-0.032	-0.039	-0.036	-0.046
$E_{\text{disp}}^{(20)}$	-7.470	-7.653	-8.173	-7.377	-7.909
$E_{\text{exch-disp}}^{(20)}$	0.942	0.933	0.985	0.888	0.899
$E_{\text{int}}(\text{SAPT2})$	-1.798	-2.289	-2.405	-2.403	-3.357
$E_{\text{int}}(\text{MP2})$ ^b	-1.744	-2.223	-2.345	-2.326	-3.273

^a Using MP2/aug-cc-pVTZ optimized monomer geometry and intermolecular distance of 3.7 angstroms.

^bMP2/aug-cc-pVDZ' counterpoise-corrected binding energies.

Because the electrostatic energies for benzene-phenol and benzene-toluene are so similar to that of benzene dimer, clearly electrostatics alone cannot explain why the total binding energies of those dimers are significantly larger than that of benzene dimer. Other energetic components are therefore important in determining the energetic order of the dimers.

The exchange energy terms calculated here are $E_{\text{exch}}^{(10)}$, $E_{\text{exch-ind},r}^{(20)}$, $E_{\text{exch}}^{(11)}$, $E_{\text{exch}}^{(12)}$, ${}^t E_{\text{exch-ind}}^{(22)}$, and $E_{\text{exch-disp}}^{(20)}$. The first term, $E_{\text{exch}}^{(10)}$, accounts for the repulsion due to the Pauli exclusion principle and arises from the antisymmetry requirement of the wavefunction, $E_{\text{exch}}^{(11)}$ and $E_{\text{exch}}^{(12)}$ account for the effects of intramonomer correlation on the exchange

repulsion, and $E_{exch-ind,r}^{(20)}$ and other second order terms (${}^tE_{exch-ind}^{(22)}$ and $E_{exch-disp}^{(20)}$) account for additional exchange repulsion arising from the coupling of electron exchange and the induction and dispersion interactions. The exchange energy is slightly more repulsive for benzene-toluene and slightly less repulsive for benzene-phenol, benzene-fluorobenzene, and benzene-benzonitrile, than in benzene dimer. This is consistent with the reduced π - π overlap for the electron withdrawing CN and F substituents and the increased π - π overlap for the weakly electron donating CH_3 group. Kim and co-workers previously noted a similar reduction in exchange-repulsion energies for fluorobenzene-argon as compared to benzene-argon.⁹¹

The relative exchange energies are larger in magnitude than the relative electrostatic energies for CH_3 and OH substituents, indicating the importance of exchange terms. When the exchange-induction and exchange-dispersion cross terms are counted as induction and dispersion, respectively, the relative exchange energy for the OH substituent (compared to benzene dimer) is almost twice as large as the relative electrostatic energy.

The induction contribution to the binding energy is mainly contained in $E_{ind,r}^{(20)}$. This is a second-order energy correction that results from the distortion of the charge distribution of one monomer by the electrostatic charge distribution of the other monomer, and vice versa. This mutual polarization of the monomer by the static electric field of the other is proportional to the multipole moments and static polarizabilities of the monomers. The leading intramonomer correlation contribution is included in ${}^tE_{ind}^{(22)}$ and accounts for only 2% of the induction energy. Table 5.3 shows that the dominant

induction term $E_{ind,r}^{(20)}$ is very similar for all sandwich dimers. The attractive part of the induction energy is substantially quenched by the repulsive exchange-induction energy (represented by $E_{exch-ind,r}^{(20)}$ and ${}^tE_{exch-ind}^{(22)}$). As noted by Jeziorski and coworkers,¹⁹ any quantitatively accurate calculation of the induction energy cannot neglect the exchange-induction contribution. If we account for this repulsive term and also add the third- and higher-order induction and exchange-induction terms in dE_{int}^{HF} , induction stabilizes the total binding energy by 0.3-0.4 kcal mol⁻¹ for the sandwich dimers investigated. The shifts in the induction energies relative to benzene dimer due to substitution are less than 0.07 kcal mol⁻¹.

Dispersion stabilizes the binding energy of the sandwich dimers by 6.5 to 7.2 kcal mol⁻¹ after the exchange-dispersion correction. This is by far the largest attractive contribution. Figure 5.3 displays the good correlation between dispersion energies and the π components of the polarizabilities of the substituted monomers, computed at the HF/aug-cc-pVDZ level (46.2 **1a**, 47.4 **1b**, 54.9 **1c**, 44.7 **1d**, and 51.8 e²a₀²/E_h **1e**). The dispersion energies relative to benzene dimer **2a** are significant for all but the F substituent, ranging from -0.66 to 0.04 kcal mol⁻¹. For methyl-substituted **2c**, dispersion is the most important contributor to the energy lowering relative to benzene dimer **2a**. For hydroxyl-substituted **2b**, dispersion is twice as important as electrostatics in contributing to the stabilization relative to **2a**.

To summarize the above results, we plot in Figure 5.4 the relative changes in electrostatic, exchange-repulsion, induction, and dispersion components of the interaction energy for different sandwich dimers compared to benzene dimer. The dE_{int}^{HF} term was left out and *not* added to any of the four energy components.

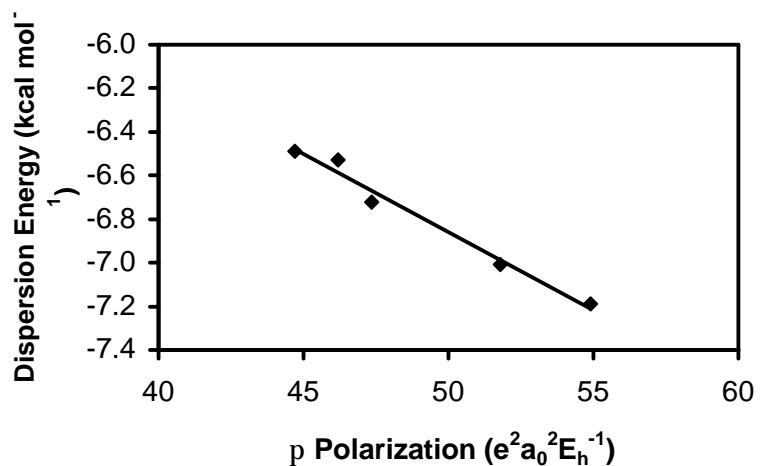


Figure 5.3. Plot of dispersion energy for sandwich dimers (SAPT2/aug-cc-pVDZ', exchange-corrected) vs. polarizability in the direction perpendicular to the aromatic plane (HF/aug-cc-pVDZ).

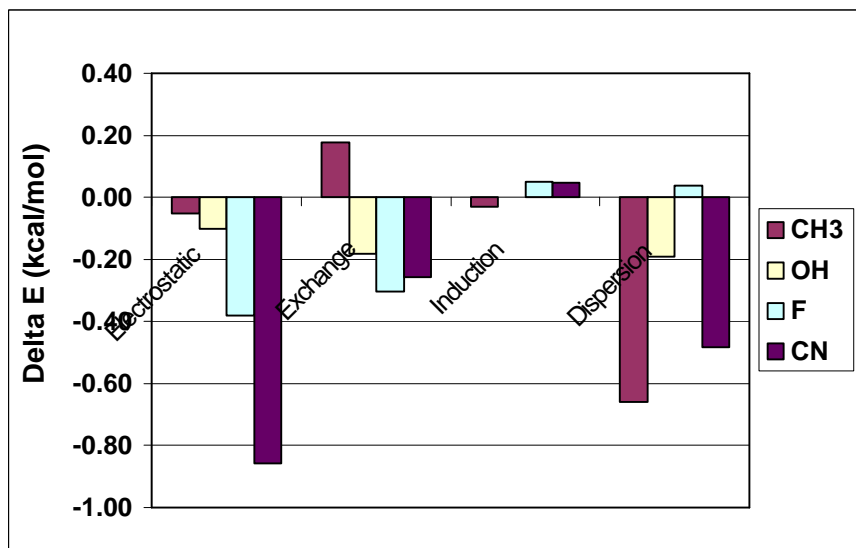


Figure 5.4. The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different sandwich dimers compared to benzene dimer.

T-shaped Dimers

Table 5.4 shows the SAPT contributions to the interaction energy for T-shaped dimers **3a-e**. It should be noted that, in general, the T-shaped dimers have a larger electrostatic component than their sandwich dimers counterparts because of favorable quadrupole-quadrupole interactions. They also exhibit smaller destabilizing exchange-repulsion energies and smaller stabilizing dispersion energies than the sandwich dimers. Nevertheless, the dispersion and exchange-repulsion energies remain larger than the electrostatic energies for **3a-e**.

Examining Table 5.4, we see that the sum of electrostatic terms ($E_{elst}^{(10)} + E_{elst,r}^{(12)}$) is very similar for all of the T-shaped dimers, with differences of 0.04 kcal mol⁻¹ or less from benzene dimer except for benzene-benzonitrile. The trend in electrostatic energies follows the trend in computed para hydrogen charges, with the exception of the hydroxyl substituent. The trend in electrostatic energies also happens to match the trend in total binding energies, except for a reversal of the order for the CH₃ and OH substituents. However, due to the very small changes in electrostatic energies due to substitution, and the much larger changes in other energy components (below), this appears fortuitous.

The exchange interactions for the T-shaped dimers are considerably less repulsive than for the sandwiches; this is due to the reduced overlap between the orbitals of the two monomers. Benzene-fluorobenzene and benzene-benzonitrile have smaller (0.32 and 0.27 kcal mol⁻¹, respectively) exchange-repulsion energies than benzene dimer because the F and CN substituents are electron withdrawing and reduce the electron density available to

interact with the other benzene. Such a reduction of electron density in the ring also reduces the favorable dispersion contribution by about 0.1 kcal mol⁻¹. An opposite effect

Table 5.4. Contributions to the Interaction Energy (kcal mol⁻¹) for Different T-shaped Dimer Configurations **3a-e**^a

	X=H	OH	CH ₃	F	CN
$E_{\text{int}}^{\text{HF}}$	1.618	1.677	1.817	1.228	0.648
$E_{\text{elst}}^{(10)}$	-2.190	-2.135	-2.131	-2.293	-2.770
$E_{\text{exch}}^{(10)}$	4.447	4.442	4.588	4.171	4.219
$E_{\text{ind},r}^{(20)}$	-1.161	-1.152	-1.189	-1.105	-1.210
$E_{\text{exch-ind},r}^{(20)}$	0.914	0.907	0.946	0.836	0.847
$dE_{\text{int},r}^{\text{HF}}$	-0.392	-0.385	-0.397	-0.382	-0.438
$E_{\text{elst},r}^{(12)}$	-0.054	-0.071	-0.104	0.022	0.181
$E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)}$	0.418	0.405	0.431	0.376	0.374
${}^t E_{\text{ind}}^{(22)}$	-0.144	-0.139	-0.154	-0.117	-0.105
${}^t E_{\text{exch-ind}}^{(22)}$	0.113	0.109	0.123	0.089	0.073
$E_{\text{disp}}^{(20)}$	-4.893	-4.896	-5.004	-4.713	-4.772
$E_{\text{exch-disp}}^{(20)}$	0.526	0.529	0.547	0.489	0.478
$E_{\text{int}}(\text{SAPT2})$	-2.415	-2.385	-2.344	-2.626	-3.122
$E_{\text{int}}(\text{MP2})$ ^b	-2.248	-2.241	-2.189	-2.464	-2.888

^a Using MP2/aug-cc-pVTZ optimized monomer geometry and intermolecular distance of 4.9 angstroms.

^bMP2/aug-cc-pVDZ' counterpoise-corrected binding energies.

is seen for benzene-toluene because methyl is a slightly electron donating substituent. The increased electron density in the toluene ring leads to a slightly larger exchange repulsion (~0.2 kcal mol⁻¹) and more favorable dispersion energy (0.1 kcal mol⁻¹) than in benzene dimer. Although substituent effects for exchange and dispersion are of opposite sign, the shifts in the exchange energies are usually 2-4 times larger. Therefore, dispersion is less important in determining relative energies for the T-shaped dimers than for the sandwich dimers. Substitution by OH leads to very small changes in exchange and

dispersion energies relative to benzene dimer. Shifts in exchange-corrected induction energies relative to benzene dimer are relatively small and generally have the same sign as the electrostatic shifts. The largest effect due to induction is a $0.1 \text{ kcal mol}^{-1}$ increase in binding for the CN substituent.

As for the sandwich dimer, we summarize the above results for the T-shaped dimers by plotting in Figure 5.5 the relative changes in electrostatic, exchange-repulsion, induction, and dispersion components of the interaction energy for different T-shaped dimers compared to benzene dimer.

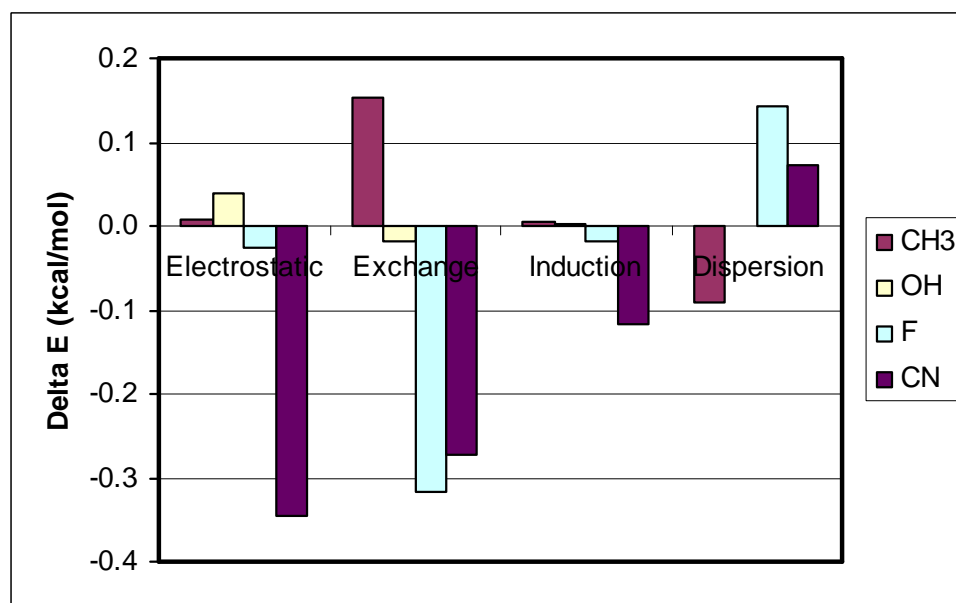


Figure 5.5. The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different T-shaped dimers compared to benzene dimer.

T-shaped(2) Dimers

Table 5.5 shows the SAPT contributions to the interaction energy for substituted T-shaped(2) dimers **4a-e**. In this case the substituent will either enhance or reduce the π -electron density on the lower benzene ring, and consequently, increase or decrease its electrostatic interaction with the partial positive charge of the hydrogen on the upper benzene. The electron withdrawing F and CN substituents reduce the amount of π -electron density on the lower ring and decrease both the electrostatic interaction (by 0.27 and 0.52 kcal mol⁻¹, respectively) and the total binding energy (by 0.30 and 0.33 kcal mol⁻¹ at the SAPT2 level). The electron donating methyl substituent has the opposite effect, increasing the electrostatic attraction by 0.14 kcal mol⁻¹ and the total binding energy by 0.44 kcal mol⁻¹. Once again the OH substituent has little effect, with a decrease in electrostatic attraction of 0.06 kcal mol⁻¹ and an increase in total binding of the same size. Other than this sign change for OH, the trend in total binding energies follows that of the electrostatic interaction. We observe once again that the exchange-corrected induction energies are relatively small, and their shifts relative to benzene dimer have the same sign but are several times smaller than the shifts in the electrostatic energies.

Although the relative energies generally follow the same trend as the electrostatic energies, this does not mean that the electrostatic contribution to the total energy is the dominant one. As was seen for **3**, both the exchange and dispersion energies are greater in magnitude than the electrostatic energies for **4**. This agrees with Wilcox's conclusion,⁸² based on the molecular torsion balance, that "the electrostatic potential of the aromatic ring is *not* a dominant aspect of the aryl-aryl interaction."

We might expect that the F and CN substituents, by reducing the amount of π -electron density directly below the hydrogen from the other benzene, should also decrease the exchange repulsion relative to benzene dimer. This expectation is fulfilled by reductions of 0.13 and 0.18 kcal mol⁻¹ in the exchange energies for fluorine and cyano

Table 5.5. Contributions to the Interaction Energy (kcal mol⁻¹) for Different T-shaped(2) Dimer Configurations **4a-e**^a

	X=H	OH	CH ₃	F	CN
$E_{\text{int}}^{\text{HF}}$	1.618	1.674	1.365	1.971	2.295
$E_{\text{elst}}^{(10)}$	-2.190	-2.065	-2.343	-1.816	-1.510
$E_{\text{exch}}^{(10)}$	4.447	4.365	4.374	4.333	4.293
$E_{\text{ind},r}^{(20)}$	-1.161	-1.167	-1.179	-1.118	-1.110
$E_{\text{exch-ind},r}^{(20)}$	0.914	0.930	0.914	0.929	0.939
$dE_{\text{int},r}^{\text{HF}}$	-0.392	-0.389	-0.401	-0.355	-0.316
$E_{\text{elst},r}^{(12)}$	-0.054	-0.119	-0.036	-0.161	-0.218
$E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)}$	0.418	0.417	0.435	0.400	0.393
${}^t E_{\text{ind}}^{(22)}$	-0.144	-0.158	-0.146	-0.158	-0.154
${}^t E_{\text{exch-ind}}^{(22)}$	0.113	0.126	0.113	0.131	0.130
$E_{\text{disp}}^{(20)}$	-4.893	-4.932	-5.120	-4.802	-5.032
$E_{\text{exch-disp}}^{(20)}$	0.526	0.519	0.534	0.500	0.498
$E_{\text{int}}(\text{SAPT2})$	-2.415	-2.474	-2.854	-2.117	-2.087
$E_{\text{int}}(\text{MP2})^{\text{b}}$	-2.248	-2.316	-2.687	-1.967	-1.963

^a Using MP2/aug-cc-pVTZ optimized monomer geometry and intermolecular distance of 4.9 angstroms.

^bMP2/aug-cc-pVDZ' counterpoise-corrected binding energies.

substituents, respectively. On the other hand, the exchange energy for the CH₃ substituent is also decreased (by a small 0.06 kcal mol⁻¹), in contrast to its modest electron donating character.

As for the sandwich configuration, we observe a correlation between dispersion energies and the computed π polarizabilities of the substituted monomers (see Figure

5.6). The relative dispersion energy is comparable in magnitude to the relative electrostatic energy for hydroxyl-substituted **4b**, and it is larger than the relative electrostatic energy for methyl-substituted **4c**. Hence, the electrostatic energy largely determines the energetic order of the T-shaped dimers **4**, but other effects remain important in determining the size of the relative energies. Finally, Figure 5.7 display the relative changes in electrostatic, exchange-repulsion, induction, and dispersion components of the interaction energy for different T-shaped(2) dimers compared to benzene dimer.

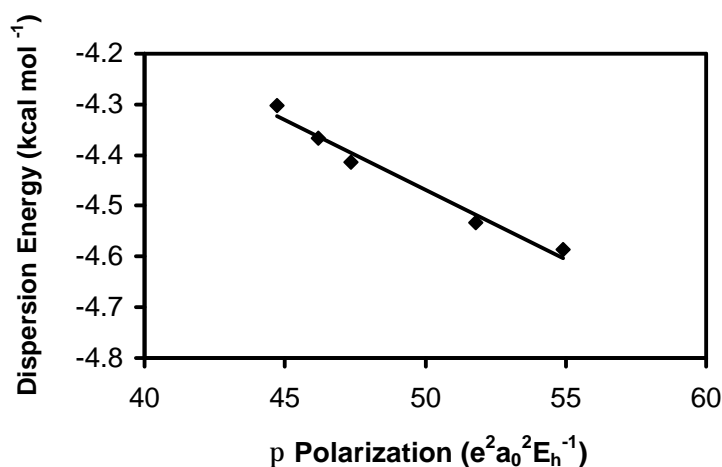


Figure 5.6. Plot of dispersion energy for T-shaped(2) dimers (SAPT2/aug-cc-pVDZ', exchange-corrected) vs. polarizability in the direction perpendicular to the aromatic plane (HF/aug-cc-pVDZ).

5.4 Conclusions

A better understanding of π - π interactions will aid rational design efforts in biological chemistry and crystal engineering. Substituent effects in the sandwich and T-

shaped configurations of benzene dimer have been quantified using an additive scheme to obtain high-quality CCSD(T)/aug-cc-pVTZ binding energies and interfragment distances.

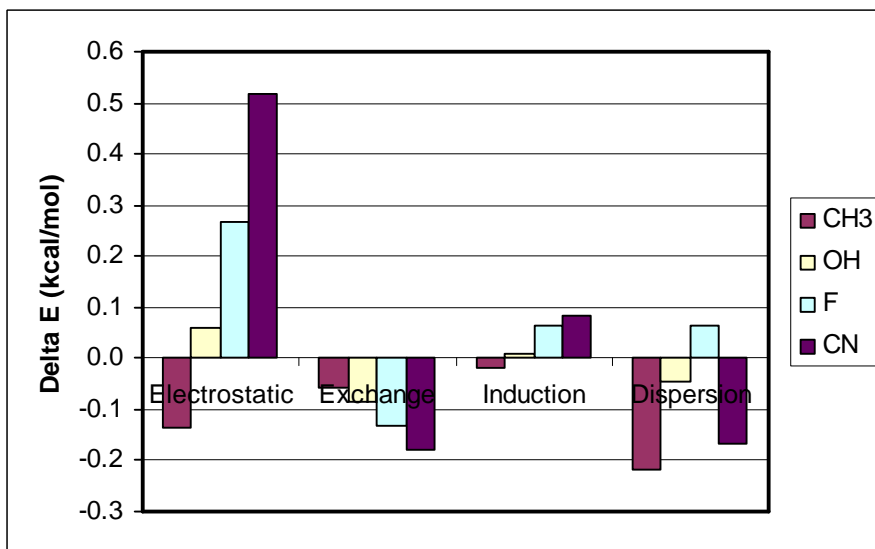


Figure 5.7. The relative changes in electrostatic, exchange, induction, and dispersion components of the interaction energy for different T-shaped(2) dimers compared to benzene dimer.

Of all the substituents studied, the cyano substituent has by far the largest effect, changing the binding energy by more than 1 kcal mol⁻¹ relative to benzene dimer in some cases; this substituent is even capable of making the sandwich configuration drop lower in energy than one of the T-shaped configurations. In general, fluorine has the next largest effect, followed by methyl and then hydroxyl. Unlike previous experimental studies, the present work does not show a good correlation between binding energies and Hammett parameters (although a rough correlation with σ_{meta} is seen for the T-shaped dimers).

To uncover the origin of the observed trends in binding energies, we performed the first symmetry-adapted perturbation theory (SAPT2) decomposition of π - π

interaction energies into their electrostatic, dispersion, induction, and exchange-repulsion components. Dispersion and exchange-repulsion are more important than electrostatics in determining the total binding energies of the dimers considered. Induction energies are largely quenched by exchange-induction coupling, and they contribute very little to differences in binding energies between the substituted dimers. Contrary to the predictions of the Hunter-Sanders rules, electrostatics alone is insufficient to predict the correct trends in binding. For the sandwich configuration, electrostatics suggest that phenol and toluene should bind to benzene about as well as benzene does; however, both of them actually bind more strongly by 0.4-0.5 kcal mol⁻¹. For several T-shaped dimers, either exchange or dispersion makes larger contributions to the relative energy than electrostatics. This suggests that models based solely on electrostatic effects will have difficulty in reliably predicting substituent effects in π - π interactions.

CHAPTER 6

HIGHLY ACCURATE COUPLED-CLUSTER POTENTIAL ENERGY CURVES FOR BENZENE DIMER

In this work, state-of-the-art electronic structure theory has been applied to generate potential energy curves (PEC's) for the sandwich, T-shaped, and parallel-displaced configurations of the simplest prototype of aromatic π - π interactions, the benzene dimer. Potential energy curves as a function of intermonomer geometries for the dimers were obtained at the second-order Møller-Plesset perturbation theory (MP2), and coupled-cluster with singles, doubles and perturbative triples [CCSD(T)] levels with different augmented, correlation-consistent basis sets. At the MP2 level, the smallest basis set used (a modified aug-cc-pVDZ basis) underestimates the binding by ~ 0.5 kcal mol⁻¹ at equilibrium and by ~ 1 kcal mol⁻¹ at smaller intermonomer distances compared to the MP2 values with a modified aug-cc-pVQZ basis (denoted aug-cc-pVQZ*). By comparison with the more accurate CCSD(T) PEC's, the best MP2 binding energies differ from the CCSD(T) ones by about 1.7 to 2.0 kcal mol⁻¹ at equilibrium and by more than 2.5 kcal mol⁻¹ at smaller intermonomer distances, highlighting the importance of relying on the coupled-cluster PEC's to achieve higher accuracy in estimating binding energies. When comparing the binding energies at equilibrium using the current approach to the single point complete basis set (CBS) CCSD(T) estimates of the binding energies from our previous work, the current CCSD(T)/aug-cc-pVQZ* binding energies are within 0.1-0.15 kcal mol⁻¹ of the more rigorous CBS estimates. These high-quality potential

energy curves for benzene dimer provide a better understanding of how the strength of π - π interactions varies with distance and orientation of the rings, and will assist in the development of approximate methods capable of modeling weakly bound π - π systems.

6.1 Introduction

Benzene dimer, the simplest prototype of aromatic π - π interactions, has been extensively studied, both theoretically^{46,47,49,50,67,88} and experimentally,³⁷⁻⁴¹ in an attempt to obtain a clear picture of the strength and directionality of π - π interactions. The small binding energy of the gas phase benzene dimer (~ 2 - 3 kcal mol⁻¹) and the shallowness of the potential energy surface make it a challenge for both theory and experiment. Additionally, although microwave spectroscopy has provided a structure for the T-shaped configuration,³⁸ other configurations such as the sandwich and parallel-displaced configurations are microwave undetectable because they lack permanent dipole moments.

In various complex chemical and biological systems, the aromatic rings comprising these systems can be found at different orientations and distances from each other. Due to steric constraints imposed on such systems, these geometries might not correspond to the potential energy minima for π - π interactions. Nevertheless, the aromatic rings might still interact favorably enough to contribute significantly to the overall stability of the system. In their notable X-ray crystallographic study of side-chain aromatic interactions in 34 high-resolution protein structures, Burely and Petsko⁴ analyzed the frequency of aromatic pairs and their interaction geometry (distance and dihedral angle); they found that around 60% of aromatic side chains of phenylalanine, tyrosine, and tryptophan, were involved in aromatic pairs. Aromatic rings separated by

distances ranging from 4.5-7 Å, and dihedral angles near 90° were found to be the most common. Pairwise nonbonded potential energy estimations indicated that fifty four percent of the aromatic interactions have a stabilizing energy between 1 and 2 kcal mol⁻¹. In a study of a larger database of 52 proteins, Hunter, Singh, and Thornton⁵ examined the orientational preferences of phenylalanine side chains in proteins using crystallographically derived atomic coordinates. They observed that these interacting pairs are found in a wide range of T-shaped (edge-to-face) and parallel-displaced (offset-stacked) arrangements, and they are scattered over a wide variety of conformational space with no strongly preferred single orientation.

From the above discussion, it becomes clear that high-quality potential energy curves for prototype systems would be very helpful in better understanding how π - π interactions depend on both the orientation and distance between these aromatic rings. Due to computational limitations, previous attempts to obtain ab initio potential energy surfaces for benzene dimer^{32,33,50,67,92} were mostly performed at the MP2 level and involved the use of relatively small basis sets, hence they lack the high accuracy needed to model π - π interactions reliably. Also, these curves estimated the binding energies only at a small set of intermonomer distances, and are therefore, at best, incomplete. Jaffe and Smith³² used the MP2 method along with the 6-311G(2d,2p) basis set to evaluate the potential energy curves of the sandwich, T-shaped, parallel-displaced, and V-shaped configurations of benzene dimer, and to determine the interconversion path of the parallel-displaced configuration. Even though the MP2 method is qualitatively correct, it tends to overestimate the correlation energy. Also, the basis set used in these calculations is of medium size (384 basis functions), and does not contain any diffuse functions.

Špirko, Hobza, and co-workers⁹³ evaluated the CCSD(T) binding energies for different configurations of benzene dimer to parameterize the nonempirical model (NEMO) intermolecular potential, and then compared their theoretical structures and barriers to rotation with microwave and Raman spectral data. They noted sizable differences between theoretical and experimental predictions, especially the height of the barrier opposing the hindered internal rotation in the T-shaped geometry. In their study, they computed 107 CCSD(T) single point energies for the fitting of the NEMO potential using a modified cc-pVDZ basis with 228 basis functions. Even though these calculations are impressive, the modified cc-pVDZ basis set used in the CCSD(T) computations is too small to obtain highly accurate binding energies for π - π systems.

Tsuzuki and co-workers^{50,92} recently studied the energy profile of the interconversion between the T-shaped and parallel-displaced configurations of benzene dimer. Binding energies for a number of horizontal and vertical displacements were computed at the estimated CCSD(T) basis set limit for four tilt angles ($\mathbf{f} = 0^\circ, 30^\circ, 60^\circ, \text{ and } 90^\circ$) along the interconversion path. Additionally, an aug(d,p)-6-311G** basis (384 basis functions) was used to estimate the MP2 interaction energies. Our previous study on benzene dimer showed that the aug-cc-pVDZ basis set, which has the same number of basis functions as the aug(d,p)-6-311G** basis, is a significant ~ 0.5 - $0.7 \text{ kcal mol}^{-1}$ away from the MP2 complete basis set (CBS) limit. Furthermore, the basis set used to estimate the $\Delta\text{CCSD(T)}$ correction is of medium size (6-311G*) and lacks the diffuse functions needed to obtain more accurate values for this correction. These factors could affect the accuracy of the energy profile and the energetic ordering of the different

configurations examined, especially if one considers that the interconversion barrier height is only 0.2 kcal mol⁻¹.

In our previous study of benzene dimer,⁶⁷ we computed MP2 potential energy curves as a function of intermonomer distance for the sandwich (S), T-shaped (T), and parallel-displaced (PD) configurations using the aug-cc-pVTZ basis set (828 basis functions), a much larger basis than previously used. At the equilibrium geometries, we also combined CBS MP2 binding energies (obtained using MP2-R12/A theory) with a Δ CCSD(T) correction computed in a smaller basis (aug-cc-pVDZ) to estimate the CBS CCSD(T) results, which should provide binding energies accurate to a few tenths of one kcal mol⁻¹. However, no coupled-cluster potential energy curves were reported in that study because it was not feasible to obtain entire potential energy curves at the very high CBS CCSD(T) level of theory.

In this work, we aim to obtain potential curves that are very close to the ab initio limit while remaining computationally feasible. High-quality potential energy curves (PEC's) are obtained as a function of the intermonomer distance R for the S and T configurations, and of R_1 and R_2 for the PD configuration (see Figure 6.1). At the MP2 level, PEC's are computed using correlation-consistent basis sets as large as aug-cc-pVQZ (minus g functions on carbon and f 's on hydrogen). The CCSD(T) potential energy curves are computed using a truncated aug-cc-pVDZ basis to produce reliable results for the Δ CCSD(T) correction, and combined with the best MP2 values to estimate high-quality CCSD(T) potential energy curves.

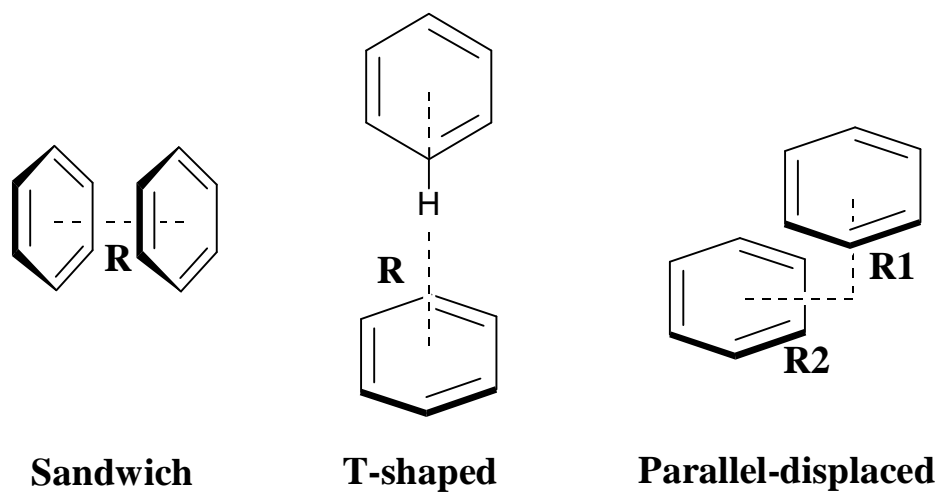


Figure 6.1. Sandwich, T-shaped, and parallel-displaced configurations of the benzene dimer.

The availability of high quality curves will be critical to the calibration of molecular mechanics force field methods meant to accurately model biochemical systems exhibiting π -stacking interactions. The Lennard-Jones component of these force fields is usually fitted to *ab initio* gas phase binding energies and equilibrium geometries.⁹⁴⁻⁹⁹ In many force field calibrations, the MP2 method with polarized double-zeta or larger basis sets are used for geometry optimizations and gas phase dimerization energy estimates.⁹⁵⁻⁹⁹ Our present results use a considerably more reliable methodology. It is worth noting that our high-quality benzene dimer binding energies have already been used by Friesner and co-workers in calibrating their new polarizable force field (PFF).⁹⁶

6.2 Theoretical Methods

All computations in this work were performed using Dunning's correlation-consistent split valence basis sets of contracted Gaussian functions.⁵³ Multiple polarization and diffuse functions were added to better describe the polarizability of the monomers and the delocalized nature of electrons in the benzene ring. Specifically, we used an aug-cc-pVDZ basis set without diffuse s and p functions on hydrogen (denoted aug-cc-pVDZ*) containing 336 basis functions, an aug-cc-pVTZ basis set (828 basis functions), and an aug-cc-pVQZ basis set minus all g functions on carbon and all f functions on hydrogen (denoted aug-cc-pVQZ*) with 1128 basis functions.

Potential energy curves (PEC's) as a function of intermonomer distances for the sandwich, T-shaped, and parallel-displaced configurations were computed via second order perturbation theory (MP2) with the aug-cc-pVDZ*, the aug-cc-pVTZ, and the aug-cc-pVQZ* basis sets. For the sandwich and T-shaped configurations, the center-to-center distance, R , was systematically varied, and for the parallel-displaced configuration both the vertical, R_1 , and horizontal, R_2 , distances between the centers of mass were systematically varied (see Figure 6.1). All calculations were performed with the recommended benzene monomer geometry of Gauss and Stanton⁵⁵ (C-C = 1.3915 Å and C-H = 1.0800 Å), and this monomer geometry was not allowed to vary in the calculations. Our previous study has shown that the effect of freezing the monomer geometry is minimal on both the equilibrium geometry and binding energy of benzene dimer⁶⁷. The basis set superposition error (BSSE), which results from the use of an incomplete basis set, was corrected for by the counterpoise (CP) method of Boys and

Bernardi.⁵⁴ Only CP-corrected binding energies are reported in this work. Core orbitals were constrained to remain doubly occupied in all correlated computations.

To account for the effect of triple excitations on the binding energies of benzene dimer, CCSD(T) potential energy curves were computed using the above-mentioned aug-cc-pVDZ* basis. Due to the prohibitive computational cost, it was not possible to obtain CCSD(T) PEC's using the full aug-cc-pVDZ basis set. The Δ CCSD(T) correction is computed in an aug-cc-pVDZ* basis as

$$\Delta CCSD(T) = E_{CCSD(T)}^{aug-cc-pVDZ*} - E_{MP2}^{aug-cc-pVDZ*} \quad [6.1]$$

This correction is combined with the MP2/aug-cc-pVQZ* curves to estimate high-quality CCSD(T)/aug-cc-pVQZ* potential energy curves for benzene dimer according to the equation

$$E_{CCSD(T)}^{aug-cc-pVQZ*} = E_{MP2}^{aug-cc-pVQZ*} + \Delta CCSD(T) \quad [6.2]$$

If we compare our current Δ CCSD(T) values evaluated at the same intermonomer distances as the ones used in our previous work⁶⁷ for the sandwich, the T-shaped and the PD dimer configurations, we find that the two approaches agree to within 0.02 kcal mol⁻¹. All calculations reported in this work were carried out using Q-Chem 2.0,⁵⁹ PSI 3.2,⁵⁸ and MOLPRO⁷³ programs.

6.3 Results and Discussion

Sandwich Configuration

The potential energy curves for the sandwich configuration of benzene dimer are plotted in Figure 6.2 along with the Δ CCSD(T) correction. At the MP2 level, the aug-cc-pVTZ and aug-cc-pVQZ* curves are nearly parallel and give nearly the same equilibrium intermonomer distances of 3.7 Å. The aug-cc-pVDZ* curve is also parallel to the above curves and gives a slightly larger equilibrium intermonomer distance of 3.8 Å (see Table 6.1). The aug-cc-pVTZ basis stabilizes the dimer by 0.42 kcal mol⁻¹ relative to the much smaller aug-cc-pVDZ* basis at their corresponding minima, with the difference in binding energies being larger than 1 kcal mol⁻¹ at shorter intermonomer distances (3.2 Å or less). The aug-cc-pVQZ* basis stabilizes the dimer by only an additional 0.10 kcal mol⁻¹ compared to the aug-cc-pVTZ basis at the corresponding minima, and by about 0.2 kcal mol⁻¹ at shorter intermonomer distances. This suggests that our high quality MP2/aug-cc-pVQZ* potential energy curves are very close to the complete basis set (CBS) limit at the MP2 level.

Table 6.1. Benzene Dimer Geometries (R in Å)^a

Method	Basis	S	T	PD	
				R ₁	R ₂
MP2	aug-cc-pVDZ* ^b	3.8	5.0	3.4	1.6
	aug-cc-pVTZ	3.7	4.9	3.4	1.6
	aug-cc-pVQZ* ^c	3.7	4.9	3.4	1.6
CCSD(T)	aug-cc-pVDZ*	4.0	5.1	3.6	1.8
Estcd.	CCSD(T)/aug-cc-pVQZ*	3.9	5.0	3.6	1.6

^a All intermonomer parameters, in angstroms, obtained using the best estimates of monomer geometry (C-C = 1.3915, C-H = 1.0800Å).

^b This is an aug-cc-pVDZ basis without diffuse functions on hydrogen.

^c This is an aug-cc-pVQZ basis without g functions on carbon and f functions on hydrogen.

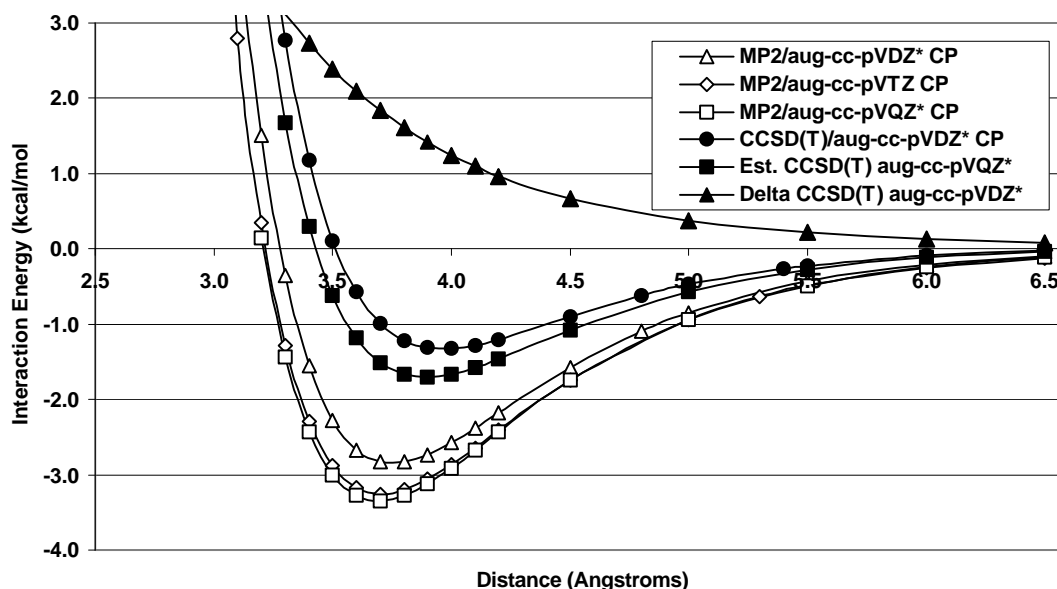


Figure 6.2. MP2 and CCSD(T) potential energy curves for the sandwich configuration of the benzene dimer.

To better account for electron correlation, the CCSD(T) potential energy curve was computed using the aug-cc-pVDZ* basis set. By computing the Δ CCSD(T) correction in that basis, and adding it to the MP2/aug-cc-pVQZ* results, an estimate of the CCSD(T)/aug-cc-pVQZ* PEC was obtained. It is clear from Figure 6.2 that Δ CCSD(T) is very large at smaller R (e.g. Δ CCSD(T) = -3.57 kcal mol⁻¹ at R = 3.2 Å), and it remains large (-1.42 kcal mol⁻¹) at the estimated CCSD(T)/aug-cc-pVQZ* equilibrium geometry (R = 3.9 Å). This confirms earlier observations^{32,46,67} that MP2 overestimates the electron correlation energy compared to CCSD(T) at different intermonomer geometries. Additionally, the difference between the CCSD(T)/aug-cc-pVDZ* and the estimated CCSD(T)/aug-cc-pVQZ* interaction energies is larger than one kcal mol⁻¹ at smaller R, and is about 0.4 kcal mol⁻¹ at the estimated CCSD(T)/aug-cc-

pVQZ* equilibrium geometry. This highlights the need to use larger basis sets to obtain CCSD(T) potential curves since the smaller ones (like aug-cc-pVDZ*) are far from saturation.

Our current best estimate of the binding energy for the sandwich benzene dimer, evaluated at the equilibrium intermonomer distance ($R = 3.9 \text{ \AA}$) of the CCSD(T)/aug-cc-pVQZ* PEC, gives a binding energy of $1.70 \text{ kcal mol}^{-1}$; this is only $0.1 \text{ kcal mol}^{-1}$ smaller than the estimated complete basis set (CBS) CCSD(T) binding energy of $1.81 \text{ kcal mol}^{-1}$ from our previous work⁶⁷ evaluated at the MP2/aug-cc-pVTZ optimized dimer geometry of $R = 3.7 \text{ \AA}$ (see Table 6.2 below).

Table 6.2. Binding Energies (kcal mol^{-1}) for Different Configurations of Benzene Dimer^a

Method	Basis	S	T	PD
MP2	aug-cc-pVDZ* ^b	2.83	3.00	4.12
	aug-cc-pVTZ	3.25	3.44	4.65
	aug-cc-pVQZ* ^c	3.35	3.48	4.73
	aug-cc-pVQZ ^d	3.37	3.54	4.79
CCSD(T)	aug-cc-pVDZ*	1.33	2.24	2.22
Estd. CCSD(T)/aug-cc-pVQZ*		1.70	2.61	2.63
Estd. CBS CCSD(T) ^{d,e}		1.81	2.74	2.78

^a Unless otherwise noted, all calculations used the optimized dimer geometry at each level of theory, and rigid monomer geometry ($\text{C-C} = 1.3915$, $\text{C-H} = 1.0800 \text{ \AA}$).

^b This is an aug-cc-pVDZ basis without diffuse functions on hydrogen.

^c This is an aug-cc-pVQZ basis without g functions on carbon and f functions on hydrogen.

^d At the MP2/aug-cc-pVTZ optimized dimer geometry and using the best estimates of monomer geometry ($\text{C-C} = 1.3915$, $\text{C-H} = 1.0800 \text{ \AA}$).

^e See Chapter 3 of this thesis for full details.

T-shaped Configuration

The PEC's for the T-shaped configuration of benzene dimer are plotted in Figure 6.3 along with the Δ CCSD(T) correction. At the MP2 level, we see that the aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ* curves are nearly parallel and give similar equilibrium intermonomer distances of 5.0 Å, 4.9 Å, and 4.9 Å, respectively (see Table 6.1). By examining Figure 6.3 and Table 6.2, we see that at the MP2 level, the aug-cc-pVTZ basis stabilizes the dimer by 0.44 kcal mol⁻¹ relative to the aug-cc-pVDZ* basis at their corresponding minima; the difference in binding energies is larger at shorter R. The aug-cc-pVQZ* basis stabilizes the dimer by a small 0.04 kcal mol⁻¹ compared to the aug-cc-pVTZ basis at equilibrium, again showing that the MP2/aug-cc-pVQZ* curves are very close to the CBS limit at the MP2 level.

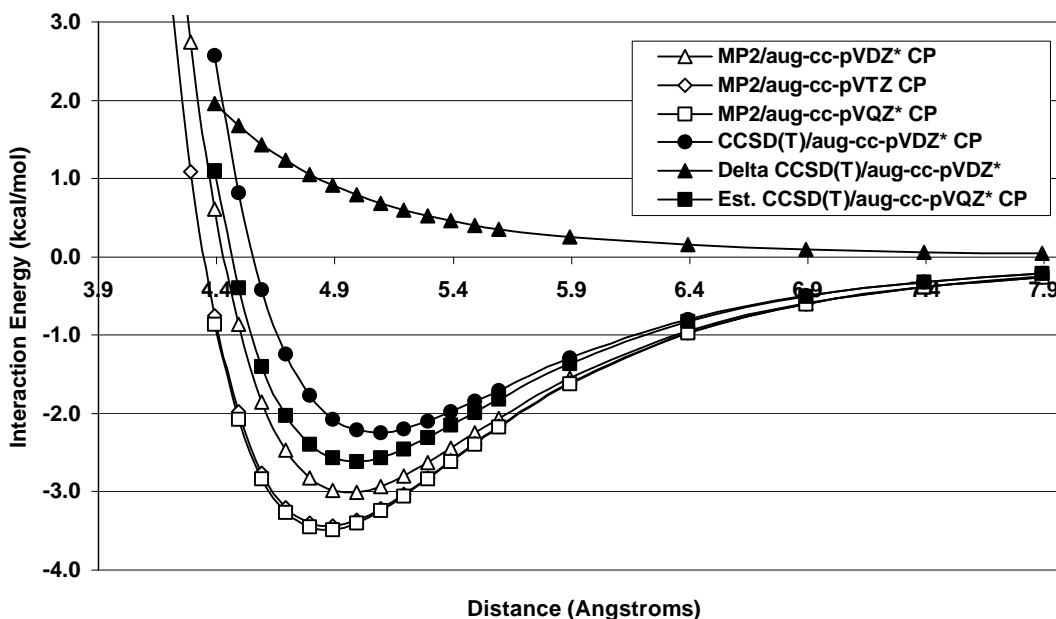


Figure 6.3. MP2 and CCSD(T) potential energy curves for the T-shaped configuration of the benzene dimer.

The equilibrium intermonomer distances are 5.1 Å and 5.0 Å at the CCSD(T)/aug-cc-pVDZ* and the estimated CCSD(T)/aug-cc-pVQZ* levels of theory, respectively. This is in good accord with the microwave results of Arunan and Gutowsky,³⁸ who found a distance of 4.96 Å between the centers of mass of gas-phase benzene dimer. The difference between the CCSD(T) and the MP2 equilibrium geometries is in agreement with the trend observed with the sandwich dimer, where the CCSD(T) equilibrium distances were found to be 0.1-0.3 Å larger than the MP2 ones.

The Δ CCSD(T) correction is large at R smaller than equilibrium (e.g. Δ CCSD(T) is about -2 kcal mol⁻¹ at R = 4.4 Å). At the estimated CCSD(T)/aug-cc-pVQZ* equilibrium geometry, Δ CCSD(T) is -0.79 kcal mol⁻¹, compared with a much larger value of -1.42 kcal mol⁻¹ for the sandwich configuration of benzene dimer. The difference between CCSD(T)/aug-cc-pVDZ* and estimated CCSD(T)/aug-cc-pVQZ* values is larger than one kcal mol⁻¹ at smaller R, about 0.4 kcal mol⁻¹ at equilibrium R = 5.0 Å, and is less than 0.1 kcal mol⁻¹ for R > 6 Å.

Burley and Petsko⁴ observed that, in proteins, aromatic side chains separated by distances ranging from 4.5-7.0 Å and dihedral angles near 90° are the most common. Results from our estimated CCSD(T)/aug-cc-pVQZ* curve (see Figure 6.3) for the T-shaped dimer (where the dihedral angle between the benzene planes is 90°) are in agreement with these observations. At this level of theory, dimers separated by intermonomer distances larger than 4.5 Å have favorable binding energies, with the binding energy becoming small (<0.4 kcal mol⁻¹) for dimers separated by a distance R greater than 7.0 Å.

Our current best estimate of the binding energy for the T-shaped benzene dimer, evaluated at the equilibrium $R = 5.0 \text{ \AA}$ of the CCSD(T)/aug-cc-pVQZ* PEC, gives a binding energy of $2.61 \text{ kcal mol}^{-1}$, which is only $0.13 \text{ kcal mol}^{-1}$ smaller than the estimated complete basis set (CBS) CCSD(T) binding energy of $2.74 \text{ kcal mol}^{-1}$ from our previous work⁶⁷ (see Table 6.2).

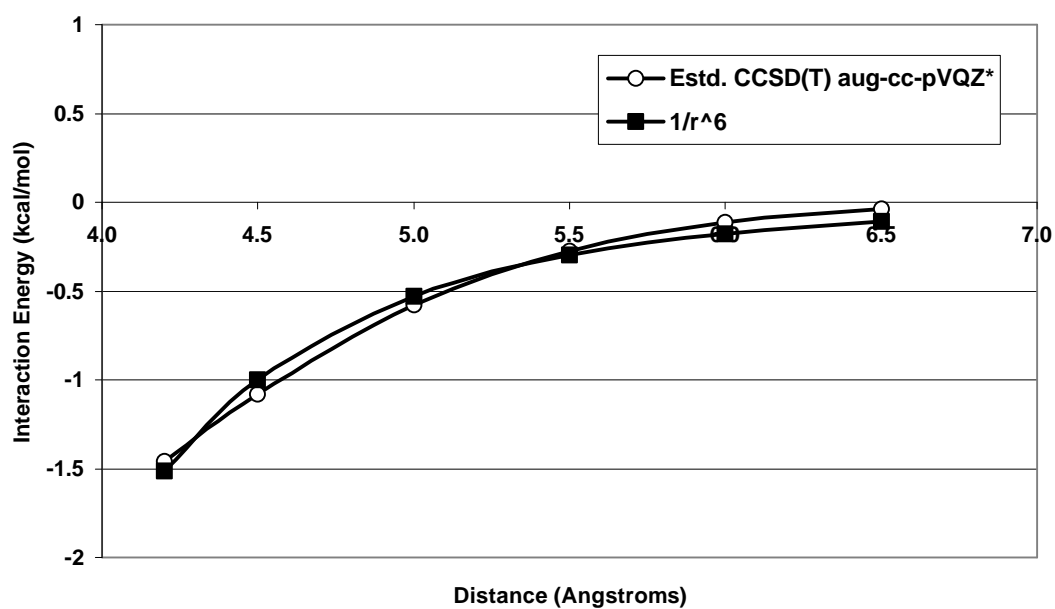


Figure 6.4. Asymptotic $1/r^6$ fit of the estimated CCSD(T)/aug-cc-pVQZ* potential energy curve for the sandwich configuration of the benzene dimer.

Since the potential of the attractive van der Waals interactions falls off with distance as $1/R^6$, we performed a least-squares fit of the tail of the potential energy curve for the sandwich and T-shaped configurations using the equation $DE = -B/R^6$, where B is a constant to be determined from the binding energy fit. For the sandwich benzene dimer,

we have fitted the estimated CCSD(T)/aug-cc-pVQZ* curve from $R = 4.2 \text{ \AA}$ to $R = 6.5 \text{ \AA}$ (see Figure 6.4) and obtained a B value of $8.3 \times 10^3 \text{ kcal mol}^{-1} \text{ \AA}^6$. Similarly, for the T-shaped dimer, we have fitted the estimated CCSD(T)/aug-cc-pVQZ* curves from $R = 5.3 \text{ \AA}$ to $R = 7.9 \text{ \AA}$ (see Figure 6.5) and obtained a B value of $5.39 \times 10^4 \text{ kcal mol}^{-1} \text{ \AA}^6$.

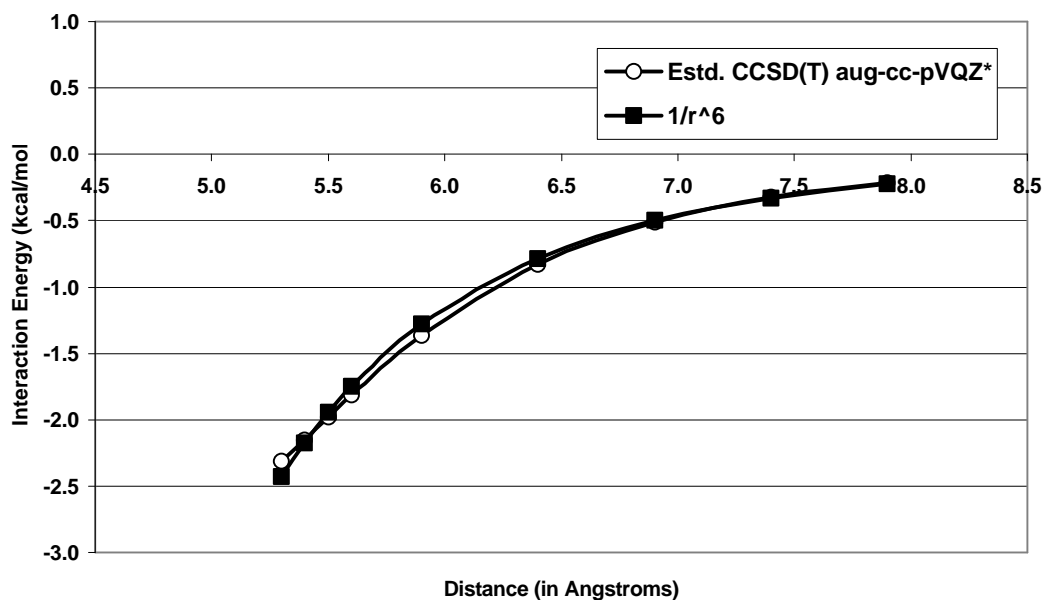


Figure 6.5. Asymptotic $1/r^6$ fit of the estimated CCSD(T)/aug-cc-pVQZ* potential energy curve for the T-shaped configuration of the benzene dimer.

Parallel-displaced Configuration

Figures 6.6 and 6.7 present the potential energy curves for the parallel-displaced configuration of benzene dimer at the MP2/aug-cc-pVTZ and the MP2/aug-cc-pVQZ* levels of theory (MP2/aug-cc-pVDZ* PEC's are not displayed). At the MP2 level, the

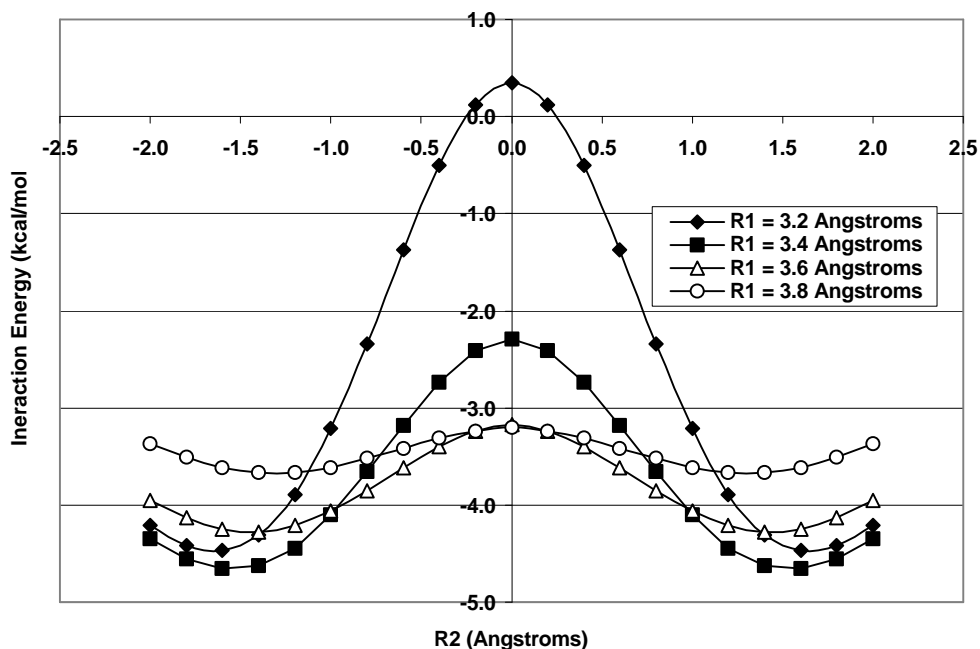


Figure 6.6. Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) MP2/aug-cc-pVTZ level of theory.

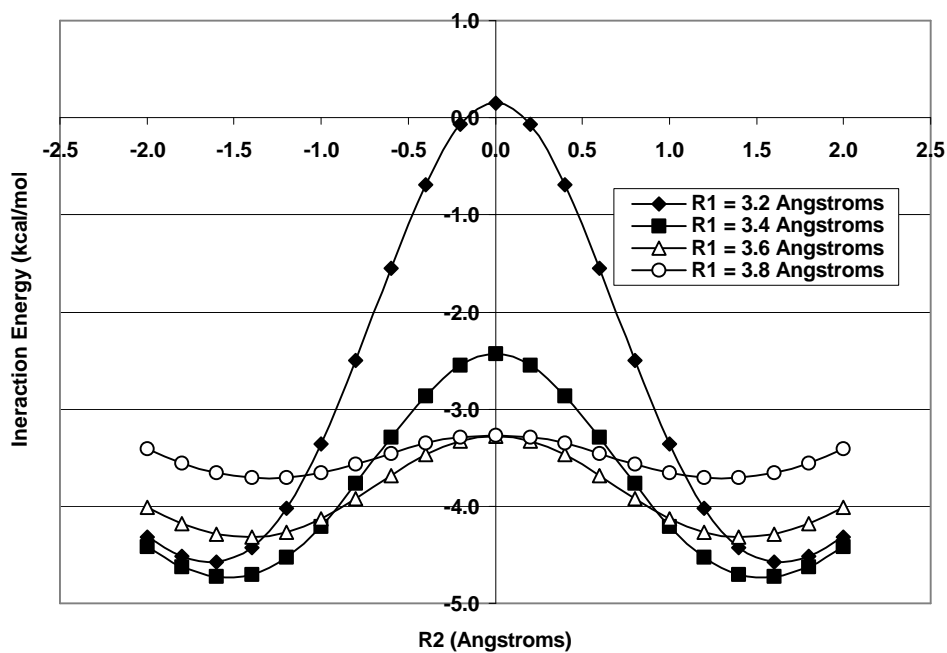


Figure 6.7. Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) MP2/aug-cc-pVQZ* level of theory.

CP-corrected aug-cc-pVDZ*, aug-cc-pVTZ, and aug-cc-pVQZ* curves are nearly parallel, and give nearly the same equilibrium intermonomer distance of $R_1 = 3.4 \text{ \AA}$ and $R_2 = 1.6 \text{ \AA}$, suggesting again that smaller basis sets such as aug-cc-pVDZ* are sufficient for intermonomer geometry optimizations. With regard to the MP2 binding energies, the aug-cc-pVTZ basis stabilizes the dimer by $0.43 \text{ kcal mol}^{-1}$ relative to the aug-cc-pVDZ* basis at their corresponding minima, with the aug-cc-pVQZ* basis adding another $0.08 \text{ kcal mol}^{-1}$ to the equilibrium binding energy. As shown in Figures 6.6-6.9, the sandwich configuration of benzene dimer represents a maximum (saddle point) along the horizontal displacement coordinate R_2 which connects two equivalent parallel-displaced benzene dimer configurations.

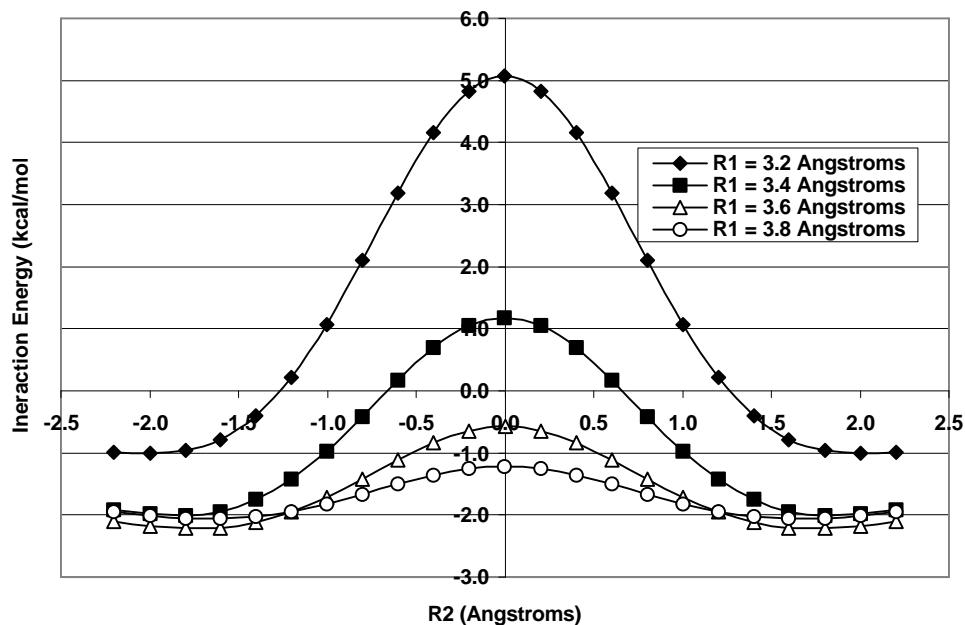


Figure 6.8. Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) CCSD(T)/aug-cc-pVDZ* level of theory.

Figures 6.8 and 6.9 display the CCSD(T)/aug-cc-pVDZ* and the estimated CCSD(T)/aug-cc-pVQZ* potential energy curves, respectively (for clarity Δ CCSD(T) curves are not shown). In contrast to the MP2 potential energy curves, these two curves are not totally parallel, with equilibrium distances of $R_1 = 3.6 \text{ \AA}$ and 3.6 \AA , and $R_2 = 1.8 \text{ \AA}$ and 1.6 \AA , respectively. These results are consistent with the sandwich and T-shaped cases, where the CCSD(T) equilibrium distances were 0.1-0.3 \AA larger than the MP2 values.

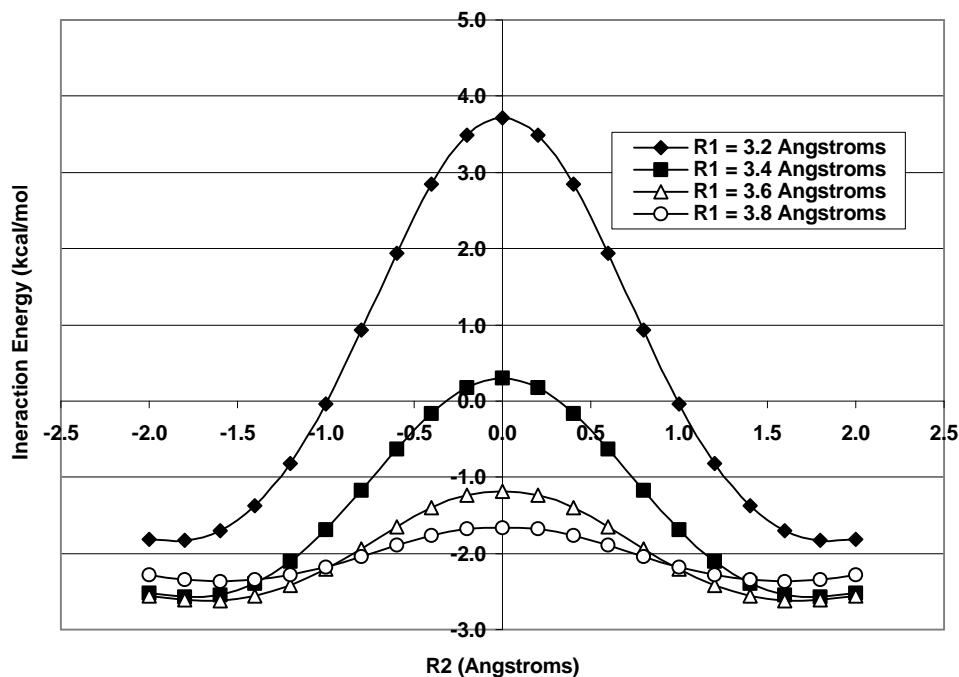


Figure 6.9. Potential energy curves for the parallel-displaced configuration of benzene dimer at the (counterpoise-corrected) estimated CCSD(T)/aug-cc-pVQZ* level of theory.

At the estimated CCSD(T)/aug-cc-pVQZ* equilibrium geometry, $\Delta\text{CCSD(T)}$ is $1.67 \text{ kcal mol}^{-1}$, with larger values at shorter intermonomer distances. This correction is comparable in size to the sandwich $\Delta\text{CCSD(T)}$, especially at smaller R_2 values, and is about twice as large as the T-shaped one. The CCSD(T)/aug-cc-pVQZ* equilibrium binding energy is $2.63 \text{ kcal mol}^{-1}$, and this is only $0.15 \text{ kcal mol}^{-1}$ smaller than the estimated complete basis set (CBS) CCSD(T) binding energy of $2.78 \text{ kcal mol}^{-1}$ from our previous work. These results, combined with the sandwich and the T-shaped ones, indicate that the CCSD(T)/aug-cc-pVQZ* potential energy curves should be accurate to a few tenths of a kilocalorie per mole from the ab initio limit.

6.4 Conclusions

In this work, we have generated high-quality potential energy curves for the sandwich, T-shaped, and parallel-displaced configurations of the simplest prototype of aromatic π - π interactions, the benzene dimer. At the MP2 level, the aug-cc-pVTZ and aug-cc-pVQZ* (truncated aug-cc-pVQZ) basis sets with 828 and 1128 basis functions, respectively, are much larger than previous basis sets used to compute potential energy curves (PEC's) for benzene dimer. We find that even though the intermonomer geometries can be accurately predicted using smaller basis sets at the MP2 level, the binding energies are sensitive to the improvement of the basis set. Therefore, large basis sets (such as the aug-cc-pVQZ* basis) are needed to approach the complete basis set (CBS) limit at the MP2 level to within a few tenths of one kcal mol^{-1} . By combining the MP2/aug-cc-pVQZ* results with a correction for the difference between the CCSD(T) and MP2 interaction energies (the $\Delta\text{CCSD(T)}$ correction) determined in a smaller basis,

estimates of the CCSD(T)/aug-cc-pVQZ* potential energy curves were obtained. The Δ CCSD(T) correction is large at distances around or shorter than the equilibrium distance. At the estimated CCSD(T)/aug-cc-pVQZ* equilibrium geometries, Δ CCSD(T) is $-0.79 \text{ kcal mol}^{-1}$ for the T-shaped configuration, compared with a much larger values of $-1.42 \text{ kcal mol}^{-1}$ and $-1.67 \text{ kcal mol}^{-1}$ for the sandwich and parallel-displaced configurations of benzene dimer, respectively.

By comparing the equilibrium binding energies of the estimated CCSD(T)/aug-cc-pVQZ* PEC's with the estimated complete basis set CCSD(T) equilibrium binding energies,⁶⁷ we conclude that our CCSD(T)/aug-cc-pVQZ* PEC's for the three dimer configurations are accurate to within a few tenths of one kilocalorie per mole of the ab initio limit. By performing a least-squares fit of the tail of the potential energy curve for the sandwich and T-shaped configurations, we find a good correlation between the interaction energy and a $-B/R^6$ function, where R represents the intermonomer distance.

For all dimer configurations studied, the CCSD(T) equilibrium intermonomer distances R (or R_1 for the parallel-displaced configuration) were found to be 0.1-0.3 Å larger than the MP2 values. For the T-shaped dimer, the equilibrium intermonomer distances are in good accord with the microwave results of Arunan and Gutowsky³⁸ (who found a distance of 4.96 Å between the centers of mass of gas-phase benzene dimer) and with the observed mean distance of 5.05 Å between the phenyl ring centroids for interacting aromatic side chains in proteins.⁴ Finally, we expect these high-quality potential energy curves to assist in the development of approximate methods that are computationally inexpensive and capable of modeling π - π interactions in biomolecules.

CHAPTER 7

SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS

In this work, we have applied state-of-the-art ab initio electronic structure methods to elucidate the nature of aromatic π - π interactions. These interactions are of great importance in many areas of science and molecular engineering processes and play a key role in biological and molecular recognition.

We first considered the simplest prototype of aromatic π - π interactions, the benzene dimer. We have provided definitive high-quality ab initio estimates of the geometries and the first converged results of the binding energies for the sandwich, T-shaped, and parallel-displaced configurations of benzene dimer. Because a major component of the binding energy of benzene dimer is due to London dispersion interactions, Hartree-Fock molecular orbital theory and density functional theory are both incapable of describing the instantaneous charge fluctuations that give rise to dispersion forces. Therefore, we employed the second-order Møller-Plesset perturbation theory (MP2) and the coupled-cluster theory with perturbative triples [CCSD(T)] to obtain a highly accurate description of benzene dimer. In particular, we have investigated the effects of basis sets and correlation methods on the geometries and binding energies. At the MP2 level, the aug-cc-pVTZ basis set is much larger than any basis sets employed in previous optimizations of benzene dimer. For the T-shaped configuration, where experimental data is available, our MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ

geometries are in very good agreement with the microwave results of Arunan and Gutowsky.³⁸

We have used basis sets as large as aug-cc-pVQZ along with the MP2 method and obtained the first converged values of the binding energies for benzene dimer. A new shared-memory parallel implementation of the explicitly correlated MP2-R12/A method was developed to estimate the complete basis set (CBS) limit at the MP2 level. To account for higher-order electron correlations, these results were combined with a correction for the difference between CCSD(T) and MP2 in a smaller basis set to estimate the CBS CCSD(T) results which should provide binding energies that are accurate to within a few tenths of one kcal mol⁻¹.

Our best estimates of the binding energies indicate that the T-shaped and parallel-displaced dimers are close in energy with the parallel-displaced configuration slightly more bound. The sandwich configuration is ~1 kcal mol⁻¹ higher in energy. Our binding energy values of $D_0 = 2.4$ kcal mol⁻¹ and 2.7 kcal mol⁻¹ for the T-shaped and parallel-displaced configurations, respectively, strongly indicate that the experimental binding energy of Krause and co-workers ($D_0 = 1.6 \pm 0.2$ kcal mol⁻¹) is too small, and support an older, less commonly cited results of Grover and co-workers ($D_0 = 2.4 \pm 0.4$ kcal mol⁻¹).

The close binding energies of the T-shaped and the PD configurations of benzene dimer, and the higher binding energy of the sandwich configuration are entirely consistent with crystal structure studies of proteins where the interacting pairs of phenylalanines are found in mostly T- and PD-like configurations.

Guided by our success with benzene dimer, we have examined the role of substituents in tuning π - π interactions. An understanding of this issue is of prime

importance for molecular design and supramolecular synthesis efforts. Because none of the experiments examining the role of substituents were performed in the gas phase, it is very hard to decouple the intrinsic binding energy from contributions due to the solvent or environment, which will change from system to system. Additionally, due to secondary intramolecular interactions or steric constraints, the model system itself may complicate the interpretation of results.⁸⁶ Our studies on benzene-monosubstituted benzene dimers were designed to present a high-level systematic theoretical investigation of substituent effects on π - π interactions. In these studies, we first examined the sandwich dimer configurations of benzene with substituted benzenes, with substituents OH, CH₃, F, and CN. Using frozen monomer geometries, intermonomer distances were optimized at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory. These results were combined with the CCSD(T) results (in a smaller basis) to estimate high-quality CCSD(T)/aug-cc-pVTZ binding energies.

According to the Hunter-Sanders electrostatic rules, for face-to-face sandwich dimers, the electron-withdrawing substituents will cause the binding to become more favorable by reducing the electrostatic repulsion between the two negatively charged π -clouds of monomers. Electron-donating substituents will have the opposite effect. Our MP2 and CCSD(T) results using different basis sets show that *all* of the substituted sandwich dimers bind more strongly than benzene dimer regardless of whether the substituents are considered electron donating (OH, CH₃) or electron withdrawing (CN, F), in apparent contradiction to the Hunter-Sanders model.² To gain increased understanding of substituents effects in π - π interactions, and to help uncover the origin of the above trends in binding energies, we have employed symmetry adapted perturbation

theory (SAPT) to analyze the interaction energy in terms of physically meaningful components such as electrostatic, induction, dispersion, and exchange-repulsion energies. This analysis demonstrates that, contrary to the predictions of the Hunter-Sanders rules, electrostatics alone are insufficient to predict the correct trends in binding. The increased binding in benzene-toluene (compared to benzene dimer) is mainly attributed to more favorable dispersion interactions. For benzene phenol, dispersion is twice as important as electrostatics in contributing to the stabilization relative to benzene dimer. Only for heterodimers containing the electron withdrawing F and CN substituents are electrostatic interactions the dominant contributor to their increased stability.

In addition to the study of sandwich monosubstituted dimers, we have also explored how the same substituents (OH, CH₃, F, and CN) affect the binding of two T-shaped configurations (see Chapter 5). For these dimers, we have found that dispersion is less important in determining relative energies than for the sandwich dimers, with the shifts in the exchange energies usually 2-4 times larger. For one of the T-shaped configurations (denoted by “T-shaped” in the text), both exchange-corrected induction and electrostatic energies relative to benzene dimer are found to be small except for benzene-benzonitrile dimer, where the relative change in electrostatic energy is fairly large. For the sandwich and the T-shaped(2) dimers, a strong correlation is observed between their dispersion energy components and the computed π polarizability of the monosubstituted monomers.

In various complex chemical and biological systems, aromatic rings can be found at different orientations and distances from each other which might not correspond to the potential energy minima for π - π interactions. Nevertheless, the aromatic rings might still

interact favorably enough to contribute significantly to the overall stability of the system. Therefore, it is essential to obtain the potential energy curves for prototype systems in order to determine how π - π interactions depend on both the orientation and distance between these aromatic rings. To achieve that goal, we have computed the potential energy curves as a function of intermonomer distances for three important configurations of benzene dimer, namely the sandwich (S), the T-shaped (T), and the parallel-displaced (PD) configurations. Both the MP2 and the CCSD(T) methods combined with different basis sets were used in this study. Our best curves for the T-shaped configuration are in agreement with the observed geometric orientations of aromatic rings in proteins⁴ and with the microwave results of Arunan and Gutowsky.³⁸ By comparing the current results with the estimated complete basis set CCSD(T) equilibrium binding energies,⁶⁷ we concluded that our best CCSD(T) potential energy curves are accurate to within a few tenths of one kilocalorie per mole of the ab initio limit. We expect these high-quality potential energy curves to serve as benchmarks and assist in the development of approximate methods that are computationally inexpensive and capable of modeling π - π interactions in biomolecules.

Future theoretical research in the area of aromatic π - π interactions shall investigate different parallel-displaced configurations of monosubstituted dimers. Lower level calculations can be performed to investigate different locations for placing the substituents, and then determine the optimized geometries and binding energies for the most stable dimers. These computations will be more demanding than the sandwich and T-shaped ones, because of the lower symmetry of the dimers, and the multiple locations where the substituents can be placed. Also, for the parallel-displaced configurations, both

the horizontal and vertical intermonomer distances need to be optimized, adding to the challenge of accurately predicting geometries and binding energies for these dimers.

It is also very important to investigate the role of multiple substituents in tuning π - π interactions. Of particular interest is the effect of successive fluorination of one benzene ring on the geometry and binding energy of benzene-fluorinated benzene dimers. An experimental study by Cozzi and co-workers⁷¹ indicates that each fluorine added to the aromatic ring of substituted 1,8-diarylnaphthalenes contributes an extra 0.5-0.6 kcal mol⁻¹ of binding energy. Our preliminary MP2/aug-cc-pVDZ computations on benzene-fluorinated benzene dimers all the way up to benzene-hexafluorobenzene give similar results, with each fluorine contributing an additional 0.5-0.6 kcal mol⁻¹ to the binding energy. This additivity of the binding energy can also be tested to see if it holds with other substituents (e.g. OH, CH₃, CN), where dimers with mixed substituents are used. Since our study on the substituent effects showed that the binding energies relative to benzene dimer are accurately predicted at *any* of the levels of theory used, lower levels of theory, such as MP2/aug-cc-pVDZ should be sufficient in the study of additivity effects.

We expect that a better understanding of π - π interactions and how they may be tuned will play a major role in the advancement of rational supermolecular design. Finally, future development of approximate methods (such as molecular mechanics force fields) that are computationally inexpensive and capable of modeling π - π interactions in larger systems will require very accurate benchmark results for prototype systems, such as benzene dimer.

REFERENCES

- (1) Mulliken, R. S. *J. Am. Chem. Soc.* **1952**, *74*, 811.
- (2) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525.
- (3) Kumpf, R. A.; Dougherty, D. A. *Science* **1993**, *261*, 1708.
- (4) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23.
- (5) Hunter, C. A.; Singh, J.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 837.
- (6) Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18.
- (7) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (8) Zimm, B. H. *J. Chem. Phys.* **1960**, *33*, 1349.
- (9) Crothers, D. M.; Zimm, B. H. *J. Mol. Biol.* **1967**, *9*, 1.
- (10) Hunter, C. A. *Chem. Soc. Rev.* **1994**, *23*, 101.
- (11) Kryger, G.; Silman, I.; Sussman, J. L. *J. Physiol.* **1998**, *92*, 191.
- (12) Claessens, C. G.; Stoddart, J. F. *J. Phys. Org. Chem.* **1997**, *10*, 254.
- (13) Glaser, R.; Dendi, L. R.; Knotts, N.; Barnes, C. L. *Crystal Growth & Design* **2003**, *3*, 291.
- (14) Lehn, J., -M. *Supramolecular Chemistry: Concepts and Perspectives*; VCH: New York, 1995.
- (15) Kutzelnigg, W.; Klopper, W. *J. Chem. Phys.* **1991**, *94*, 1985.
- (16) Klopper, W. *Chem. Phys. Lett.* **1991**, *186*, 583.
- (17) Valeev, E. F.; Schaeffer, H. F. *J. Chem. Phys.* **2000**, *113*, 3990.
- (18) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638.
- (19) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.
- (20) Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Rybak, S. *J. Chem. Phys.* **1993**, *98*, 1279.

- (21) Levine, I. N. *Quantum Chemistry*; 4th ed.; Prentice-Hall, Inc.: New Jersey, 1991.
- (22) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons Ltd.: West Sussex, England, 2003.
- (23) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: New York, 1996.
- (24) Boys, S. F. *Proc. R. Soc. London* **1950**, *542*, A200.
- (25) Scuseria, G. E.; Lee, T. J. *J. Chem. Phys.* **1990**, *93*, 5851.
- (26) Bukowski, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **1996**, *104*, 3306.
- (27) Mas, E. M.; Szalewicz, K. *J. Chem. Phys.* **1996**, *104*, 7606.
- (28) Rybak, S.; Szalewicz, K.; Jeziorski, B. *J. Chem. Phys.* **1989**, *91*, 4779.
- (29) Williams, H. L.; Korona, T.; Bukowski, R.; Jeziorski, B.; Szalewicz, K. *Chem. Phys. Lett.* **1996**, *262*, 431.
- (30) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorski, B.; Jeziorska, M.; Kucharski, S. A.; Misquitta, A. J.; Moszynski, R.; Patkowski, K.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wormer, P. E. S. *SAPT2002: An Ab Initio Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies. Sequential and Parallel Versions.*
- (31) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 9751.
- (32) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780.
- (33) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500.
- (34) Dahl, T. *Acta Chem. Scand.* **1994**, *48*, 95.
- (35) Cox, E. G.; Cruikshank, D. W. J.; Smith, J. A. S. *Proc. R. Soc. London* **1958**, *A247*, 1.
- (36) Janada, K. C.; Hemminger, J. C.; Winn, J. S.; Novich, S. E.; Harris, S. J.; Klemperer, W. *J. Chem. Phys.* **1975**, *63*, 1419.
- (37) Steed, J. M.; Dixon, T. A.; Klemperer, W. *J. Chem. Phys.* **1979**, *70*, 4940.
- (38) Arunan, E.; Gutowsky, H. S. *J. Chem. Phys.* **1993**, 4294.
- (39) Ventura, V. A.; Felker, P. M. *J. Chem. Phys.* **1993**, *99*, 748.
- (40) Felker, P. M.; Maxton, P. M.; Schaeffer, M. W. *Chem. Rev.* **1994**, *94*, 1787.

- (41) Law, K. S.; Schauer, M.; Bernstein, E. R. *J. Chem. Phys.* **1984**, *81*, 4871.
- (42) Börnsen, K. O.; Selzle, H. L.; Schlag, E. W. *J. Chem. Phys.* **1986**, *85*, 1726.
- (43) Scherzer, W.; Krätzschar, O.; Selzle, H. L.; Schlag, E. W. *Z. Naturforsch* **1992**, *47A*, 1248.
- (44) Krause, H.; Ernstberger, B.; Neusser, H. J. *Chem. Phys. Lett.* **1991**, *184*, 411.
- (45) Grover, J. R.; Walters, E. A.; Hui, E. T. *J. Phys. Chem.* **1987**, *91*, 3233.
- (46) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
- (47) Tsuzuki, T.; Uchimaru, T.; Tanabe, K. *J. Mol. Struct. (THEOCHEM)* **1994**, *307*, 107.
- (48) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **2000**, *319*, 547.
- (49) Tsuzuki, S.; Luthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.
- (50) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104.
- (51) Kristyàn, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.
- (52) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (53) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (54) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (55) Gauss, J.; Stanton, J. F. *J. Phys. Chem. A* **2000**, *104*, 2865.
- (56) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
- (57) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.
- (58) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. T.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P.; Allen, W. D. In *PSI 3.2*, 2003.
- (59) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Daschel, H.; Zhang, W.; Korambath, P. P.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Baker, J.; Byrd, E. F. C.; Voorhis, T. V.; Oumi, M.; Hirata, S.; Hsu, C.-P.;

- Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532.
- (60) Colvin, M. E.; Janssen, C. L.; Whiteside, R. A.; Tong, C. H. *Theor. Chim. Acta* **1993**, *84*, 301.
- (61) Nielsen, I. M. B.; Seidl, E. T. *J. Comput. Chem.* **1995**, *16*, 1301.
- (62) Nielsen, I. M. B. *Chem. Phys. Lett.* **1996**, *225*, 210.
- (63) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.
- (64) Hobza, P.; Selzle, H. L.; Schlag, E. W. *Collect. Czech. Chem. Commun.* **1992**, *57*, 1186.
- (65) Sun, S.; Bernstein, E. R. *J. Phys. Chem.* **1996**, *100*, 13348.
- (66) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 1210.
- (67) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887.
- (68) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc. Perkin Trans. 2* **2001**, 651.
- (69) Chipot, C.; Jaffe, R.; Maignet, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217.
- (70) Cozzi, F.; Siegel, J. S. *Pure Appl. Chem.* **1995**, *67*, 683.
- (71) Cozzi, F.; Ponzini, F.; Annunziata, R.; Cinquini, M.; Siegel, J. S. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019.
- (72) Rashkin, M. J.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 1860.
- (73) Amos, R. D.; Bernhardsson, A.; Berning, A.; Celani, P.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Knowles, P. J.; Korona, T.; Lindh, R.; Lloyd, A. W.; McNicholas, S. J.; Manby, F. R.; Meyer, W.; Mura, M. E.; Nicklaß, A.; Palmieri, P.; Pitzer, R.; Rauhut, G.; Schütz, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Werner, H.-J. *MOLPRO, a package of ab initio programs designed by Werner, H. -J.; Knowles, P. J. Version 2002.*
- (74) Lide, D. R. *CRC Handbook of Chemistry and Physics*; 77th ed.; CRC Press: Boca Raton, FL, 1997.
- (75) *SPARTAN '02*, Wavefunction Inc., 18401 Von Karman Ave., Ste. 370, Irvine, CA 92612, U. S.

- (76) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.
- (77) Mecozzi, S.; West, A. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10566.
- (78) McGaughey, G. B.; Gagné, M.; Rappé, A. K. *J. Biol. Chem.* **1998**, *273*, 15458.
- (79) Cozzi, F.; Cinquini, M.; Annunziata, R.; Dwyer, T.; Siegel, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 5729.
- (80) Cozzi, F.; Cinquini, M.; Annunziata, R.; Siegel, J. S. *J. Am. Chem. Soc.* **1993**, *115*, 5330.
- (81) Paliwal, S.; Geib, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1994**, *116*, 4497.
- (82) Kim, E.; Paliwal, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 11192.
- (83) Adams, H.; Carver, F. J.; Hunter, C. A.; Morales, J. C.; Seward, E. M. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1542.
- (84) Carver, F. J.; Hunter, C. A.; Livingstone, D. J.; McCabe, J. F.; Seward, E. M. *Chem. Eur. J.* **2002**, *8*, 2848.
- (85) Nakamura, K.; Houk, K. N. *Org. Lett.* **1999**, *1*, 2049.
- (86) Ribas, J.; Cubero, E.; Luque, F. J.; Orozco, M. *J. Org. Chem.* **2002**, *67*, 7057.
- (87) Hong, B. H. M. S. Thesis, Pohang University of Science and Technology, 1999.
- (88) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A.* **2003**, *107*, 8377.
- (89) Hopkins, B. W.; Tschumper, G. S. *J. Phys. Chem.* **2004**, *108*, 2941.
- (90) Hammett, L. P. *J. Am. Chem. Soc.* **1937**, *59*, 96.
- (91) Tarakeshwar, P.; Kim, K. S.; Kraka, E.; Cremer, D. *J. Chem. Phys.* **2001**, *115*, 6018.
- (92) Tsuzuki, S.; Uchimaru, T.; Sugawara, K.; Mikami, M. *J. Chem. Phys.* **2002**, *117*, 11216.
- (93) Špirko, V.; Engkvist, O.; Slodán, P.; Selzle, H. L.; Schlag, E. W.; Hobza, P. *J. Chem. Phys.* **1999**, *111*, 572.
- (94) Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. *J. Chem. Phys.* **1995**, *103*, 1481.
- (95) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515.

- (96) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A.* **2004**, *108*, 621.
- (97) Liu, Y. P.; Kim, K. S.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.
- (98) Yin, D.; MacKerrel Jr., A. D. *J. Comput. Chem.* **1998**, *19*, 334.
- (99) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *J. Phys. Chem.* **1994**, *98*, 1830.