FROM SMALL TO BIG: UNDERSTANDING NONCOVALENT INTERACTIONS IN CHEMICAL SYSTEMS FROM QUANTUM MECHANICAL MODELS

A Thesis
Presented to
The Academic Faculty

by

Ashley L. Ringer

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
May 2009

FROM SMALL TO BIG: UNDERSTANDING NONCOVALENT INTERACTIONS IN CHEMICAL SYSTEMS FROM QUANTUM MECHANICAL MODELS

Approved by:

C. David Sherrill, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Jean-Luc Brédas
School of Chemistry and Biochemistry
Georgia Institute of Technology

Mostafa A. El-Sayed
School of Chemistry and Biochemistry
Georgia Institute of Technology

Stephen Harvey
School of Biology
Georgia Institute of Technology

Rigoberto Hernandez
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: 5 March 2009

Dedicated to Mrs. Joan Benton

ACKNOWLEDGEMENTS

The work in this thesis has been supported by the generosity of the Philanthropic and Educational Organization through a Scholar Award and by the American Association of University Women through a Dissertation Fellowship. However, even more valuable than the financial support these awards provided, were the encouragement and personal support I received from the amazing women in these organizations, especially the ladies of PEO Georgia Section B and the Atlanta area AAUW chapter. I sincerely appreciate the very personal investment you have made in my life, not to mention the cards, letters, coffees, lunches, and countless desserts you fed me over the last two years. I have been blessed to get to know all of you, and it was my honor to represent your organizations.

There are many other individuals to whom I owe significant thanks for enabling the completion of this work, none more so than my ever patient and encouraging adviser, Prof. David Sherrill. There is no truly sufficient way to thank you for the support, advice, and guidance you have provided to me over the last five years. Your mentorship and expertise have enabled me to develop a range of professional skills that will benefit me in every aspect of my career for years to come.

My graduate experience has been significantly enhanced by my active and very involved thesis committee. Profs. Rigoberto Hernandez, Jean-Luc Brèdas, Mostafa El-Sayed, and Stephen Harvey, you all have been an invaluable resource to me throughout my time at Georgia Tech, and I appreciate the significant interest you have taken in my work and my future. I owe particular thanks to Prof. Hernandez for his thoughtful counsel about my post-doc search and future professional career.

I have had the pleasure to work with several outstanding collaborators during the last five years, including Prof. Mutasem Sinnokrot, Anastasia Senenko, Michelle Figgs, Bill March, and Jim Waters, and Prof. Alexander Gray. I have enjoyed working with all of you and appreciate the unique skills and talents you each brought to our collaborations.

To all the gentlemen of the Sherrill group, past and present, I am very grateful for the many discussions and insights you have shared with me over the years. Dr. Arteum Bochevarov, Dr. Berhane Temelso, Stephen Arnstein, Tait Takatani, Edward Hohenstein, and Michael Marshall, it has been a pleasure to work with each of you, and I wish you all the very best. Special thanks is also owed to Dr. Edward Valeev, Dr. John Sears, and particularly Dr. Micah Abrams and Dr. Jeremy Moix for their mentorship and friendship when I was a first-year student and was especially need of their advice and encouragement.

I have also enjoyed working with and learning from many of my other colleagues in the School of Chemistry and Biochemistry including Yanping Qin, Gungor Ozer, Denise Enekwa, Megan Damm, and especially Ashley Tucker and Rusty Nicovich. You guys have been my colleagues and my friends throughout this experience, and I thank each of you. I am also deeply appreciative for the support of many other friends and colleagues throughout Georgia Tech, particularly Dr. Karen Feigh and Dr. Brian German, Dr. Maria-Isabel Carnasciali, Dr. Teena Carroll, Janine Johnson, Farhana Zamen, Dr. Shannon Watt, and Dr. Caroline Burd.

I have also received significant support from many members of the Georgia Tech campus community, who have supported my professional development over last five years, including Yvette Upton and Colleen Riggle of the Women's Resource Center, Dr. Karen Head of the Center for the Enhancement of Teaching and Learning, and Dr. Amanda Gable from the Fellowship Communication Office. Thanks to all of you for helping me discern my own role in the campus community and supporting my efforts in applying for external funding.

To my parents, David and Susan Ringer, I am forever grateful for your support and encouragement of my educational achievement. To my grandfathers, Dr. Robert Ringer and Dr. George Spencer, I appreciate your continual guidance and advice. And finally, I am forever indebted to Mrs. Joan Benton, my high school chemistry teacher to whom this thesis is dedicated, and to Prof. David Magers, my undergraduate adviser who encouraged me to pursue my Ph.D. Without your encouragement, I would never even have attempted the goals I am now achieving. Your investment in my life has forever changed it, and I am forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this thesis, I examine noncovalent interactions in complex chemical systems by considering model systems which capture the essential physics of the interactions and applying correlated correlated electronic structure techniques to these systems. Noncovalent interactions are critical to understanding a host of energetic and structural properties in complex chemical systems, from base pair stacking in DNA and protein folding to crystal packing in organic solids. Complex chemical and biophysical systems, such as enzymes and proteins, are too large to be studied using the computational techniques rigorous enough to capture the subtleties of noncovalent interactions. Thus, the larger chemical system must be truncated to a smaller model system to which the rigorous methods can be applied in order to capture the essential physics of the interaction. Computational methodologies which can account for high levels of electron correlation, such as second-order perturbation theory and coupled-cluster theory, must be used. These computational techniques are used to study several types (π stacking, S/π, and C-H/π) of noncovalent interactions in two chemical contexts: biophysical systems and organic solids.

The effect of substituent effects on sandwich and T-shaped configurations of substituted benzene dimers are studied by second-order perturbation theory to determine how substituents tune π-π interactions. Remarkably, multiple substituents have an additive effect on the binding energy of sandwich dimers except in some cases when substituents are aligned on top of each other. T-shaped configurations are more complex, but nevertheless a simple model that accounts for electrostatic and dispersion interactions (and direct contacts between substituents on one ring and hydrogens on the other), provides a good match to the quantum mechanical results. The additivity of substituent effects in sandwich configurations also counters assertions that substituent effects are governed solely by electrostatic control, as the differential dispersion contributions accumulate with multiple substituents and give molecules with very different electrostatic potentials very similar interactions with

benzene.

The preferred interaction geometries of S/π interactions are evaluated through coupled-cluster computations for the H_2S-benzene complex. Geometries of cysteine/aromatic interactions found in crystal structures from the Brookhaven Protein Databank (PDB) are analyzed and compared to the equilibrium configurations predicted by high-level quantum mechanical results for the H_2S-benzene complex. A correlation is observed between the energetically favorable configurations on the quantum mechanical potential energy surface of the H_2S-benzene model and the cysteine/aromatic configurations most frequently found in crystal structures of the PDB. This result suggests that accurate quantum computations on models of noncovalent interactions may be helpful in understanding the structures of proteins and other complex systems.

Prototypical C-H/π interactions are examined by determining potential energy curves for methane-benzene, methane-phenol, and methane-indole complexes as prototypes for interactions between C-H bonds and the aromatic components of phenylalanine, tyrosine, and tryptophan. Second-order perturbation theory (MP2) is used in conjunction with the aug-cc-pVDZ and aug-cc-pVTZ basis sets to determine the counterpoise-corrected interaction energy for selected complex configurations. Using corrections for higher-order electron correlation determined with coupled-cluster theory through perturbative triples [CCSD(T)] in the aug-cc-pVDZ basis set, results are estimated, through an additive approximation, at the very accurate CCSD(T)/aug-cc-pVTZ level of theory. The fundamental C-H/π interaction is relatively insensitive to the type of aromatic ring involved in the interaction and thus, a general C-H/π interaction can be modeled as a five- or six-membered aromatic ring.

Finally, π stacking as a fundamental stabilizing force of organic crystals is considered. A first-principles methodology to obtain converged results for the lattice energy of small, neutral organic crystals is developed. In particular, the lattice energy of crystalline benzene is computed using an additive system based on the individual interaction energies of benzene dimers. Enthalpy corrections are estimated so that the lattice energy can be directly compared to the experimentally determined sublimation energy. The best estimate of the sublimation energy is 49.4 kJ mol^{-1}, just over the typical experimentally reported values of

43-47 kJ mol^{-1}. These results underscore the necessity of using highly correlated electronic structure methods to determine thermodynamic properties within chemical accuracy. The first coordination sphere contributes about 90% of the total lattice energy, and the second coordination sphere contributes the remaining 10%.

In this work, I have capitalized on theory and computation to gain unique insight about noncovalent interactions in chemical systems in a way that bridges accurate computational methods used to characterize small systems to large scale chemical systems. This "small to big" methodology is the framework for bottom-up development of chemical systems and is of ever increasing importance in the development of molecular engineering. Creating this bridge is a vital step to understanding how molecular properties can be utilized to solve chemical problems.

CHAPTER I

INTRODUCTION AND COMPUTATIONAL METHODS

1.1 Noncovalent interactions in chemical systems

Noncovalent interactions are critical to understanding a host of energetic and structural properties in complex chemical systems, from base pair stacking in DNA and protein folding to crystal packing in organic solids. Nobel Laureate Jean-Marie Lehn highlights the importance of such interactions in his book *Supermolecular Chemistry* [63] saying, "Noncovalent interactions define the inter-component bond, the action and reaction, in brief, the behaviour of molecular individuals and populations... Molecular interactions form the basis of the highly specific recognition, reaction, transport, and regulation processes that occur in biology." Interactions such as π stacking, S/π, and CH/π interactions contribute to the energetic stability and structure proteins. Small molecule binding events, such as occur when drug molecules enter the binding pocket of proteins, often involve molecular recognition through noncovalent interactions. For example, noncovalent interactions, particularly π-stacking, stabilize the association of the drug Aricept (which treats symptoms of Alzheimer's disease) with its enzyme receptor [69]. Understanding these interactions and how they can be utilized to increase the affinity and specificity of binding events is a key to unlocking the possibility of rational drug design.

Understanding the noncovalent interactions that stabilize organic crystal structures is vital to understanding self-assembly phenomena and is foundational to crystal engineering. The cohesive energy of crystals provides a means to rank competing crystals structures and identify low energy products. A reliable methodology to describe how individual interactions between molecules in a crystal contribute to the overall stability of the crystal structure could pave the way for crystal structure predictions based on molecular information.

Thus, despite the overwhelming importance of noncovalent interactions, they are only

beginning to be well understood, in part because of the difficulty in determining and separating individual interactions within the complex chemical system. Questions remain about the strength, directionality, and preferred geometric conformations of noncovalent interactions, plus how such interactions can be modified by substituent effects. Computational and experimental techniques can be used to explore these questions, and several relevant methods will be described.

1.2 Experimental determination of noncovalent interactions

Supermolecular chemistry is "chemistry beyond the molecule" and examines systems of increased complexity which are organized through intermolecular binding interactions. Given the importance these noncovalent interactions, a variety of experimental techniques can be used to quantify these interactions. A variety of soft ionization mass spectroscopy techniques can be used to determine noncovalent binding interactions [28]. Mass spectroscopy techniques can be used to study many types of noncovalent interactions including those bound by electrostatic and dispersion interactions, such as the systems examined in this thesis.

Also of particular importance to determine the magnitude of π-π interactions are molecular torsion balance experiments [14] and chemical double mutant complex cycles [22]. In a molecular torsion balance experiment, a flexible molecule with two aromatic moieties is developed which can adopt a folded conformation, in which the two aromatic groups interact, and an unfolded conformation where the interaction is removed. The ratio between the folded and unfolded conformations is quantified, often by the integration of an NMR spectrum, and this ratio is used to calculate the strength of the arene-arene interaction. Different molecular balances are developed to study different configurations of aromatic interactions, and sometimes a rigid molecular framework must be used to restrict the interaction to a particular geometry.

Chemical double mutant cycles are a general thermodynamic cycle first proposed in 1984 by Fersht [15] as a way to measure cooperative interactions for binding events to enzymes. The cycle considers a single X-mutation to a protein and the same mutation to a

single Y-mutated protein and compares the free energy changes for both these mutations. If these free energy changes are not the same, then an interaction exists between residues X and Y. This interaction can be quantified by making two mutations, one to remove the primary interaction and one to quantify the effect removing the primary mutation had on secondary interactions in the protein. By subtracting free energy changes for any two parallel mutations, the interaction between the residues X and Y can be determined.

However, such techniques have some limitations. Generally, the experimental methods determine the ΔG of the total reaction, and equate this free energy change to the ΔE of the aromatic interaction. This equality is not precise, as there could be additional entropy changes which affect the overall ΔG of the reaction. Additionally, even if entropy effects are not significant, the measured interaction energy is always modulated by solvent effects [73], which can disproportionately affect one conformation over another. Additionally, different solvents may affect noncovalent interactions differently, making it difficult to compare results from different experiments. In some cases, differing solvation effects can lead to different conclusions about how substituents effects affect π stacking and can lead to conflicting conclusions about the trends in the interaction energies [21]. Solvent rearragement is also a significant factor in how entropy changes (ΔS's) contribute with enthalpy changes (the ΔH's) to determine the overall free energy change (ΔG) of a noncovalent interaction. This rearrangement is decidedly different for noncovelent complexes in a homogenous solvent versus noncovalent interactions in complex, constrained chemical environments such as the interior of a protein. In the latter case, the solvent rearragement will likely make a smaller contribution to the ΔS (compared to a homogeneous solvent) and the ΔH will be a better approximation of the ΔG.

1.3 Computational determination of noncovalent interactions

Computational techniques can alleviate some of these complications by enabling the direct computation of the interaction energy between two systems. Complex chemical and biophysical systems, such as enzymes and proteins, are too large to be studied using the

computational techniques rigorous enough to capture the subtleties of noncovalent interactions. Thus, the larger chemical system must be truncated to a smaller model system to which the rigorous methods can be applied in order to capture the essential physics of the interaction. However, when truncated system are used, environmental effects from the rest of the chemical system are lost. Thus, it is critical to validate the use of such models by comparing information gained from the model systems to information about the macroscopic chemical system.

Alternatively, large chemical systems can be studied with less rigorous computational techniques such as semi-empirical methods, force field methods, or density functional theory. However, such methods must be calibrated for the types of chemical systems to which they are to be applied, thus correlated electronic structure studies such as the present work also provide valuable benchmark data to calibrate such lower-cost techniques. Highly accurate representations of the electronic structure of atoms and molecules requires a quantum mechanical description of the particles by the Schrödinger equation. This fundamental equation, and the techniques to solve it, are described below.

1.4 Overview of theoretical methods

1.4.1 The Schrödinger equation

Chemical systems can be described by the nuclear geometry and electronic structure of the molecules in the systems. The electronic structure of these systems can not be described by classical mechanics because the electrons posses both wave and particle-like characteristics. Rather, the physical information about the electronic structure of a chemical system is contained in wavefunction describing the system, denoted Ψ. To compute the energy of the system, the energy operator, the Hamiltonian (denoted **H**), is applied to the wavefunction. The wavefunction is an eigenfunction of the Hamiltonian, and the eigenvalues generated are the energy of the system. This relationship is the known as the Schrödinger equation

$$H\Psi = E\Psi.$$

(Additional details and derivations can be found in any physical chemistry text book such as Reference 99.)

The Schrödinger equation can only be solved exactly for a small number of systems. For realistic chemical systems, various approximate methods are employed. Most such methods utilize the Born-Oppenheimer approximation, which separates the wavefunction into a nuclear wavefunction and an electronic wavefunction by assuming that the nuclear velocities are sufficiently slower than the electronic velocities that these motions are not correlated with each other. The primary difference in the methods is in the way in which electron correlation is treated. Several approximate methods for solving the Schrödinger equation are used in this thesis and each will be described briefly below.

1.4.2 Hartree-Fock theory

All the correlated electronic structure methods used in this work build upon a Hartree-Fock (HF) reference wavefunction. Hartree-Fock theory is an approximate method for solving the Schrödinger equation for atoms and molecules that is based on the fundamental approximation that each electron feels only the *average* field of all the other electrons; thus, there is no explicit electron correlation. The wavefunction in HF methods is represented as a single Slater determinant, written in terms of the occupied molecular orbitals (MOs) and is invariant to unitary transformations of the MOs.

$$\Phi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}$$

The total energy is not the sum of the energies of these MOs, but is found variationally by applying the Fock operator (which has terms describing the repulsion between electrons) to the wavefunction, and solving variationally for the lowest energy set of molecular orbitals. Additional information and detailed derivations can be found in Chapter 3 of Reference 55.

1.4.3 Perturbation theory

From the HF reference, a variety of other theoretical methods can be derived which give better approximation to the exact solution to the Schrödinger equation. Perturbation theory methods take a problem for which the solution is known and improve it by adding a

perturbation to the operator.

$$H = H_0 + \lambda H'$$

where H_0 is the Hamiltonian for the known solution and H′ is a perturbing operator, and λ varies from 0 to 1 and maps the Hamiltonian for the known solution into the Hamiltonian for the improved, unknown solution.

The perturbed Schrödinger equation is now

$$H\Psi = W\Psi$$

where the wavefunction (Ψ) and the energy (W) can be written as expansions in terms of the perturbation, λ.

$$W = \lambda^0 W_0 + \lambda^1 W_1 + \lambda^2 W_2 + \lambda^3 W_3 + \ldots$$

$$\Psi = \lambda^0 \Psi_0 + \lambda^1 \Psi_1 + \lambda^2 \Psi_2 + \lambda^3 \Psi_3 + \ldots$$

The most common implementation of perturbation theory, called Møller-Plesset (MP) theory [6], takes the sum over the Fock operators as the unperturbed Hamiltonian. The wavefunctions Ψ are written as sums of molecular orbitals Φ. The zeroth-order energy is then the sum of the energy of the MOs. This double counts the electron-electron repulsion for each pair of electrons, thus the perturbation is the exact electron repulsion operator minus twice the average electron repulsion operator.

$$H' = V_{ee} - 2\langle V_{ee} \rangle$$

The first-order energy correction is the average of the first-order perturbation operator over the zeroth-order wavefunction

$$W_1 = \langle \Phi_0 | H' | \Phi_0 \rangle.$$

The total energy through first-order is then

$$E(MP1) = W_0 + W_1 = \langle \Phi_0 | H_0 + H' | \Phi_0 \rangle = E_{HF}$$

which gives the same result as HF theory. Since the so-called MP1 energy *is* the HF energy, electron correlation begins with the second-order energy correction which can be determined

from matrix elements involving the perturbation operator, the HF reference, and the excited states. For an excited state where 2 electrons have been promoted from orbitals i and j (from the set of occupied orbitals) to virtual orbitals a and b, the second order energy correction is given by

$$W_2 = \sum_{i<j}^{occ} \sum_{a<b}^{vir} \frac{\langle \Phi_0 | H' | \Phi_{ij}^{ab} \rangle \langle \Phi_{ij}^{ab} | H' | \Phi_0 \rangle}{E_0 - E_{ij}^{ab}}$$

and the total MP2 energy is

$$E(MP2) = E_{HF} + W_2.$$

1.4.4 Coupled-Cluster theory

Coupled-cluster (CC) theory [85, 7, 26] improves on HF theory by using more than one determinant to represent the wavefunction. (This is also true of other types of electronic structure theory such as configuration interaction.) An exponential projection operator

$$e^T = 1 + T + \frac{1}{2}T^2 + \frac{1}{6}T^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} T^k$$

is applied to the HF reference wavefunction to generate excited Slater determinants. The cluster operator **T** is given by

$$T = T_1 + T_2 + T_3 + \cdots + T_N$$

where each cluster operator produces determinants with N excitations relative to the HF reference determinant. In practice, all the cluster operators up to T_N can not be used unless the system is very small. In this work, the CCSD [86, 98, 96] (where the cluster operator has been truncated at T_2 thus only excited determinants that can be generated from a single and double excitations from the references wavefunction are used) and CCSD(T) [88, 97] variants of coupled-cluster theory are used. CCSD(T) does not actually determined all the excited determinants which would results from triple excitations self-consistently (such an approach would be called CCSDT), but rather the contribution of the connected triples is found using perturbation theory and added to the CCSD energy. CCSD(T) is popularly called the "gold standard" in quantum chemistry and is frequently used to benchmark other computational strategies.

1.5 Organization of thesis

The thesis contains six chapters, including this introductory chapter and a concluding chapter discussing the context and significance of this work. The remaining four chapters each discuss specific noncovalent interactions in different types of chemical systems. These four chapters are adapted from several papers (listed below) previously published about this work.

- "The Effect of Multiple Substituents On Sandwich and T-Shaped π-π Interactions," Ashley L. Ringer, Mutasem O. Sinnokrot, and C. David Sherrill, *Chem. Eur. J.* **12**, 3821-3828 (2006)

- "Substituent Effects in Sandwich Configurations of Multiply-substituted Benzene Dimers are Not Governed by Electrostatic Control," Ashley L. Ringer and C. David Sherrill, submitted to *J. Am. Chem. Soc.*

- "Models of S/π Interactions in Protein Structures: Comparison of the H_2S-benzene Complex with PDB Data," Ashley L. Ringer, Anastasia Senenko, and C. David Sherrill, *Protein Sci.* **16**, 2216-2223 (2007)

- "Aliphatic C-H/π Interactions: Methane-Benzene, Methane-Phenol, and Methane-Indole Complexes," Ashley L. Ringer, Michelle S. Figgs, Mutasem O. Sinnokrot, and C. David Sherrill, *J. Phys. Chem. A* **110**, 10822-10828 (2006)

- "First Principles Computation of Lattice Energies of Organic Solids: The Benzene Crystal," Ashley L. Ringer and C. David Sherrill, *Chem. Eur. J.* **14**, 2542-2547 (2008)

CHAPTER II

THE EFFECT OF MULTIPLE SUBSTITUENTS ON SANDWICH AND T-SHAPED π-π INTERACTIONS

2.1 Introduction

Noncovalent π-π interactions are involved in a wide variety of chemical and biological processes [69], ranging from self-assembly of synthetic molecules [20] to drug intercalation into DNA [94]. However, these important interactions are weak and feature shallow potential energy landscapes. Substituents can significantly alter the energy landscape and provide a way to tune π-π interactions. An understanding of how substituents can be used to adjust π-π interactions could be helpful in crystal engineering and the design of supramolecular architectures.

A few experiments have probed the effect of substituents on π-π interactions using nuclear magnetic resonance (NMR) spectroscopy techniques. Cozzi, Siegel, and coworkers [23, 24, 25] have measured barriers to rotation in substituted 1,8-diarylnaphthalenes featuring a nearly face-to-face (sandwich) configuration. Other experiments by Rashkin and Waters [89], Hunter and coworkers [1, 16], and Wilcox and coworkers [81, 58] examined π-π interactions in other (parallel-displaced and T-shaped) configurations. Other studies have examined the structures of benzene-hexafluorobenzene dimers [122] or 1:1 crystals [129]. None of these experiments were performed in the gas phase, so characterizing the intrinsic binding energy is difficult due to the inevitable presence of secondary interactions and solvent effects [73, 91]. Unfortunately, these experiments do not agree about how substituents alter π-π interactions: some of them indicate that electrostatic effects are dominant [23, 24, 25, 1, 16], while others argue for dispersion effects [81, 58].

Approximately perpendicular and offset parallel configurations are frequently observed in the crystal structures of simple aromatic compounds [52, 27], and interacting sidechains in proteins exhibit both orientations [13, 52]. Tsuzuki and coworkers [113] also noted in

their examination of toluene dimers that unlike benzene dimer, toluene dimers favor stacked configuration over T-shaped configurations. Here sandwich and T-shaped configurations of substituted benzene dimers will be investigated.

Conventional wisdom about geometric and substituent effects in π-π interactions is currently based upon the Hunter-Sanders model [51], which argues that although dispersion effects are important to the total binding energy, changes due to geometry or substitution are governed by electrostatic forces. This simple model describes an aromatic ring as a positively charged σ framework and a negatively charged π cloud. For sandwich configurations of substituted benzenes, this model predicts that electron withdrawing substituents strengthen the interaction because they decrease the electrostatic repulsion between the negatively charged π clouds. The reverse effect is predicted for electron donating substituents. High-level theoretical studies of substituted benzene dimers [102, 104] demonstrate that *all* substituted sandwich benzene dimers have a stronger attraction than the unsubstituted benzene dimer, regardless of the electron-donating or electron-withdrawing nature of the substituent, in contradiction to the Hunter-Sanders rules. Geerlings and coworkers [70] find similar results in their theoretical study of the interaction between mono-substituted benzenes with pyrimidine and imidazole. These unconventional prediction that electron donating substituents *increase* binding in face-to-face π-π interactions has been confirmed in in a recent study by Mei and Wolf [68]. These workers have synthesized a new, highly congested 1,8-diacridylnapthalene system to serve as a more robust experimental model of face-to-face π-π interactions. They find that oxides of their parent system feature increased π-π interactions, in agreement with the theoretical predictions.

The binding energies of substituted sandwich and T-shaped benzene dimers were analyzed using symmetry-adapted perturbation theory (SAPT) [56, 128], which provides the electrostatic, dispersion, induction, and exchange-repulsion components of the interaction energy. This analysis showed that not only is dispersion more important than electrostatics in the overall binding, but it can also be more important in determining substituent effects [104]. This conclusion is supported by previous studies of substituent effects in solute-solvent interactions in nematic liquid crystals by Williams and Lemieux [130].

So far, only monosubstituted benzene dimers have been considered. Here this work is extended to explore the effect of multiple substituents on sandwich and T-shaped configurations. Experimental work on multiple fluorination of 1,8-diarylnaphthalenes by Cozzi, Siegel, and coworkers [25] suggests that substituent effects in π-π interactions may be *additive*: these workers measured the barrier to rotation of phenyl groups about the naphthyl-phenyl bond, which they argue is related to the strength of the π-π interaction between phenyl groups.[1] In their studies of mono- through trifluorinated phenyl rings, they found that each fluorine contributes about -0.5 kcal mol^{-1} to the barrier to rotation. This is a remarkable result and suggests that, if additivity holds more generally, it might be possible to predict the energy change in π-π interactions based simply on the number and type of substituents using tabulated substituent values and/or a very simple equation involving molecular quantities for the monomers. In recent theoretical work, Kim and coworkers [61] demonstrated additivity in a single example in which they substituted both aromatic rings in a T-shaped benzene dimer and found that change in total interaction energy was nearly equal to the sum of the changes caused by the individual substitutions. However, by considering only a single disubstituted dimer, this work did not address the question of additivity in a general fashion. Riley and Merz [92] demonstrated the need to carefully consider direct hydrogen-substituent interaction in their extensive study of fluorosubstituted dimers, in which they consider every possible substitution pattern through hexasubstitution for benzene-n-fluorobenzene dimers. In this work, a broader investigation of the additivity of substituent effects on π-π interactions is presented through consideration of sandwich and T-shaped dimers of benzenes which are up to hexasubstituted for five different substituents. Further, a mathematical model is developed to predict relative interaction energies for substituted dimers that is a function of parameters correlating to electrostatic and dispersion contributions of the substituents.

[1] While it is certainly true that substitution will affect the strength of the π-π interaction between the phenyl groups in their minimum energy conformation, the substituents may *also* affect the energy of the rotation barrier, leading to the possibility of a nontrivial relationship between the strength of the π-π interaction and the rotational barrier height.

2.2 Computational details

All computations were performed using second-order Møller-Plesset perturbation theory (MP2) in conjunction with Dunning's augmented polarized correlation-consistent basis set aug-cc-pVDZ [57]. The aug- prefix denotes that this basis set has an extra set of diffuse functions for each angular momentum appearing in the basis. This basis set was chosen because the low symmetry of the dimers in this study, ranging in size from 24 to 33 atoms, limited the level of theory that could be applied. Previous work [106] on the benzene dimer indicates that it is more important to include additional diffuse functions rather than use a triple-ζ quality basis set. Fortunately, previous study of the relative changes caused by substitution of the benzene dimer shows that the *change* in interaction energy due to the substituents can be accurately determined at this computational level [104], even though the *total* binding energies are not as reliable as those computed using coupled-cluster theory with large basis sets. Monomers (Ph-X_n where X = H, F, CH_3, OH, NH_2, and CN) were optimized at the MP2/aug-cc-pVDZ level of theory, and sandwich dimers were constructed by maintaining these monomer geometries and varying the distance between the monomers over the range 3.0 to 4.0 Å. For the T-shaped configurations, the monomers were aligned at a 90° angle as shown in Figure 1, and the distance between the centers of the rings was varied over the range 4.5 to 5.5 Å. The monomer separation was initially varied by 0.2 Å increments to give the general shape of the potential energy curve, and then the resolution of the curve was increased to 0.05 Å near the equilibrium point. When substituting the benzene ring, the symmetrical substitution patterns, illustrated in Figure 2, were used. Disubstituted systems were substituted in the para-1,4 positions, and trisubstitutions were in the 1,3,5 positions. Hexasubstituted systems were also considered in some cases.

In the sandwich configurations, the monomers were aligned at their centers such that the C-X bonds of the substituted-benzene were coplanar to the C-H bond of benzene. In this procedure, the geometric center of each ring was used for alignment.[2] This configuration

[2]For substituted monomers, the ring is slightly deformed from hexagonal. The geometric center may be obtained by computing the center of mass of a ring with the same geometry but with equal masses for each atom.

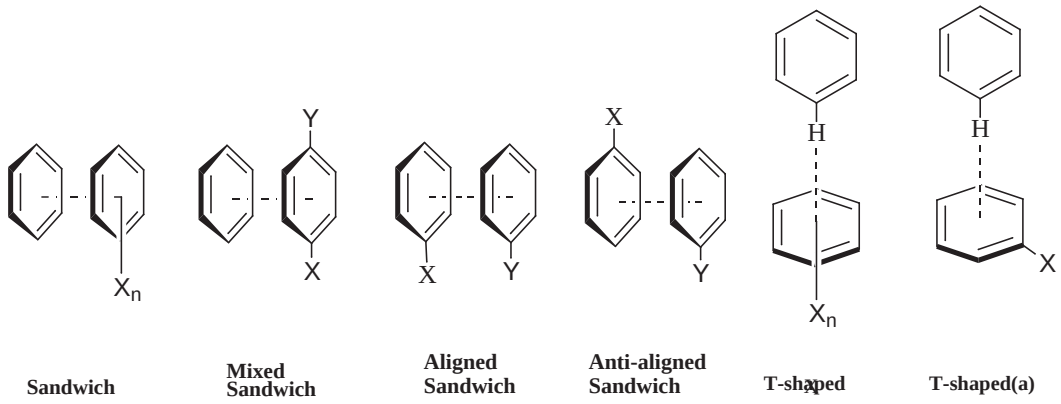
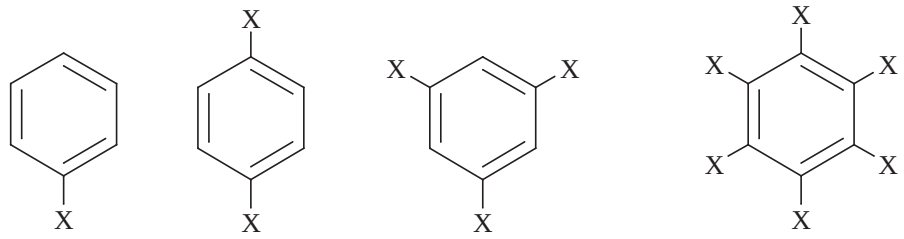


Figure 1: Dimer construction configurations for sandwich and T-shaped configurations.



X=H
X=OH
X=CH_3
X=F
X=CN
X=NH_2

Figure 2: Symmetric substitution patterns for substituted-dimers

was chosen as representative and is of course not the only possibility, but rotation of the sandwich dimer caused no more than 0.01 kcal mol^{-1} difference in the total interaction energy, even in the hexasubstituted dimers. Rotations of the lower ring in the T-shaped configurations is dicussed below. The CH_3 substituents had nearly free rotation around the C-C single bond, so the C_s configuration with one H up and two H down was chosen as representative. For the amino-substituted systems, the optimized configuration in which the hydrogens are directed away from the benzene ring was chosen.

Most of the dimers in this study are heterodimers between a benzene and a substituted benzene where the substituents are all of the same type. However, several "mixed" sandwich dimers with two different types of substituents were also considered; these are depicted in Figure 1. These dimers allowed the evaluation of, among other factors, the possible importance of direct interactions between substituents on different rings. Mixed sandwiches of benzene and para-disubstituted benzene and also dimers of two different monosubstituted benzenes were considered. In the latter case, the substituents were allowed to be aligned on top of each other or to be opposite each other in an "anti-aligned" configuration (see Figure 1).

Previous work on the benzene dimer [106, 104] demonstrates that interaction energies converge more rapidly when the Boys-Bernardi counterpoise correction [10] is employed (although this is not necessarily the case for all weakly bound systems); hence, the counterpoise correction is applied to all results reported here. Optimizations of monomer geometries were performed using Q-Chem 2.1 [59], and dimer computations were performed using MOLPRO [126].

Symmetry-adapted perturbation theory (SAPT) [56, 128] was applied using the program package SAPT2002 [12] to selected dimers to analyze their total interaction energies in terms of electrostatic, induction, dispersion, and exchange energies. The total interaction energy can be represented by the sum

$$E_{int} = E_{int}^{HF} + E_{int}^{CORR}$$

where E_{int}^{HF} describes the interactions at the Hartree-Fock level. This term can be further

expanded to yield

$$E_{int}^{HF} = E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + \delta E_{int,resp}^{HF}.$$

The superscripts (ab) indicate the order of the perturbation with respect to the intermolecular and intramonomer parts of the Hamiltonian, respectively. The subscript "resp" indicates that the term contains contributions from the coupled-perturbed Hartree-Fock response.

In the SAPT2 method employed here, the contribution of electron correlation to the interaction energy is nearly equivalent to that from a supermolecular MP2 computation and can be represented as

$$E_{int}^{CORR} = E_{elst,resp}^{(12)} + E_{exch}^{(11)} + E_{exch}^{(12)} + {}^{t}E_{ind}^{(22)} + {}^{t}E_{exch-ind}^{22} + E_{disp}^{(20)} + E_{exch-disp}^{(20)}.$$

where ${}^{t}E_{ind}^{(22)}$ represents the part of $E_{ind}^{(22)}$ that is not included in $E_{ind,resp}^{(20)}$. To simplify our discussion of the SAPT results, the exchange-induction and exchange-dispersion cross terms will be considered as induction and dispersion contributions, respectively. Additionally, the $\delta E_{int,resp}^{HF}$ term, which includes the third- and higher-order induction and exchange-induction contributions, is counted as induction. To make the SAPT computations feasible, a less expensive basis set was used, denoted cc-pVDZ+, which is the cc-pVDZ basis for hydrogen and an aug-cc-pVDZ basis minus diffuse d functions for all other atoms.

2.3 Results and discussion

2.3.1 Sandwich dimers

First, sandwich heterodimers consisting of one benzene and one substituted benzene (leftmost dimer of Figure 1) are considered. The optimum intermonomer distances are presented in Table 1 along with the change in the interaction energy (relative to the benzene dimer) due to substitution. As seen in previous work [102, 104] all substituted sandwich dimers have a greater interaction energy than the sandwich benzene dimer, regardless of the electron-donating or electron-withdrawing nature of the substituent. It is remarkable that the energy lowering due to two substituents is very nearly twice the energy lowering due to one substituent in all cases; i.e., the substituent effects are nearly additive for these sandwich heterodimers. Moreover, this additivity persists up through hexasubstituted dimers.

Table 1: Optimum intermonomer distances (in Å) and changes in the interaction energy (in kcal mol^{-1}, relative to benzene dimer) due to n substituents for sandwich heterodimers of benzene with multiply-substituted benzenes.

	n=1		n=2		n=3		n=6	
	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$
H	3.80	0.00						
OH	3.70	-0.49	3.65	-1.05	3.60	-1.50		
CH_3	3.70	-0.70	3.65	-1.23	3.60	-1.98		
F	3.70	-0.60	3.65	-1.24	3.60	-1.89	3.45	-4.29
CN	3.65	-1.58	3.60	-3.28	3.55	-4.82	3.40	-10.46
NH_2	3.65	-0.64	3.60	-1.39	3.50	-2.20		

[a] Equilibrium monomer separation (using rigid monomers). [b] All data computed at MP2/aug-cc-pVDZ level of theory; interaction energy of benzene dimer at this level is -2.90 kcal mol^{-1}.

This result is illustrated more clearly by Figures 3 and 4, which show the total interaction energy versus the number of substituents. The average change in the interaction energy per substituent can be determined from the slope of the best fit line for each functional group (-OH, 0.50; -CH_3, 0.66; -F, 0.64; -CN, 1.61; -NH_2, 0.69 kcal mol^{-1}). These values are in good agreement with the value simply determined from the monosubstituted system by subtracting the total interaction energy of benzene dimer from the interaction energy of the monosubstituted dimer (see Table 1). This indicates that interaction energies of these heterodimers might be accurately estimated using only information from the monosubstituted dimers. The results for multiple fluorination are of particular interest because they relate to the NMR experiments on multiply-fluorinated, biarylnaphthalenes by Cozzi, Siegel, and coworkers [25]. Those experiments indicated that the barrier to rotation about the arylnaphthyl bond was increased by 0.5 kcal mol^{-1} for each fluorine substituent (presumably due to increased π-π interactions between the two aryl groups). A near-linearity in the energies for multiple fluorinations is also found in this work, with the π-π interaction increasing by 0.6 kcal mol^{-1} per fluorine, in excellent agreement with the experimental findings.

Like the changes in the energies, the optimum geometries also show a systematic pattern with respect to the number of substituents. For monosubstituted dimers, the optimized distance between the rings ranges from 3.80 (benzene dimer) to 3.65 Å (benzene-benzonitrile and benzene-aniline). However, in nearly all cases, each additional substituent, regardless

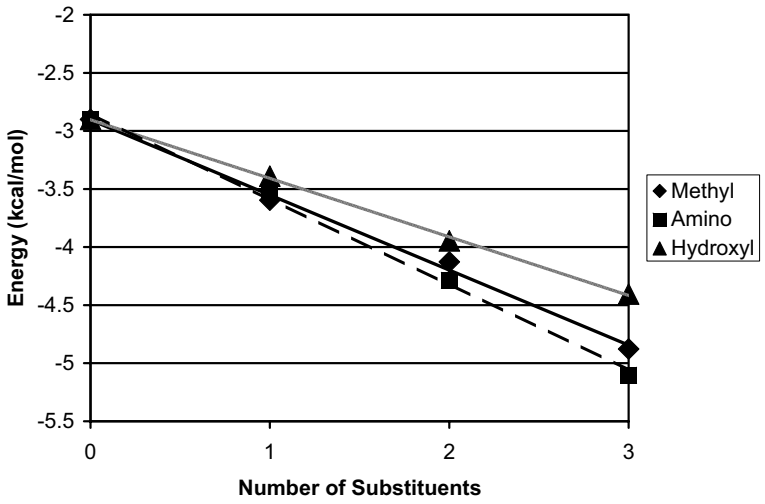


Figure 3: Total interaction energy vs. number of substituents (through trisubstitution) for sandwich configurations.

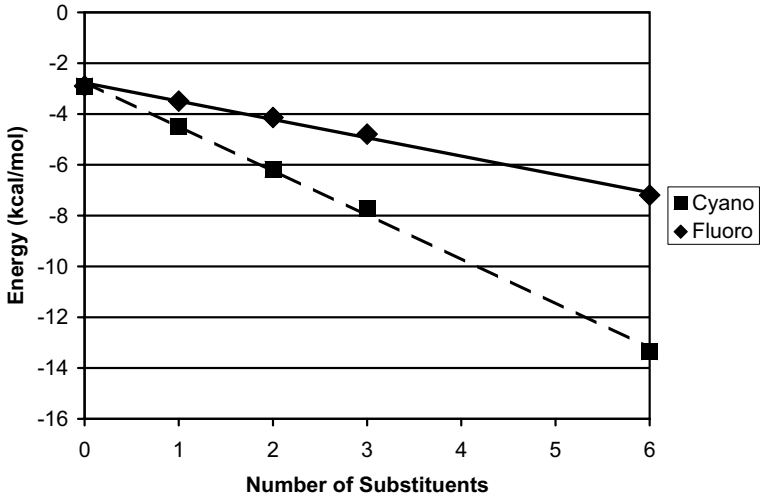


Figure 4: Total interaction energy vs. number of substituents (through hexasubstitution) for sandwich configurations.

of type, decreases the equilibrium distance between the rings by 0.05 Å (note that this is the resolution used in determining the potential curves); for example, the equilibrium distance in benzene-hexacyanobenzene is 0.25 Å less than that in benzene-benzonitrile, which has five fewer CN substituents.

Table 2 presents the SAPT results for the benzene dimer and several fluorinated dimers. In agreement with the MP2/aug-cc-pVDZ supermolecule computations, the SAPT2/cc-pVDZ+ results show that one fluorine in the sandwich fluorobenzene-benzene dimer stabilizes the complex by about 0.6 kcal mol^{-1} relative to the benzene dimer sandwich, and two fluorines in 1,4-difluorobenzene-benzene dimer lead to almost twice this stabilization. One might suppose that this doubling of the stabilization might be reflected in each of the SAPT energy components, but this is not the case. For example, considering the electrostatic stabilization of substitution relative to the sandwich benzene dimer is -1.145 kcal mol^{-1} for the 1,4-difluorobenzene-benzene dimer, which is significantly more than twice the stabilization of -0.395 kcal mol^{-1} found for the fluorobenzene-benzene dimer. On the other hand, the change in the induction term relative to benzene dimer is almost the same for both fluorinated dimers. Both the exchange-repulsion and dispersion terms are much larger in magnitude for the 1,4-difluorobenzene-benzene sandwich because its shorter intermonomer distance leads to greater overlap between the π clouds.

2.3.2 Substituent effects in sandwich configurations of multiply-substituted benzene dimers are not solely governed by electrostatic control

As the SAPT analysis indicates, dispersion is critically important to the overall stabilization of the substituted dimers. However, several other works which examined substituent effects using correlated electronic structure techniques or density functional theory [61, 62, 2, 127] came to the conculsion that while dispersion interactions are important in particular cases and often contribute significantly to the overall stability of noncovalent complexes, the *trend* of substituent effects can be related to simple electrostatic parameters of the substituents. A particularly extensive study of this type was conducted by Wheeler and Houk [127], who considered 25 different monosubstituted sandwich benzene dimers by using the computationally economical density functional M05-2X [135]. When the unsubstituted case (benzene

Table 2: Physical components (in kcal mol^{-1}) of total interaction energy determined using SAPT for benzene and substituted fluorobenzene dimers.

Configuration[a]	R	Elst.	Exch.	Ind.	Disp.	SAPT2[b]
Benzene-Benzene(S)	3.70	-0.974	6.034	-0.331	-6.528	-1.799
Fluorobenzene-Benzene(S)	3.70	-1.369	5.890	-0.305	-6.630	-2.414
Difluorobenzene-Benzene(S)	3.65	-2.119	6.425	-0.311	-7.012	-3.017
Fluorobenzene-Fluorobenzene(S aligned)	3.70	-1.066	5.582	-0.237	-6.538	-2.259
Fluorobenzene-Fluorobenzene(S anti)	3.65	-2.068	6.412	-0.285	-7.013	-2.954
Benzene-Benzene(T)	4.90	-2.244	4.865	-0.670	-4.367	-2.416
Fluorobenzene-Benzene(T)	5.00	-1.639	3.777	-0.487	-3.876	-2.225
Fluorobenzene-Benzene(T(a))[c]	5.00	-1.748	3.778	-0.483	-3.867	-2.320
Difluorobenzene-Benzene(T)	5.00	-1.368	3.706	-0.420	-3.834	-1.916

[a] S = sandwich configuration; T = T-shaped configuration [b] All data computed using cc-pVDZ+ using optimized MP2/aug-cc-pVDZ monomer geometries with optimum intermonomer separations. [c] Configuration depicted by rightmost dimer in Figure 1

dimer) was not included, the relative interaction energies showed a reasonably good linear correlation with the Hammett parameter σ_m for each substitutent, which represents that substituent's electron donating or electron withdrawing character. However, the M05-2X results also corroborated previous findings [102] that all substituents increase the interaction energy relative to benzene dimer. Wheeler and Houk suggest that this is due to a relatively constant dispersion stabilization for all substituents considered. Because most substituents will lead to larger dispersion interactions than hydrogen, this has the effect of shifting the relative interaction energies down (becoming more stabilizing), so that the linear fit line determined (ΔE_{int} = 2.71 σ_mx - 0.57) does not cross through the origin, but has a negative intercept. However, when the effect of dispersion was explicitly subtracted using previously published results for four of the substituted dimers [104], the linear model now nicely fit not only these points, but also that for the parent benzene dimer (see Table S1 in the supplmental material of Reference 127).

Although the dispersion contributions complicate the picture somewhat, the linear correlation with Hammett parameters led Wheeler and Houk to conclude that, "the trend in substituent effects can be qualitatively understood in terms of the electron-donating or withdrawing character of the substituents." This is certainly true for the data presented

in that work, but by evaulating the data for the multiply-substituted cases presented in this work, it shown that this is not true for π-π interactions in general. Instead, differential dispersion effects can be so large that even molecules with wildly different electrostatic potentials can exhibit similar attractions to benzene.

A similar analysis as in the work of Wheeler and Houk is performed for multiply-substituted sandwich benzene dimers. The counterpoise-corrected MP2/aug-cc-pVDZ interaction energies (relative to bezene dimer) for the mono-substituted, 1,3,5-tri-substituted and hexa-substituted benzene complexes for six different substitutents (CH_3, F, OH, NH_2, CH_2OH, and CN; the hexa-substituted cases for OH and CH_2OH were not included) versus the sum of the Hammett parameters ($\Sigma\sigma_m$) for all the substituents is shown in Figure 5. Previous work [108] has shown that an additivity rule is applicable when using Hammett parameters to capture inductive effects for multiple substituents in quinuclidine and bicyclo[2.2.2]octane carboxylic derivatives, and a summation of Hammett parameters has been used to represent electrostatic character of multiply-substituted complexes in other work which calculated stacking interactions for substituted sandwich complexes [9]. A linear correlation is not observed, and Figure 5 is striking evidence that substituent effects in face-to-face π-π interactions are not governed solely by electrostatic control.

To further examine the relationship between the electrostatic nature of the substituted systems and the interaction energy, Hartree-Fock/6-31G* electrostatic potential maps were computed for three of the complexes with similar relative interaction energies (Figure 6). The hexa-substituted NH_2 complex, with six highly electron donating groups, still has an interaction energy that is 5.1 kcal mol^{-1} more bound than benzene dimer. Such a result is impossible to explain on the basis of the Hunter-Sanders rules, which posit that electron-donating substituents increase the negative charge in the π-electron cloud and thus lead to less favorable electrostatic interactions with an unsubstituted benzene. The electrostatic potential map (Figure 6) confirms an electron-rich π cloud for the hexa-amino-substituted complex, whereas the tri-cyano-substituted (4.8 kcal mol^{-1} more bound than benzene dimer) and the hexa-fluoro-substituted complex (4.3 kcal mol^{-1} more bound than

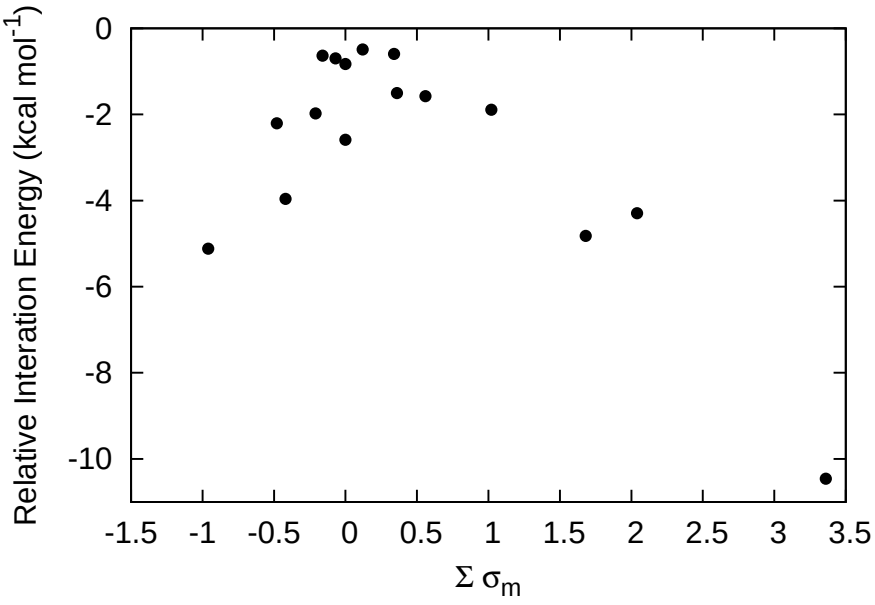


Figure 5: Interaction energies (relative to benzene dimer) vs. $\Sigma\sigma_m$ parameters for substituted face-to-face benzene dimers.

benzene dimer) have similar interaction energies with benzene but noticeably depleted electron density in the center of the substituted rings.

The notable stabilization of the hexa-amino substituted complex demonstrates the significant effect that the differential dispersion effects can have on the overall stability of the substituted complex. Changes in the dispersion energy due to substituents in monosubstituted sandwich benzene dimers, while relatively small, are not roughly constant nor even always stabilizing. The relative dispersion contributions (in kcal mol^{-1}) for the monosubstituted cases are -0.66 (CH_3), 0.039 (F), -0.192 (OH), -0.482 (CN) at the SAPT/aug-cc-pVDZ' level of theory [104]. As demonstrated above, the relative interaction energy is additive for increasing numbers of substitutions in sandwich configurations; thus, differences in the dispersion contributions of various substituents would become magnified for multiply substituted dimers and correlations with electrostatic parameters will be erased for multiply-substituted dimers unless the dispersion contribution is explicitly accounted for also.

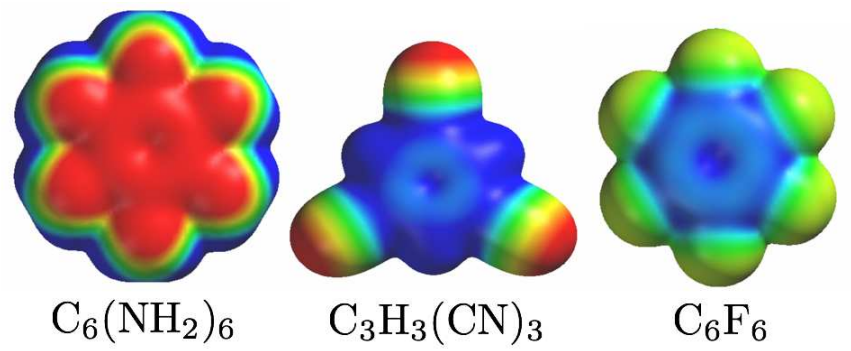


Figure 6: Hartree-Fock/6-31G* electrostatic potential maps [-25 kcal mol^{-1} (red) to +25 kcal mol^{-1} (blue)] of the hexa-amino-substituted, 1,3,5-tri-cyano-substituted, and hexa-fluoro-substituted benzene. All three have similar (within 1 kcal mol-1) interaction energies with benzene.

2.3.3 Sandwich configurations: Mixed substituent cases

Thus far, only dimers in which one ring has been substituted and which feature only one type of substituent have been considered. Let us now consider mixed sandwich dimers with two different types of substituents (Figure 1) and/or substituents on both rings. Table 3 presents equilibrium intermonomer distances and changes in the interaction energy due to substitution for five mixed sandwiches. The table also includes the change in the interaction energy which would be predicted by adding the average energy lowering of each substituent derived from the slopes of the graphs in Figures 3 and 4. For the dimers of benzene with para-disubstituted benzene, the predicted energy lowering is very close to that which is explicitly computed (within 0.1 kcal mol^{-1}). However, when the substituents are placed on two different monomers, significant deviations from the predicted values appear for the aligned CN/F, CN/CN, F/F, and NH_2/F cases. Note that the strongest deviations from the ideal values are observed when both substituents are strongly electron donating or strongly electron withdrawing; mixed cases involving methyl substituents follow the ideal behavior.

To determine the cause of this deviation from from the predicted additivity, SAPT analysis is used to obtain the physical components of the total interaction energy. The aligned fluorobenzene-benzene dimer was taken as representative of a non-additive case, and SAPT results for the aligned and anti-aligned configurations of this dimer are compared to the

Table 3: Optimum intermonomer distances (in Å) and changes in interaction energies (in kcal mol^{-1}, relative to benzene dimer) for mixed-substituent sandwich heterodimers.

	Predicted[a]	1,4-substitution		Aligned		Anti-aligned	
	$\Delta\Delta E_{int}$[b]	R[c]	$\Delta\Delta E_{int}$[b]	R[c]	$\Delta\Delta E_{int}$[b]	R[c]	$\Delta\Delta E_{int}$[b]
NH_2 and CH_3	-1.35	3.65	-1.33	3.75	-1.30	3.65	-1.32
CN and CH_3	-2.30	3.65	-2.25	3.75	-2.23	3.65	-2.20
CN and F	-2.28	3.60	-2.25	3.65	-0.98	3.60	-2.10
CN and CN	-3.28	3.60	-3.28	3.70	-0.75	3.60	-2.89
NH_2 and F	-1.33	3.60	-1.26	3.70	-0.52	3.60	-1.34
F and F	-1.28	3.65	-1.24	3.70	-0.49	3.65	-1.17

[a] Determined by adding the average change in interaction energies for each substituent as determined from Figures 3 and 4. [b] All data computed at MP2/aug-cc-pVDZ level of theory; interaction energy of benzene dimer at this level is -2.90 kcal mol^{-1}. [c] Equilibrium monomer separation (using rigid monomers).

1,4-difluorobenzene-benzene dimer in Table 2. Comparing the three cases, all components of the 1,4-substituted and the anti-aligned dimers are almost identical, thus they have nearly the same total interaction energy. However, for the aligned dimer, the electrostatic contribution is less stabilizing than the 1,4-substituted or the anti-aligned dimer by approximately one kcal mol^{-1}, despite the fact that two fluorines in any configuration should withdraw electron density from the π cloud in about the same way. However, the electrostatic potential maps presented in Reference 104 showed (not surprisingly) a concentration of negative charge on the fluorine of the fluorobenzene monomer. In the aligned dimers, this fluorine/fluorine direct interaction would have a much less favorable electrostatic contribution than a fluorine/hydrogen interaction that would be found in the 1,4-substituted or anti-aligned dimers, and this destabilization accounts for the differing electrostatic contributions. Partially compensating for this electrostatic destabilization is the significant reduction in the exchange-repulsion term (0.8 kcal mol^{-1}) due to the greater intermonomer separation in the aligned dimer. However, the greater distance also leads to a significant decrease (0.5 kcal mol^{-1}) in the dispersion stabilization, so that the aligned case is about 0.7 kcal mol^{-1} destabilized relative to the anti-aligned case. All of the aligned sandwich dimers have intermonomer distances which are at least 0.05 Å greater than those of the corresponding anti-aligned dimers.

2.3.4 T-shaped dimers

As was discussed in previous work [104], the effect of substituents on the binding energies of T-shaped dimers might be thought of, to a first approximation, in terms of the favorable electrostatic interaction between the negatively charged π cloud of the lower ring and the positively charged hydrogen of the other ring above it. One might then expect substituents on the lower ring to strengthen or weaken this interaction depending on how they tune the negative charge of the π cloud. Previous analysis [104] shows that this picture is somewhat oversimplified. First, the nominally electron-donating substituent -OH does not lead to any significant change in binding (although this is consistent with the electrostatic potential of phenol, which is very similar to that of benzene in the middle of the ring); second, -CH_3 substitution leads to significantly increased binding due to changes in the dispersion term, not the electrostatic term.

Theoretical results for multiply-substituted T-shaped dimers are summarized in Table 4 and Figures 7 and 8. Unlike the corresponding figures for the sandwich configurations, the energy shows significant nonlinearity as the number of substituents, n, increases from 0 to 6. However, the plots in Figures 7 and 8 are nearly linear through disubstitution (n=0 to n=2), suggesting that a new effect becomes operative for dimers with three or more substituents. In the T-shaped configuration, there is a possibility for direct interactions between the functional groups of the substituted benzene rings and the hydrogens of the upper benzene ring which would cause deviations from additivity. Such interactions would not be present in the mono- and disubstituted dimer configurations considered, but two interactions would be present in the trisubstituted dimers and four such interactions for hexasubstituted dimers (see Figures 2 and 1). This type of direct substituent interaction would cause an electrostatic stabilization compared to an otherwise identical dimer whose geometry did not provide such an interaction. SAPT analysis comparing the T-shaped and T-shaped(a) configurations (Figure 1) of the fluorobenzene-benzene dimer is presented in Table 2. The exchange, induction, and dispersion contributions to the total interaction energy are the same for both configurations, but the electrostatic contribution is stabilized by approximately 0.1 kcal mol^{-1}, which is consistent with a direct interaction between a

Table 4: Optimum intermonomer distances (in Å) and changes in interaction energies (in kcal mol^{-1}, relative to benzene dimer) for T-shaped heterodimers of benzene with multiply-substituted benzenes.

	n=1		n=2		n=3		n=6	
	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$	R^a	$\Delta\Delta E_{int}{}^b$
H	5.00	0.00						
OH	4.95	-0.02	4.95	-0.09	4.95	-0.22		
CH_3	4.90	-0.39	4.90	-0.72	4.85	-0.99		
F	5.00	0.33	5.00	0.56	5.00	0.64	5.00	0.90
CN	4.95	0.39	4.95	0.57	4.95	0.32	5.00	-0.68
NH_2	4.95	-0.16	4.90	-0.22	4.90	-0.90		

[a] Equilibrium monomer separation (using rigid monomers). [b] All data computed at MP2/aug-cc-pVDZ level of theory; interaction energy of benzene dimer at this level is -3.16 kcal mol^{-1}.

partially positive hydrogen and a partially negative fluorine.

SAPT energy analysis also reveals important differences in the ways that substituents affect different dimer configurations. Comparing results for the T-shaped fluorobenzene-benzene and 1,4-difluorobenzene-benzene dimers from Table 2, the only component which changes significantly with the addition of the second fluorine is the electrostatic contribution, whose almost 0.30 kcal mol^{-1} destabilization accounts for essentially the entire difference in the total interaction energy. Interestingly, the exchange-repulsion contribution, which changes by about 0.6 kcal mol^{-1} with the addition of a second fluorine in sandwich configuration dimers, is now largely unchanged by the second fluorination in a T-shaped configuration. This difference can be attributed to the sandwich configurations of these dimers having different intermonomer separations whereas the T-shaped configurations do not.

Because the T-shaped dimers do not exhibit full additivity through hexasubstitution, a simple extrapolation of interaction energies from monosubstituted dimers will not capture the correct trend as it did for the sandwich dimers. One factor which must be accounted for is the number number of direct interactions between substituents on one ring and the hydrogens of the other ring, as discussed above (in the sandwiches we considered, this direct interaction is *always* present). In previous work [104], a linear model was used to fit interaction energies of mono-substituted T-shaped benzene dimers to the Hammett

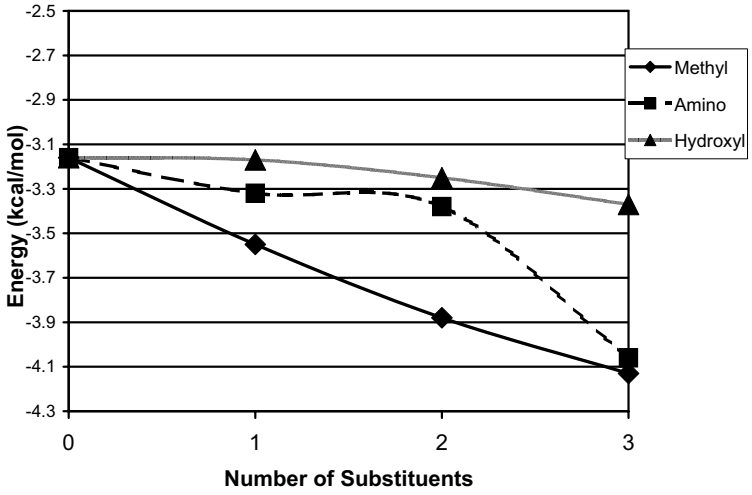


Figure 7: Total interaction energy vs. number of substituents (through trisubstitution) for T-shaped configurations.

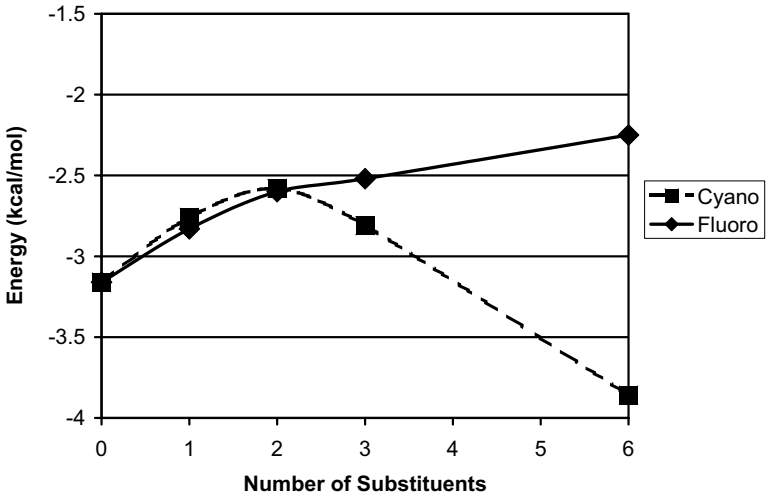


Figure 8: Total interaction energy vs. number of substituents (through hexasubstitution) for T-shaped configurations.

constants of the substituents, but only a rough correlation with σ_m was found. Because SAPT analysis showed that the two components of interaction energy most relevant in determining changes caused by substituents are dispersion and electrostatic energies, in this work a multi-linear model that uses parameters corresponding to both these interactions is developed. Williams and Lemieux [130] advanced a similar idea in a study in which they measured the shift in clearing point caused by dopant-host interactions in nematic liquids. Taking this shift as a measure of the interaction, they used a multi-linear model to describe this clearing point shift as a function of the HOMO energy for the dopant molecule and the calculated molecular polarizability. The model in this work predicts the strength of the π-π interaction directly by fitting to the Hammett parameters to describe the electrostatic character of the substituent [3] and experimentally determined molecular polarizabilities to account for the dispersive interaction.

The interaction energies (relative to benzene dimer) of the substituted, T-shaped Ph-X_n/benzene dimers (with substituents on the lower ring) were fit to a linear combination of these parameters having the form:

$$\Delta\Delta E_{int} = a\sum \sigma_m + b\Delta\alpha + d\delta.$$

$\sum \sigma_m$ is the sum of the Hammett parameters for all substituents, $\Delta\alpha$ is the change in the experimentally determined scalar molecular polarizability (in 10^{-24} cm^3) relative to benzene, and δ is a parameter to account for the direct interactions between substituents of one ring and hydrogens of the other, as described above. The experimental scalar polarizability values[4] were obtained from reference 65. As was found in previous work, a better fit is found using σ_m rather than σ_p values. To determine the value of the δ parameter, the total interaction energy is determined for another series of monosubstituted dimers in which the functional group of the substituted ring is placed closer to the interacting hydrogens of the other ring (see rightmost dimer of Figure 1), but the rest of the geometry, including

[3] As was pointed pointed out previously in Reference 104, the Hammett parameter is not always a good measure of the molecular electrostatic potential, or therefore electrostatic contribution to the interaction energy. Nevertheless, it is useful as a simple and readily available parameter.

[4] The scalar polarizabilities taken from Ref. 65 were: 10.9 (benzene), 10.3 (fluorobenzene), 9.8 (1,4-difluorobenzene), 9.74 (1,3,5-trifluorobenzene), 9.58 (hexafluorobenzene), 12.5 (benzonitrile), 19.2 (1,4-dicyanobenzene), 12.3 (toluene), 14.9 (1,4-dimethyltoluene), 11.1 (phenol), and 12.1 $\times$ 10^{-24} cm^3 (aniline).

Table 5: Interaction energies used to determine the direct interaction parameter, δ.

	T-shaped[a]	T-shaped(a)[b]	δ[c]
OH	-3.17	-3.10	0.07
CH_3	-3.55	-3.47	0.08
F	-2.83	-2.93	-0.10
CN	-2.76	-2.95	-0.19
NH_2	-3.32	-3.48	-0.16

[a] From Table 4. [b] Interaction energy of configuration shown in Figure 1 using intermonomer separations of T-shaped configuration from Table 4. [c] Determined by subtracting the interaction energies for the two configurations.

the intermonomer separation, is kept constant. The difference in the interaction energies of this configuration and the original T-shaped configuration is taken as value of a direct interaction (δ) and shown in Table 5. The coefficients a and b were determined by fitting to the MP2/aug-cc-pVDZ $\Delta\Delta E_{int}$ values for all substituted T-shaped dimers for which experimental monomer polarizabilities were available. This yielded values of $a = 0.708$ kcal mol^{-1} and b = -0.052 kcal mol^{-1} 10^{24} cm^{-3}.

Figure 9 compares the predictions of the model to the explicitly-computed MP2/aug-cc-pVDZ results. An R^2 of 0.83 is obtained for the line $y = x$, which would indicate a perfect coincidence of the $\Delta\Delta E_{int}$ values predicted by the model and as computed by the MP2 method. This value is rather similar to the R^2 of 0.81 obtained by Williams and Lemieux [130] in their fit of clearing point shifts due to substituents effects in π-π interactions in nematic liquids crystals. The largest discrepancy is for benzene-dimethylbenzene, where the model predicts a $\Delta\Delta E_{int}$ of -0.31 compared to a value of -0.72 kcal mol^{-1} computed at the MP2/aug-cc-pVDZ level of theory. Given the crudity of the model and its reliance on experimental polarizabilities which may be off by as much as 30% [65], the quality of the fit is quite good, and it may be useful in providing semiquantitative estimates of how substituents may tune the strength of T-shaped π-π interactions.

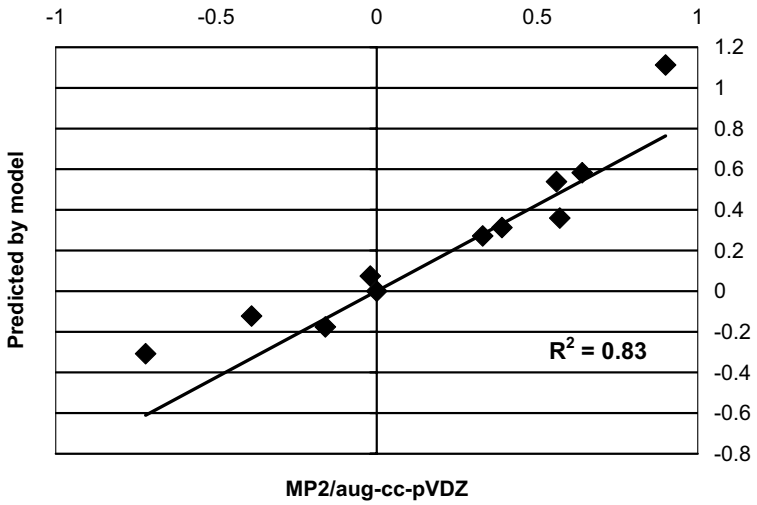


Figure 9: $\Delta\Delta E_{int}$ predicted by the simple model vs. $\Delta\Delta E_{int}$ explicity computed at the MP2/aug-cc-pVDZ level of theory.

2.4 Conclusions

In this chapter, the effect multiple substituents tuning π-π interactions has been explored. Such knowledge is foundational for rational drug design, crystal engineering, and supramolecular chemistry. Reliable *ab initio* quantum mechanical methods have been used to assess how substitution changes intermolecular geometries and binding energies in face-to-face (sandwich) and edge-to-face (T-shaped) configurations of substituted benzene dimers. Perhaps surprisingly, substituent effects are nearly additive in many sandwich configurations, allowing one to predict the results of any combination of substituents simply from the changes due to each substituent individually. An exception to this rule is the case in which substituents on different rings are aligned on top of each other, which can cause deviations from additivity. The situation for T-shaped configurations is somewhat more complex, in part because there is the additional complication of having to account for how many contacts a substituent on one benzene ring might make with hydrogens of the other ring.

Nevertheless, a simple model involving Hammett σ_m parameters and experimentally determined scalar polarizabilities provides a good fit to the *ab initio* data for the T-shaped configurations, once again suggesting that the effect of multiple substitution may be simply predicted. These results underscore the importance of accounting for direct interactions between an aromatic ring and substituents on another ring, as pointed out earlier in experimental studies of parallel-displaced interactions by Rashkin and Waters [89]. The data presented here should provide valuable guidance in how to tune π-π interactions.

CHAPTER III

MODELS OF S/π INTERACTIONS IN PROTEIN STRUCTURES

3.1 Introduction

The tertiary structure of proteins is determined by a variety of intermolecular interactions. Traditional hydrogen bonding is one critical noncovalent interaction which can play a large role in determining structure, but many other, weaker noncovalent interactions can also contribute. Understanding the underlying nature, strength, and directionality of these interactions is important for the prediction of the optimal structure of proteins and the dynamics of their folding. Unfortunately, isolating an individual interaction in a complex protein structure, and separating the effect of this interaction from that of other weak interactions and solvent effects, would be nearly impossible. Computational techniques offer a way to systematically and rigorously characterize the strength of various types of interactions by providing highly accurate potential energy curves for small model systems. For example, converged *ab initio* computations have deepened our understanding of π-π interactions through studies of the simplest possible prototype system, the benzene dimer [105, 106, 53, 49, 117, 116, 115, 119, 48].

Such an approach assumes that the model system accurately captures the essential physics of the non-bonded interaction as it would occur in larger systems. This study aims to address the validity of this assumption by providing highly accurate potential curves for several model configurations of the H_2S-benzene complex (see Figure 10) and comparing these results to the preferred geometries of cysteine/aromatic contacts observed in the Brookhaven Protein Databank (PDB).

Favorable interactions between sulfur and π aromatic systems were first suggested by Morgan et al. [71] when a series of alternating S and π bonded atoms were identified in several protein structures. Subsequent studies [134, 90] examining crystal structures from both the Protein Databank and the Cambridge Crystallographic Database have revealed

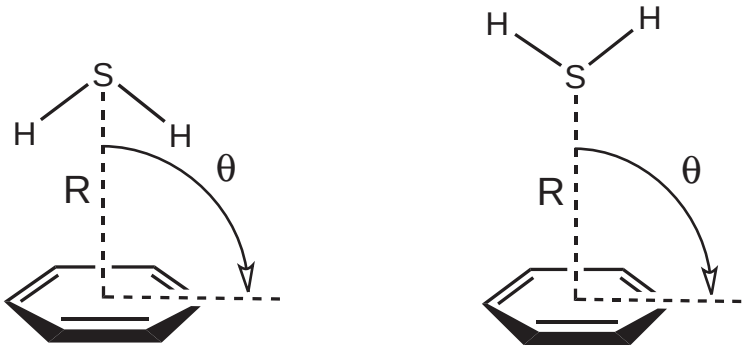


Figure 10: Variation of R and θ for C_{2v} configurations of the H_2S-benzene complex.

that sulfur-π interactions occur more commonly in protein crystal structures than would be expected from a random association of the structure. A few theoretical studies have also examined these interactions. Cheney et al. [17] investigated the methanethiol-benzene complex as a model of cysteine-aromatic interactions using Hartree-Fock (HF) and second-order Møller-Plesset perturbation theory (MP2) with several basis sets. After optimizing several initial configurations, they report that the optimum configuration has an inter-fragment separation (distance from sulfur to center of the benzene ring) of 4.4 Å and an angle between the sulfur and the plane of the benzene of 56°. More recent work by Duan et el. [31] also examined methanethiol-benzene using MP2 in conjunction with larger basis sets and found an optimum configuration in which the sulfur was directly above the benzene ring at a inter-fragment separation of 3.73 Å. For more information about S/π interactions, the reader is referred to the excellent review article by Meyer, Castellano, and Diederich [69].

Previous work with weak interactions [106] suggests that higher-order correlation techniques are required to converge the interaction energy of noncovalent complexes. Sherrill and coworkers [110] were the first to apply a highly correlated computational technique such as coupled-cluster theory through perturbative triples [CCSD(T)] with sufficiently large basis sets to achieve converged sub-chemical accuracy results for the H_2S-benzene complex as a prototype for sulfur-π interactions. Although the general term sulfur-π interactions is

used, both in the interactions of sulfur lone-pairs with the π systems, as well as the interactions of sulfur-bonded hydrogens pointed at π-systems (which could perhaps be referred to as S-H/π interactions) are of interest. [3]. The study by Tauer et al. found the inter-fragment separation for the equilibrium geometry of the hydrogens-down C_{2v} structure of H_2S-benzene was 3.8 Å, and the interaction energy of the complex at this geometry was -2.74 kcal mol^{-1} [CCSD(T)/aug-cc-pVQZ results]. In this work, the hydrogens-down C_{2v} structure (and also a hydrogens-up C_{2v} structure) are used as starting configurations and systematically vary both the inter-fragment distance (measured between the sulfur of H_2S and the geometric center of the benzene) and the angle between the sulfur and the perpendicular to the aromatic plane of benzene (see Figure 10). Based on the potential energy surfaces (in the R/θ space defined), interesting configurations are selected and analyzed using highly-correlated techniques similar to those used by Tauer et al. to determine potential curves for the selected configurations.

Seemingly at odds with the quantum mechanical results of Tauer et al. for the H_2S-benzene model, Reid et al. [90] examined 36 proteins from the Protein Databank and reported that sulfur atoms prefer to interact with the edge of aromatic rings and avoid the area in the center of the ring around the π-electrons. Zauhar et al. [134] compared probability distributions for the geometries of divalent sulfurs interacting with six-membered aromatic carbon rings with analogous probability distributions of X-CH_2-X groups interacting with aromatic rings for structures from the Crystallographic Database. From these results, they defined a preferred geometry of interaction in which the divalent sulfur is in plane with the aromatic ring and at a separation of around 5 Å.

In this work, the optimum configurations predicted by high-level quantum mechanics are directly compared with configurations which occur frequently in the PDB by performing an analysis of crystal structures from the Brookhaven Protein Databank, in which the same parameters are determined for each sulfur-π interaction in the crystal structure as were varied in the potential energy surfaces. This comparison should help us understand whether quantum mechanical calculations of small model systems can provide reliable predictions of geometric configurations for interactions found in crystal structures.

3.2 Computational details

3.2.1 Ab initio calculations

Monomer geometries for hydrogen sulfide and benzene were taken as the best values from the literature: r_e(C-C) = 1.3915 Å and r_e(C-H) = 1.0800 Å for benzene [37] and r_e(S-H) = 1.3356 Å and θ_e(H-S-H) = 92.12° for hydrogen sulfide [34]. From these monomer geometries, two initial configurations were constructed in which the sulfur of H_2S was placed directly over the center of the benzene ring: one structure with the hydrogens directed towards the ring, and one away from the ring (Figure 10). From these starting geometries the distance between the sulfur and the ring center (denoted R) was systematically varied in 0.5 Å increments from 3.5 to 7.5 Å. The angle between the sulfur and the normal to the benzene plane (denoted θ) was varied in 15° increments at every R value in the range described. At each R/θ point, the total interaction energy of the complex was determined using MP2 in conjunction with the aug-cc-pVDZ basis set. Though this method is not sufficient to determine accurate total interaction energies, the relative energetics of the configurations, and therefore the qualitative appearance of the surface, can be determined reliably. To verify this assumption, a portion of the surface for configuration B was determined using CCSD(T) in conjunction with the aug-cc-pVTZ basis, and qualitative agreement was found across the region considered. From the R/θ surfaces, interesting configurations were selected for higher level analysis. For these configurations, potential energy curves were obtained using CCSD(T) in conjunction with the aug-cc-pVTZ basis. Previous work [110] has demonstrated that reliable interaction energies can be produced for the H_2S-benzene complex using this methodology. All energy computations were performed using MOLPRO [126].

3.2.2 Protein databank analysis

The data set of PDB structures was determined by selecting protein structures which contained a cysteine residue and at least one phenylalanine, tyrosine, or tryptophan residue with better than 4.2 Å resolution. Histidine residues (which are frequently charged) were excluded to avoid entangling a sulfur-π interaction with a cation-π interaction. A custom

Perl script was developed which defined the center of each aromatic ring (for tryptophan it defined a center for both the five-membered and the six-membered ring) and determined the distance (denoted R) between that point and the sulfur of the cysteine residue. Any R less than 12 Å was considered a sulfur-π contact in this analysis. For each of these contacts, the angle (denoted θ) between the vector connecting the ring centroid and the sulfur and the normal to the aromatic ring was determined. If more than one protein structure gave duplicate contacts, only the highest resolution structure was retained in the data set. The final data set contained 753 protein structures, 642 of which had better than 2.5 Å resolution.

The resulting data was binned in 0.5 Å increments for R and 5° increments for θ. However, for each R/θ bin, the corresponding volume of the search area differs. Without correcting for this volume difference, many more contacts appear in bins with larger values for R and θ, even though these contacts are simply the result of the larger search area and not a preference for a particular geometry. To correct this effect, the number of contacts for each R/θ region is divided by the volume element:

$$V = \frac{2\pi}{3}(R_{max}^3 - R_{min}^3)(\cos\theta_{min} - \cos\theta_{max})$$

where R_{max}, R_{min}, θ_{max}, and θ_{min} represent the maximum and minimum values defining each bin. Using this normalization factor, a large number of normalized contacts would indicate more contacts were found in a particular region than would be expected from a random distribution.

3.3 Results and discussion

3.3.1 Configuration selection and ab initio results

The MP2/aug-cc-pVDZ R/θ surfaces generated for each model configuration depicted in Figure 10 are shown in Figures 11 and 12. Based on these surfaces, three local minima are identified in this R/θ space, which are depicted in Figure 13. For the hydrogens-down configuration (Figure 11), only one local minimum is found, at very short R (less than 4 Å) and θ=0° (configuration A of Figure 13). This corresponds to the configuration studied in great computational detail by Tauer et al., who found the equilibrium configuration at

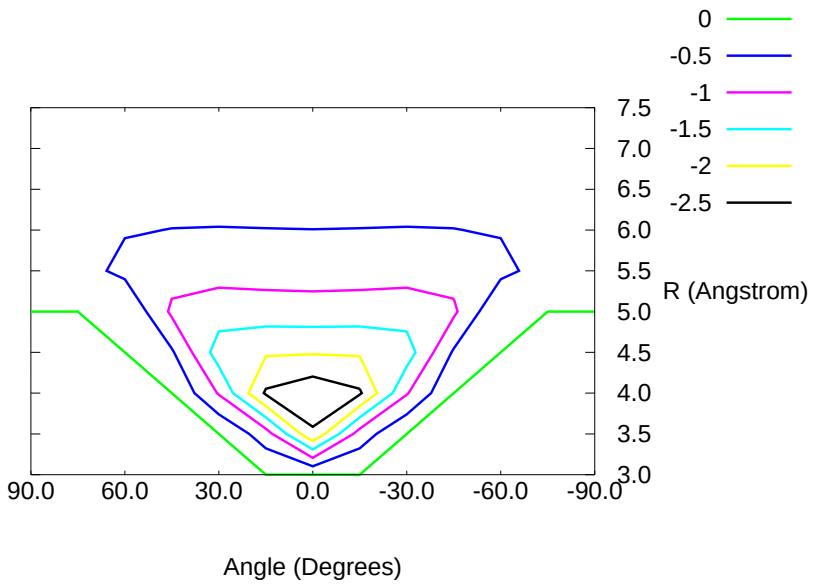


Figure 11: Contour plot of the potential energy surface for hydrogens-down configuration of H_2S-benzene; energy (kcal mol^{-1}) as a function of the distance between monomers measured from the H_2S sulfur to center of benzene and the angle between the sulfur and the normal to the benzene ring.

R=3.8 Åwith a total CCSD(T)/aug-cc-pVTZ interaction energy of -2.64 kcal mol^{-1}. For the hydrogens-up configuration (Figure 12), two local minima in R/θ space can be identified. One has a similar configuration to the hydrogen-down minimum, with R around 3.5 Å and θ=0°; the other is found at around R=5.5 Å and θ=90°.

For each of the two local minima in this R/θ space resulting from the hydrogens-up starting configuration (configurations B and C of Figure 13) which were not included in the study of Tauer et al., a complete potential energy curve using CCSD(T) with the aug-cc-pVTZ basis, was determined by fixing θ and varying R. The curves are depicted in Figures 14 and 15. The equilibrium configuration for B (θ=0°) is found at R=3.6 Å and has a total interaction energy of -1.12 kcal mol^{-1}. For C, where θ=90°, the equilibrium inter-fragment separation is R=5.5 Å, and this configuration has a total interaction energy of -0.74 kcal mol^{-1}. The equilibrium geometries and interaction energies of all three model systems are

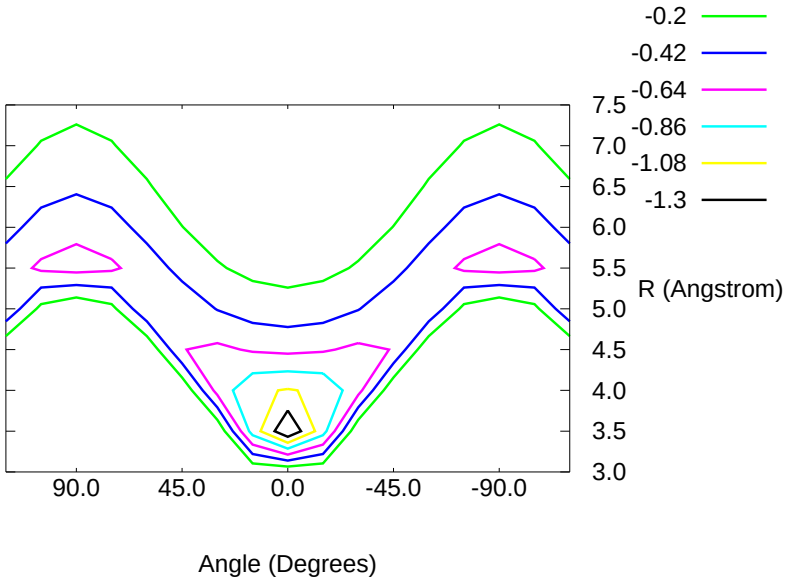


Figure 12: Contour plot of the potential energy surface for hydrogens-up configuration of H_2S-benzene; energy (kcal mol^{-1}) as a function of the distance between monomers measured from the H_2S sulfur to center of benzene and the angle between the sulfur and the normal to the benzene ring.

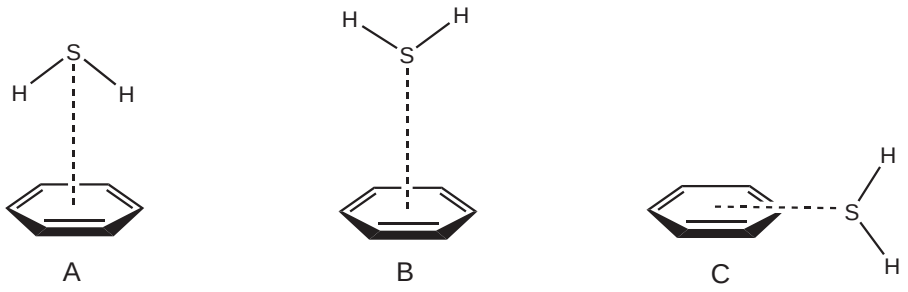


Figure 13: Configurations selected for higher-level analysis.

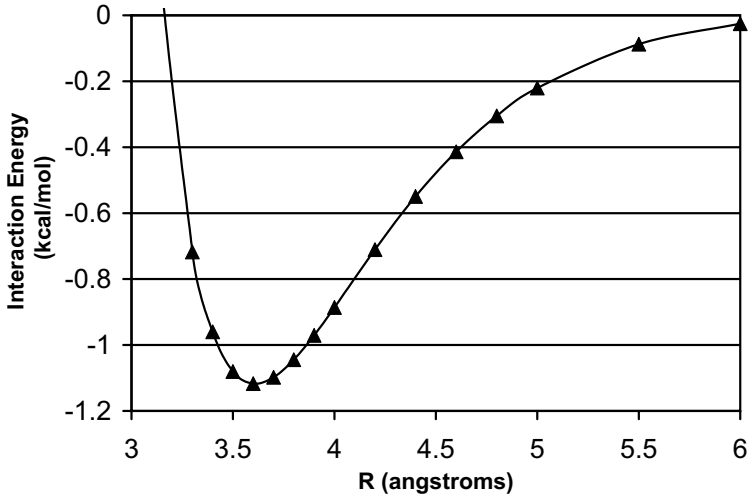


Figure 14: Potential energy curve of configuration B of the H_2S-benzene complex.

Table 6: Equilibrium geometries and CCSD(T)/aug-cc-pVTZ interaction energies for the configurations of the H_2S-benzene complex.

Configuration	R(Å)	$\theta(^\circ)$	ΔE_{int}(kcal mol^{-1})
A	3.8	0.0	-2.64
B	3.6	0.0	-1.12
C	5.5	90.0	-0.74

summarized in Table 6.

The model systems only encompass two possible orientations the hydrogen atoms could adopt relative to the aromatic ring. For this reason, although configurations A-C are local minima in the R/θ space considered, this does not mean that they are actually local minima in the full 3N-6 dimensional space of all their internal coordinates, or even local minima in the space of all intermolecular degrees of freedom with rigid monomers. Because the goal of the work is not to characterize the spectroscopic properties of the H_2S-benzene complex itself, but to understand the basic energetic properties of the sulfur-π interactions as they may occur within the constraints of protein structures, this is not problematic:

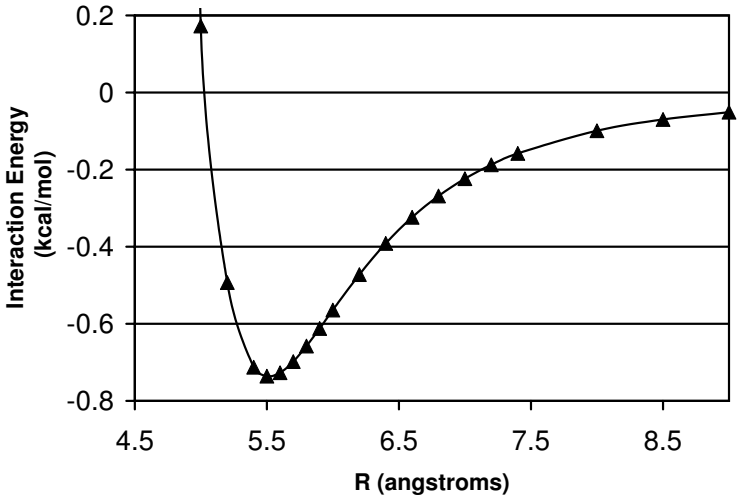


Figure 15: Potential energy curve of configuration C of the H_2S-benzene complex.

the symmetric configurations considered have very similar energies to nearby configurations in which the H_2S is rotated along symmetry-lowering coordinates. Starting from the optimal inter-fragment distances for model configurations A-C, unconstrained geometry optimizations were performed at the MP2/aug-cc-pVDZ level of theory within the appropriate point-group symmetries. Configurations A and C each had two imaginary frequencies and are therefore second-order saddle points, while configuration B is a transition state with only one imaginary frequency. For configuration A, one of the imaginary frequencies was followed to a minimum configuration, which looked like configuration A but with one hydrogen directed towards the center of the ring, as though the H_2S of configuration A had been tipped to the side. This configuration, previously identified by Sherrill and coworkers [110] was confirmed to be a minimum of the full potential surface by frequency analysis and is in agreement with a minimum configuration identified by Arunan [3]. The minimum configuration differed from configuration A by only 0.02 kcal mol^{-1}, and the optimum inter-fragment separation was very similar, 3.7 Å. Attempts were made to follow the other imaginary frequencies to their corresponding local minima, but the potential surface is so

flat in these regions that the optimizations could not converge in a reasonable amount of computational time. Therefore, to further verify that the model configurations considered appropriately describe the preferred geometries of S/π interactions generally, seven alternate configurations, which were selected to mimic the geometries observed in a random sampling of PDB entries, were examined that were similar to the model configurations except for the orientation of the hydrogen atoms. The energy for these configurations was determined at the MP2/aug-cc-pVDZ computational level and the R and θ values for the configurations were measured and compared to corresponding model configurations. For instance, one alternate configuration examined placed the sulfur directly above the ring (θ=0) with the H-S-H plane parallel to the aromatic plane. This configuration is similar in energy (within 0.1 kcal mol^{-1}) to the corresponding hydrogens up configuration (B) despite the differing orientations of the hydrogens. Overall, for all the PDB-like alternative configurations considered, good agreement was found for the interaction energy of the configuration and the symmetric model that would represent it.

If one considers a slightly larger small model system such as methanethiol-benzene, more consideration must be given to the positions of the hydrogen and methyl group than was required for the hydrogens of the simple H_2S model. For a methanethiol-benzene complex, a configuration analogous to configuration A directs a methyl group towards the aromatic ring. This configuration has destabilizing interaction energies for R values less than 4.0 Å and is not an appropriate representation of a cysteine/aromatic interaction in a protein structure because the aliphatic sidechain would likely be in contact with the aromatic ring if the sulfur were to be that close to the ring in that orientation. A more physically motivated methanethiol-benzene model would direct the single hydrogen of methanethiol towards the center of the aromatic ring, as in the minimum energy configuration of H_2S-benzene. This configuration was examined at the MP2/aug-cc-pVDZ computational level for inter-fragment separations from 3.0 to 6.0 Å. The potential energy curves for the methanethiol model and the symmetric H_2S-benzene are nearly parallel and separated by about 0.5 kcal mol^{-1}.

For methanethiol-benzene complexes in which the hydrogen and the methyl group are

directed away from the aromatic ring, configurations analogous to B and C are appropriate models for cysteine/aromatic interactions, and direct comparisons can be made between the methanethiol model and the H_2S model. For the B configurations, partial potential energy curves were compared for R values from 3.0 to 5.5 Å, and the curves were not only almost parallel, but nearly coincident, with differences in the MP2/aug-cc-pVDZ interaction energy always less than 0.1 kcal mol^{-1}. The difference between the two curves is slightly greater for configurations like C, around 0.25 kcal mol^{-1}, but the curves are still largely parallel. Overall, H_2S is qualitatively comparable to methanethiol in terms of the preferred interaction geometries it predicts for cysteine/aromatic interactions in protein structures, and, in fact, H_2S is a preferable model in the flexibility it allows in the placement of hydrogen atoms in the model system.

The difference between the methanethiol model and the H_2S model in different configurations suggests how the nature of the interaction changes with changing configuration. For the configurations where θ=0 and the hydrogens (or methyl group, in the case of methanethiol) are directed towards the aromatic ring (A), the methanethiol-benzene complex is more stabilizing than the corresponding H_2S-benzene complex, suggesting that the increased dispersion interaction of the methyl group increases the interaction energy of the complex. However, if this model is flipped (to configurations like B), the methanethiol complex is *less* stabilized than the corresponding H_2S model. In this case, the electron donating methyl group has likely increased the electron density on the sulfur atom, and the electrostatic electron repulsion is more destabilizing (though, overall, the dispersion interaction does lead to a stabilizing interaction energy for both complexes). When this model is rotated to the inplane configuration (C), the trend is reversed and methanethiol-benzene again becomes more stabilized than H_2S-benzene. In this case, the increased electron density on the sulfur atom creates a more favorable interaction with the partially positive hydrogen of the benzene ring, further stabilizing the interaction.

3.3.2 Comparison to data mining results from the Protein Databank (PDB)

Each contact located by the data-mining script was sorted into bins according to its R/θ value. Each bin has a width of 0.5 Å in R-space and 5° in θ-space. The results were normalized using the volume element described in the computational details. A 2D histogram was constructed to display the R vs. θ data and is shown in Figure 16. The histogram shows two significant clusters of peaks. The largest is found for short distances (less than 4 Å) and small angles (less than 10°). The tallest peaks in this group are found for R=3.5-4.0 Å, θ=0°-10°, which corresponds to the equilibrium geometries of model configurations A and B. A second, shorter cluster of peaks is found for large angles (θ=75°-90°) around R=5.5 Å. The largest peaks in this group are found for R=5.0-5.5 Å, θ=85°-90°, which corresponds to the equilibrium geometry of model configuration C. Overall, the results indicate that the three configurations suggested by the local minima of the R/θ-surfaces for the simple H_2S-benzene complex are, in fact, the configurations which are found in protein structures in the PDB for cysteine S/π contacts.

Interestingly, the region of the histogram between R=5.0-7.0 Å for small angles (less than around 20°) has noticeably few contacts. This is again reflected by quantum mechanical results. Considering the energetics of the transition from configuration B to configuration C (shown in the contour plot of Figure 12) as one moves to larger inter-fragment separations, the interaction energy of directly above configurations becomes less favorable. The more favorable configurations at these distances are not small angles directly above the center of the ring, but offset configurations with larger values for θ. Indeed, the histogram shows an increase in the number of contacts as one moves to larger angles at these values of R, culminating with the cluster of peaks around θ=75°-90°. This preference for offset configurations levels off at around 7.0 Å, when the interaction energy of the complex is very small and all geometries become approximately equally preferred.

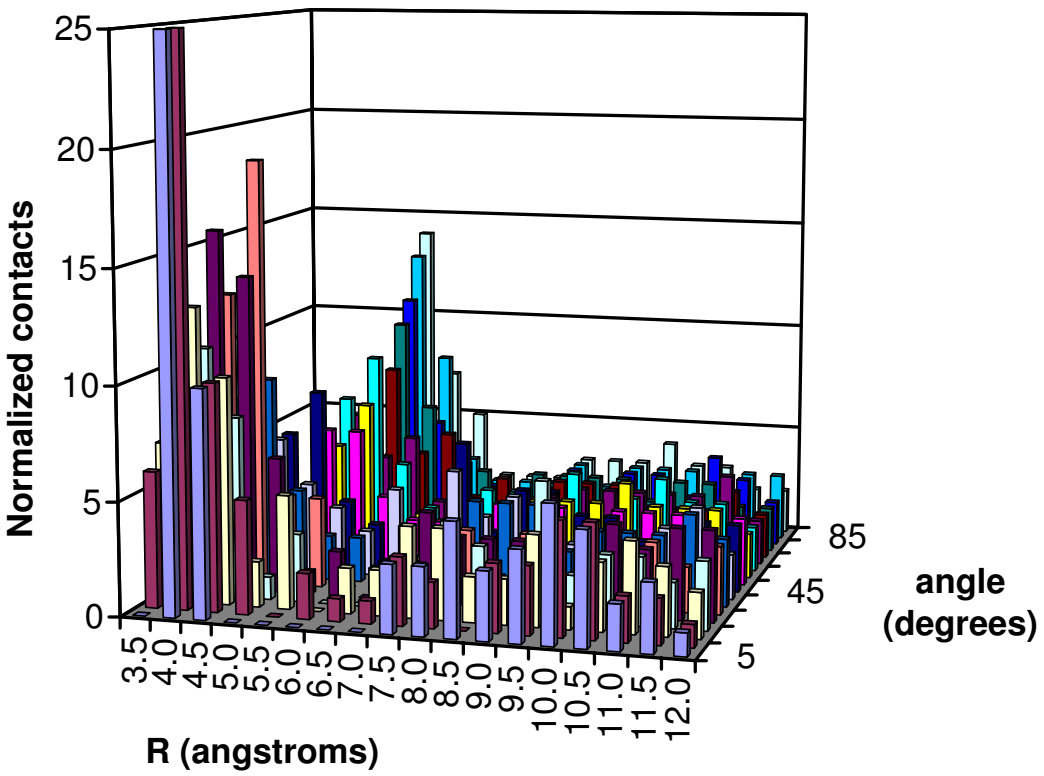


Figure 16: Histogram depicting number of normalized sulfur/π contacts from PDB data mining.

3.3.3 Predicting probability distributions using Boltzmann weighted distributions

From the relative energies of different configurations, the ratio of probabilities can be determined using the Boltzmann distribution. The ratio of the probabilities for two states A and B is given by

$$\frac{P_A}{P_B} = e^{-\beta \Delta G}.$$

Taking configuration A as the reference, the relative interaction energies are determined for all the configurations which were included in the potential energy surface scan. (These configurations are depicted by Figure 10 and the interaction energies are shown in Figures 11 and 12.) From these ΔE values (where it is assumed ΔE reasonably approximates ΔG), the probability relative to configuration A is determined. From these probability ratios, a histogram similar to the that pictured in Figure 16 is constructed. This probability distribution qualitatively agrees with the observed probabilities depicted in Figure 16, though the ratio between configuration A and configuration is slightly lower than in the observed probability distribution.

3.3.4 Comparison to other database results

Previous database studies, which identified preferred configurations for S/π interactions using only database mining results without any insight from ab initio computations, often came to differing conclusions about the preferred configuration of the interaction. In the study of Reid et al. [90], thirty-six high resolution (better than 2.5 Å resolution) crystal structures were obtained from the Protein Databank and analyzed for contacts between sulfur atoms (from cysteine or methionine) and aromatic rings (from phenylalanine, tyrosine, or tryptophan). Several geometric parameters were analyzed including the distance between the sulfur and the aromatic center (analogous to the parameter R of this study) and an angle describing the rise of the sulfur relative to the plane of the aromatic molecule. For each parameter, the number of occurrences was reported over the range of the parameter and compared to the number of occurrences that might be expected randomly, based on volume considerations. However, no two-dimensional correlation is presented to ascertain

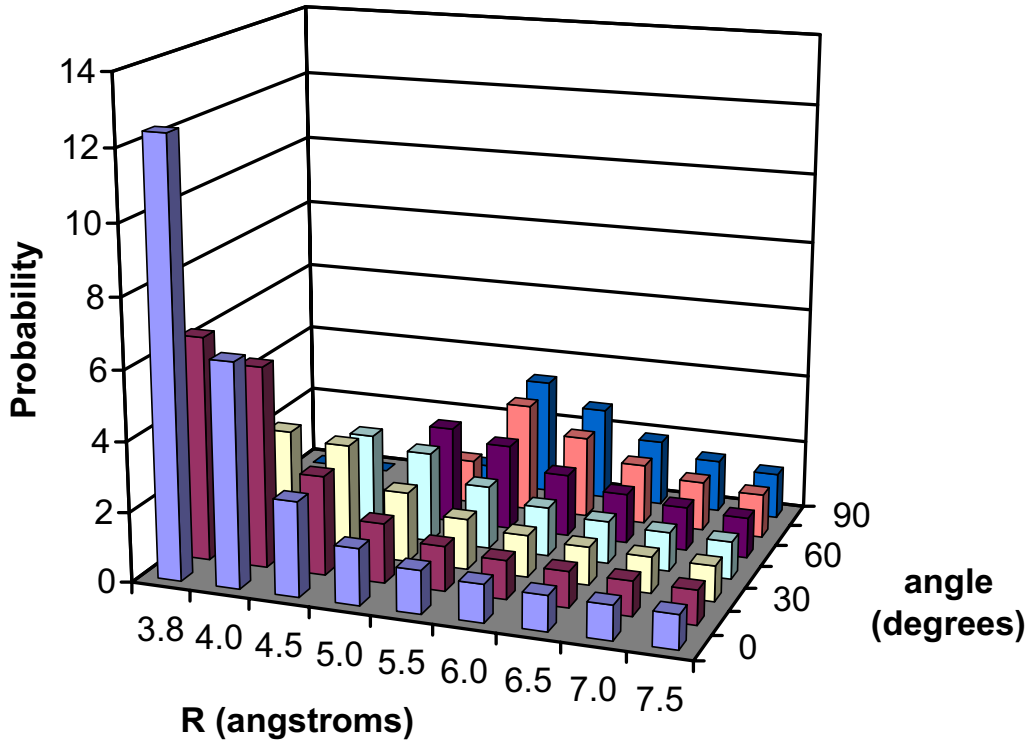


Figure 17: Histogram depicting the probabilities predicted by a Boltzmann distribution of the electronic energy for each configuration.

if particular distances appear more commonly at particular angles.

In the work of Zauhar et al., the authors made two-dimensional comparisons to correlate the relationship between the optimum separation distance and the preferred angle relative to the aromatic ring. However, their study examined divalent sulfur groups of the form X-S-X, so they do not consider the possibility of hydrogens interacting with the aromatic ring. This makes their analysis more comparable to the results for configurations B and C, in which the hydrogens are directed away from the ring and the lone pairs of the sulfur atom are interacting with the aromatic system. Their 2D histograms show a maximum at 90° angles (sulfur in-plane with the aromatic ring) for large separations, and 0° or 180° (sulfur directly above or below the aromatic ring) for short separations, in general agreement with the quantum mechanical results. Additionally, they report a local maximum from 5.0-5.5 Å, extending over the 60°-115° angle range. Therefore, they report that the "ideal" sulfur-aromatic interaction geometry (as opposed to an S-H/π interaction geometry) is an in-plane configuration at a separation of around 5 Å (similar to configuration C), while the results of this study would suggest it is a configuration in which the sulfur is directly above the aromatic ring at a shorter separation of 3.6 Å, as in the equilibrium geometry for configuration B.

This discrepancy in conclusions may lie in the normalization technique used by Zauhar et al, in which they compared their probability distributions to analogous probability distributions for C-CH_2-C group interacting with aromatic rings and looked for statistically significant differences between the two distributions. This necessarily assumes that there is no significant interaction between the CH_2 group and the aromatic system, and that it, therefore, can be used for a control. However, other work (discussed in Chapter 4 and in Reference 93) has shown that there is a significant interaction between alkyl C-H groups and aromatic rings and that this interaction has distinct geometric preferences which happen to be very similar (R=3.7 Å and θ=0°) to the preferred configuration reported in this work for configuration B of the H_2S-benzene complex. The optimum configuration identified by Zauhar et al. is not necessarily the ideal sulfur-aromatic interaction configuration; rather, it is simply the preferred interaction configuration that is dissimilar to the preferred interaction

configuration for alkyl C-H-aromatic interactions.

3.4 Conclusions

In this study, three local minima for the H_2S-benzene complex were identified on constrained MP2 potential energy surfaces which varied both the distance between the sulfur and the center of the benzene ring and the angle between the sulfur and the normal to the plane of the aromatic ring. For each configuration identified, CCSD(T) potential energy curves were generated with the aug-cc-pVTZ basis set, which should provide accurate binding energies to within a few tenths of a kcal mol^{-1}. One of these configurations centered the H_2S molecule directly above the center of the benzene ring with the hydrogens directed towards the aromatic ring; this configuration has previously been examined and the optimum complex configuration has an inter-fragment separation of 3.8 Å and a total interaction energy of -2.64 kcal mol^{-1}. In the other two local minima identified in this study, the hydrogen atoms are directed away from the aromatic ring. For the hydrogens-away configuration centered directly above the benzene ring, the best estimate of the total interaction energy is -1.12 kcal mol^{-1} with an optimum inter-fragment separation of 3.6 Å. For the hydrogens-away in-plane configuration, the best estimate of the total interaction energy is -0.74 kcal mol^{-1} with an optimum inter-fragment separation of 5.5 Å.

Taking the H_2S-benzene complex as the simplest prototype for S/π interactions, the optimum geometries predicted by these potential energy curves were compared to the sulfur-π contacts which appear in protein structures from the Brookhaven Protein Databank. The number of occurrences for each search area was normalized to account for the different volumes of each area. Two regions of the resulting histogram showed a large number of normalized contacts, indicating that significantly more contacts appear than one would expect from a random distribution of atoms. These regions corresponded to the geometries of the minimum configurations predicted by the ab initio calculations for the H_2S-benzene complex. These results validate the use of quantum mechanics calculations on small model systems to predict the geometries of interactions in protein structures.

CHAPTER IV

ALIPHATIC C-H/π INTERACTIONS: METHANE-BENZENE, METHANE-PHENOL, AND METHANE-INDOLE COMPLEXES

4.1 Introduction

Noncovalent interactions are prevalent in biochemical molecules and play a role in numerous chemical processes. Of these, the classic hydrogen bond is considered one of the most important, but over the past few decades, evidence has accumulated in support of the significance of a much weaker "hydrogen bond" occurring between an aliphatic C-H group and an aromatic π system [77]. This type of noncovalent interaction has been shown to contribute to crystal packing, stereoselectivity, and protein stability and conformation [112, 76, 11, 121]. The C-H/π bond also plays a vital role in molecular recognition for numerous ligand-binding proteins [78, 101]. Muraki reported that the interaction is common in carbohydrate binding proteins where it affects both binding affinity and conformation [72]. The interaction has already been used in drug design [125], where it is responsible for an increase in the affinity and selectivity of a thrombin inhibitor [79] and for a significant increase in the inhibitory activity of a tyrosine phosphatase inhibitor [120]. The importance of furthering the understanding of the C-H/π interaction and quantifying its energetics has been recognized [69].

Analysis of known protein structures has shown the C-H/π interaction frequently occurs between the aliphatic and aromatic groups in protein side-chains [11]. In this work, we study the simplest representation of these systems, using methane as a model of aliphatic side-chains, and benzene, phenol, and indole as the aromatic components of phenylalanine, tyrosine, and tryptophan, respectively. Full potential energy curves are of special interest given that the constrained environments of proteins give rise to individual interactions that may not be in the configurations which would be optimal if the interaction were considered in isolation. In addition to providing insight for drug design and supramolecular chemistry,

these high-accuracy computations should be helpful for the calibration of molecular force fields [66] and the development of density functional theories that attempt to accurately model dispersion interactions [44, 123, 124, 30, 82, 47, 8, 137, 42, 60, 132, 133, 35, 136].

The highest-level computations performed previously for the prototype methane-benzene complex were reported by Tsuzuki and coworkers [114, 100]. Potential energy curves were computed for six configurations of the complex, and the lowest energy orientation found was one in which the methane is centered on top of the benzene ring and one C-H bond points directly toward the center of the ring. The interaction energy for this configuration was computed using MP2 extrapolated to the complete basis set limit, with additional CCSD(T) correction terms. In recent work Tsuzuki and coworkers [100] determined potential energy curves for the complex using both correlation consistent (cc-pVXZ) and augmented correlation consistent (aug-cc-pVXZ) basis sets. The interaction energies were extrapolated to the complete basis set limit, using both the Helgaker [46] and Feller [36] basis set extrapolation techniques. To our knowledge, similar high-level studies have not been performed for the methane-phenol or methane-indole complexes.

In the present study of methane-benzene, methane-phenol, and methane-indole complexes, results are obtained using MP2 in conjunction with Dunning's augmented correlation-consistent basis sets, aug-cc-pVXZ (X = D, T). In addition, for the methane-benzene complex, basis set effects were carefully explored by using the very large aug-cc-pVQZ basis as well as extrapolation techniques to approximate the complete basis set (CBS) limit. This work expands upon the recent work of Tsuzuki and coworkers [100] for this complex by presenting high-quality aug-cc-pVTZ/aug-cc-pVQZ extrapolations to the CBS limit for the entire potential energy curve. Corrections to the MP2 energies were obtained using the robust CCSD(T) method with the smaller basis sets. These best estimates should provide binding energies accurate to within a few tenths of a kcal mol^{-1}.

4.2 Computational details

Monomer geometries were optimized using second-order perturbation theory (MP2) and the cc-pVDZ basis set, and these frozen monomer geometries were utilized in all computations

of the complexes. To verify that the monomer geometry is not significantly changed in the complex, the methane-benzene complex was fully optimized using MP2 and the cc-pVDZ and aug-cc-pVDZ basis sets. No significant geometry changes were found with either basis set; for example, the length of the C-H bond pointing to benzene varied by no more than 0.002 Å and the hydrogens of benzene were bent by only 0.3 degrees. The MP2/cc-pVDZ computational level was also used for single-point energy calculations to select low-energy complex configurations. While this basis is not sufficient to determine accurate total binding energies (because it lacks diffuse functions), it is adequate to determine which are the low energy configurations.

MP2/aug-cc-pVXZ (where X = D and T) computations were performed for five selected complex configurations, depicted in Figure 18. For these configurations, the interfragment separation distance was varied over at least a 3 Å range using a 0.1 Å stepsize to find the equilibrium distances. CCSD(T) potential curves were determined explicitly using only the aug-cc-pVDZ basis set; the CCSD(T)/aug-cc-pVTZ potential curve was estimated for each complex by calculating a correlation correction term as the difference between the MP2 and CCSD(T) energies determined in the aug-cc-pVDZ basis. This change, denoted ΔCCSD(T), is then added to the MP2/aug-cc-pVTZ results, giving an estimated CCSD(T)/aug-cc-pVTZ interaction energy. This methodology is appropriate because the ΔCCSD(T) correction term is quite insensitive to basis set effects [103]. To further verify the validity of this ΔCCSD(T) addition method, the CCSD(T)/aug-cc-pVTZ interaction energy was explicitly determined for the benzene-methane complex at an equilibrium interfragment separation of 3.8 Å and was in excellent agreement (within 0.01 kcal mol^{-1}) with the estimated value.

Previous experience with the benzene dimer [106, 104] demonstrates that the interaction energies of noncovalent complexes frequently converge more rapidly when the Boys-Bernardi counterpoise correction [10] is employed. To determine if it the counterpoise correction should be employed for C-H/π complexes, both counterpoise-corrected and non-corrected MP2 interaction energies were determined for the methane-benzene complex using the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets as shown in Figure 19. The Figure

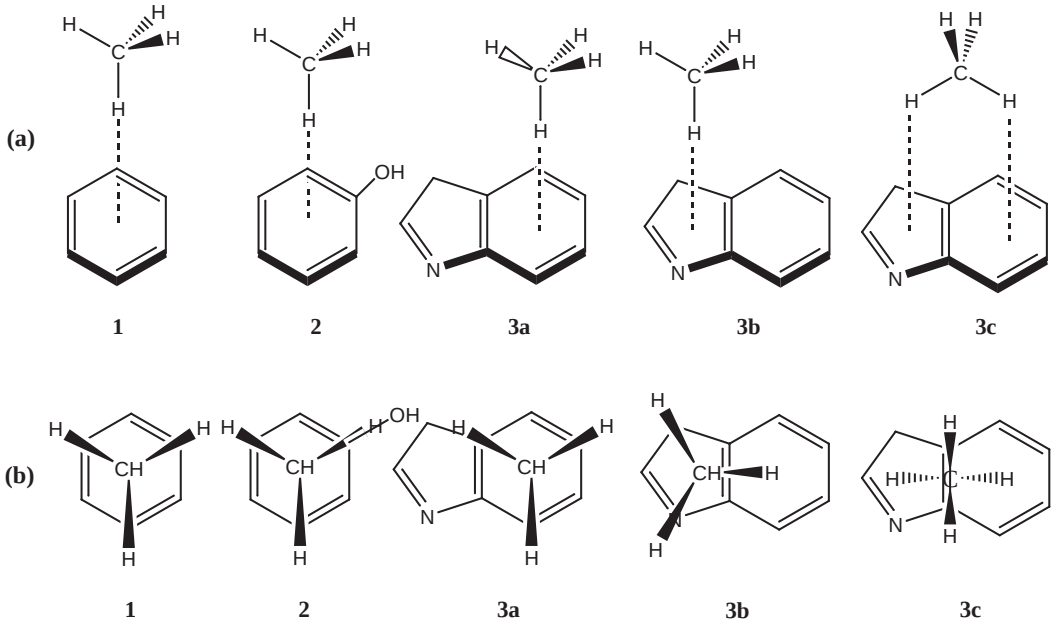


Figure 18: Configurations of methane-benzene, methane-phenol, and methane-indole complexes.

demonstrates that convergence with respect to basis set is greatly accelerated by the counterpoise correction; hence, the counterpoise correction is applied to all results reported here. Optimizations of monomer geometries were performed using Q-Chem 2.1 [59], and energy computations for the complexes were performed using MOLPRO [126].

Symmetry-adapted perturbation theory (SAPT) [56, 128] was applied using the program package SAPT2002 [12] to divide the Hartree-Fock (HF) energy and the correlation energy into physically significant components, including electrostatic, induction, dispersion, and exchange energies, plus cross-terms for exchange-induction and exchange-dispersion. The SAPT2 approach has been employed, in which the correlated portion of the interaction energy is nearly equivalent to the supermolecular MP2 correlation energy. To simplify the discussion of the SAPT results, exchange-induction and exchange-dispersion will be counted as induction and dispersion, respectively. The $\delta E^{HF}_{int,resp}$ term, which includes the third- and higher-order induction and exchange-induction contributions, is also counted as induction. Because SAPT analysis can be quite time-consuming, a less expensive basis set was used to lower the computational cost. This basis set, denoted cc-pVDZ+, is the cc-pVDZ basis

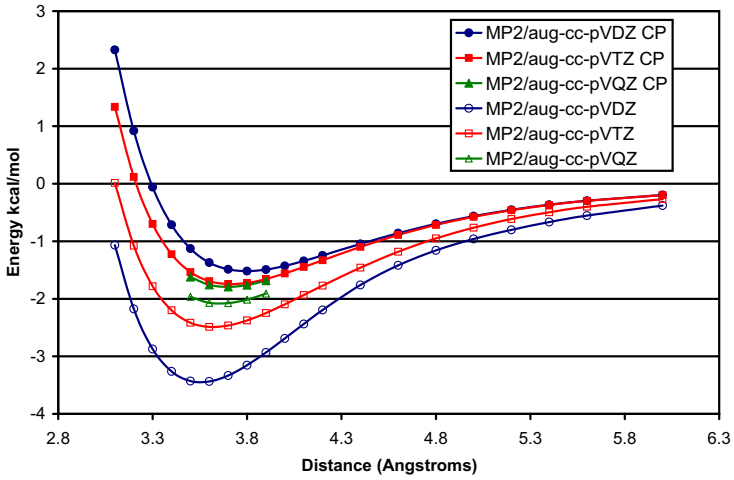


Figure 19: Effect of counterpoise (CP) correction on MP2 potential energy curves for the methane-benzene complex.

for hydrogen and an aug-cc-pVDZ basis minus diffuse d functions for all other atoms; this basis was used previously in SAPT analysis of the benzene dimer [104].

4.3 Results and discussion

4.3.1 Methane-benzene complex

Tsuzuki and coworkers [114] found that for the methane-benzene complex, the preferred configuration has the methane directly above the center of the benzene with one hydrogen pointed at the center of the ring, and three directed away from the center of the ring (complex **1** of Figure 18). Based on this result, a series of additional computations were performed to determine the effect of rotation of the methane about the axis containing the C atom of methane and the center of mass of benzene. The hydrogens of methane were rotated, in 10 degree increments, with the distance between methane carbon and the center of mass of benzene fixed at 3.8 Å. The results show less than a 0.001 kcal mol^{-1} variation in the energy. Therefore, the C_{3v} symmetric complex (as depicted as **1** of Figure 18) was

selected for higher-level analysis because of the greater computational efficiency afforded by its symmetry.

The potential energy curves determined using the MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ, MP2/aug-cc-pVQZ, and CCSD(T)/aug-cc-pVDZ levels of theory are depicted in Figure 20. The Figure demonstrates that the MP2 results are well converged with respect to basis set for the aug-cc-pVTZ and aug-cc-pVQZ basis sets. Energies for these two basis sets are then used to extrapolate to the MP2 complete basis set (CBS) limit using the method of Helgaker [46]. This extrapolation procedure was also utilized by Tsuzuki and coworkers [100] with two pairs of basis sets (cc-pVTZ/cc-pVQZ and aug-cc-pVDZ/aug-cc-pVTZ), along with an aug-cc-pVTZ/aug-cc-pVQZ extrapolation for a single optimized geometry. In this work, a complete curve was determined using an aug-cc-pVTZ/aug-cc-pVQZ Helgaker extrapolation and is shown in Figure 20. The ΔCCSD(T) correction shown in Figure 20 is determined by subtracting the CCSD(T)/aug-cc-pVDZ and MP2/aug-cc-pVDZ curves. This correction can then be added to the MP2 results to provide accurate estimations of the CCSD(T) interaction energy at the same basis set [104]. The ΔCCSD(T) correction decreases with increasing inter-fragment separation and goes to zero at large inter-fragment distances.

Results for the methane-benzene complex near equilibrium are presented in Table 7. All the results in this table are for a fixed inter-fragment separation of 3.8 Å, the equilibrium separation determined using the estimated CCSD(T) values extrapolated to the CBS limit. The MP2 results using the aug-cc-pVTZ (-1.723 kcal mol^{-1}) and aug-cc-pVQZ (-1.763 kcal mol^{-1}) basis sets show that the basis set is nearly converged, and extrapolating to the CBS limit (-1.790 kcal mol^{-1}) only changes the total interaction energy by 0.03 kcal mol^{-1}. These MP2 results are in reasonable agreement with those of Tsuzuki and coworkers [100], who determined the total interaction energy of the methane-benzene complex as -1.699 kcal mol^{-1} using MP2/aug-cc-pVTZ and -1.759 kcal mol^{-1} using MP2/aug-cc-pVQZ. The small differences in the results are most likely an effect of slightly different geometries for the complex; Tsuzuki and coworkers optimized the complex geometry using the MP2/cc-pVTZ computational level, while the geometry in this work is the equilibrium geometry

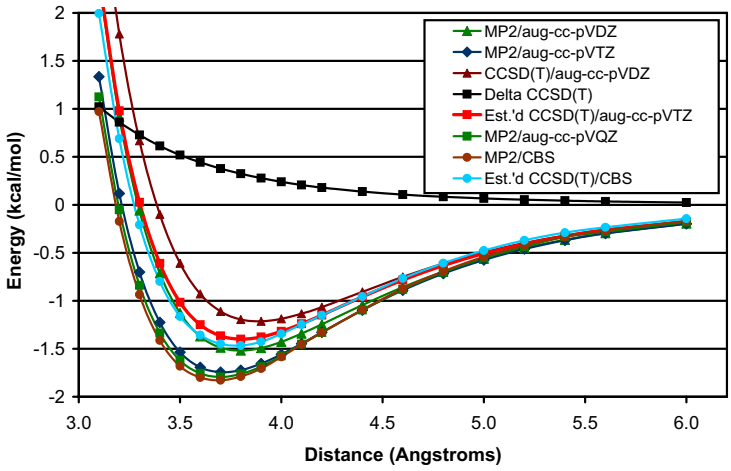


Figure 20: Potential energy curves of the benzene-methane complex.

from the estimated CCSD(T)/CBS potential energy curve. The interaction energy for the complex at an inter-fragment separation of 3.8 Å was explicitly determined using CCSD(T) for the aug-cc-pVDZ and aug-cc-pVTZ basis sets, and the ΔCCSD(T) correction is shown for both basis sets in Table 7. These results differ by about 0.01 kcal mol^{-1}, confirming that the ΔCCSD(T) correction is insensitive to basis set effects. Adding the aug-cc-pVTZ ΔCCSD(T) correction to the MP2/CBS results gives our best estimate of the total binding energy of the complex, -1.454 kcal mol^{-1} at an equilibrium inter-fragment separation of 3.8 Å.

Thus far, only considered a particular slice of the methane-benzene potential surface has been considered. To more fully explore the surface, from the initial complex configuration **1**, the angle between between the C-H bond and the normal to the aromatic plane of benzene (see Figure 21) was varied. In these computations, the original configuration (C-H bond of methane perpendicular to the π system) is denoted 0 degrees, and the configuration in which the C-H bond is in-plane with the aromatic ring is denoted 90 degrees. This angular space was scanned in 15 degree increments with the inter-fragment separation held constant

Table 7: Interaction energies (in kcal mol^{-1}) for the methane-benzene complex.

Method	ΔE_{int}[a]
MP2	
aug-cc-pVDZ	-1.519
aug-cc-pVTZ	-1.723
aug-cc-pVQZ	-1.763
CBS limit	-1.790
CCSD(T)	
aug-cc-pVDZ	-1.195
aug-cc-pVTZ	-1.387
Δ CCSD(T)	
aug-cc-pVDZ	0.324
aug-cc-pVTZ	0.336
Est.'d CCSD(T)	
aug-cc-pVTZ	-1.387
aug-cc-pVQZ	-1.400
CBS limit	-1.454

[a] At an inter-fragment (methane C to the center of the benzene ring) separation of 3.8 Å, the equilibrium distance at the estimated CCSD(T)/CBS level of theory from Figure 20.

at 3.8 Å. At this short inter-fragment separation, the total interaction energy of the in-plane configuration (relative to benzene and methane at infinite separation) was repulsive by over 50 kcal mol^{-1}; further exploration of this configuration found the most attractive interaction energy for an in-plane configuration at 5.5 Å. The inter-fragment separation was then varied in 0.1 Å increments from 3.4 to 5.7 Å, for the same angular space. The potential surface is shown in Figure 22.

The surface confirms that, among configurations which feature one hydrogen pointed directly towards the benzene center, the minimum for the methane-benzene complex is the configuration in which the C-H is directly over the aromatic ring. This is reasonable, given that this configuration provides the best access for the partially positive hydrogen to interact with the negative π system. As one moves to longer inter-fragment separations, the preferred angle changes to one in which the methane is offset from the perpendicular. Even at the equilibrium inter-fragment separation for offset configurations (40-50 degrees), these complexes are significantly less bound [maximum total CCSD(T)/aug-cc-pVDZ interaction energy is approximately -0.6 kcal mol^{-1}] than the minimum configuration where the C-H

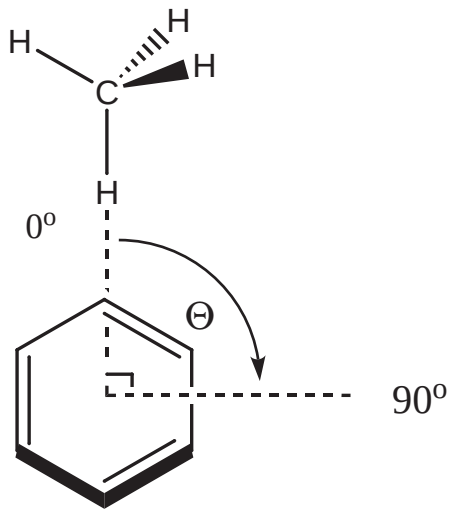


Figure 21: Angular space scanned for methane-benzene complex surface.

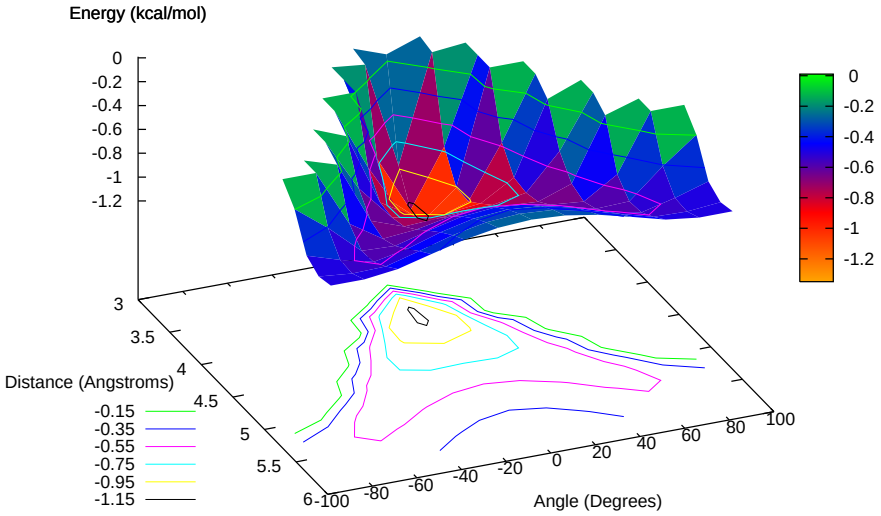


Figure 22: Methane-benzene potential energy surface; energy as a function of the distance between monomers measured from methane carbon to center of mass of benzene and the angle between C-H bond of methane and normal to the benzene ring (see Figure 21).

bond is perpendicular to the plane of the aromatic ring (-1.20 kcal mol^{-1} at the same level of theory), but they could still play a stabilizing role in proteins or other complex systems in which the geometry is constrained to non-ideal configurations.

Several studies have examined what C-H/π configurations are found in protein and peptide structures by analyzing databases of crystal structures [121, 11]. Taking the methane-benzene complex as a model system to describe a general C-H/π interaction, the computed interaction energies were compared to the results of database studies of Brandl et. al. [11] and Umezawa et. al. [121]. In the latter study, the authors examined a set of 130 peptide crystal structures from the Cambridge Structural Database (CSD) which contained a phenylalanine, tyrosine, or tryptophan residue. They counted intra- and intermolecular CH/π contacts separately, and tabulated these results according to the distance between the hydrogen of the C-H contact and the nearest carbon atom in the aromatic ring. Considering the intra- and intermolecular contacts together, the greatest number of contacts was found for the 3.02 to 3.04 Å bin, which corresponds well to the same distance in our minimum methane-benzene complex structure of 3.04 Å. However, beyond this equilibrium distance, the number of contacts falls off very quickly, whereas our results would predict a gradual decrease in the number of contacts because complexes at slightly larger inter-fragment distances retain a significant interaction energy. This discrepancy is likely due to the constraint of the searching parameters in the study, which would prevent counting of interactions with larger inter-fragment distances. In the study by Brandl et. al. [11], the authors examined a much larger set (1154) of protein structures from the Protein Data Bank (PDB) for close interactions between C-H-donors and π-acceptors. They defined a parameter d_{C-X} as the distance from the carbon of the C-H system to the center of mass of the aromatic systems (Figure 2 of Reference 11), the same parameter varied for our potential surfaces. They also constrain their selection criteria to select configurations above or below the π system, and not in-plane with the π system. This geometric search area corresponds to the well in the potential surface. The distribution of observed C-H/π contacts as a function of the d_{C-X} distance is shown in Figure 3 of Reference 11. The maximum frequency was found for d_{C-X} distances of 3.7-3.8 Å depending on the resolution of the data set considered. This

is in excellent agreement with the equilibrium distance of 3.8 Å the quantum mechanical results would predict. The frequency of contacts is low (near 0 %) for distances shorter than 3.0 Å, distances at which positive interaction energies were found. Between 3.0 Å and the maximum value at 3.7-3.8 Å there is a steady increase in the frequency of contacts, as the predicted interaction energy becomes more attractive. At distances greater than 3.8 Å the frequency of contacts again begins to decrease, corresponding to less bound complexes on the potential energy surface. The qualitative agreement of this distribution with our potential energy surface is very encouraging and suggests that, despite a number of serious complicating factors (solvent effects, steric constraints, secondary interactions, etc.), there may nevertheless be a good correlation between the observed properties of noncovalent interactions in complex systems and the predicted properties of these interactions in small model systems.

4.3.2 Methane-phenol complex

The electrostatic potential above the ring in phenol is similar to that of benzene [102]; therefore it seems reasonable to expect that the C-H/π interaction in the methane-phenol complex might have similar geometric preferences as the methane-benzene complex. An analogous configuration (complex **2** of Figure 18) was examined, along with two additional configurations, both of which had two hydrogens directed towards the aromatic system. Both of these additional configurations positioned methane over the phenol ring and placed two hydrogens coplanar to the C-O bond of phenol. One configuration centered the methane carbon over the center of the ring, while the other configuration was shifted such that the methane carbon was over the substituted carbon of phenol. All three configurations were similar in energy (differences of about 0.1 kcal mol^{-1} at the MP2/cc-pVDZ level of theory), but the one hydrogen down configuration (**2** of Figure 18) was the only configuration chosen for higher-level analysis because it was the lowest in energy and was the most similar to the equilibrium benzene-methane configuration. A similar configuration with methane directly above the center of the ring and with two hydrogens directed down towards benzene was

examined for the benzene-methane complex by Tsuzuki [114], who also found this configuration slightly higher in energy than the one-hydrogen down configuration, except at short inter-fragment distances. The effect of rotating the methane over the phenol was examined in the same manner as for the methane-benzene complex, and at a separation distance of 3.8 Å the energy of the complex varied at most 0.007 kcal mol^{-1}. It is interesting to note that while rotational effects were not significant for the structure in which one hydrogen was directed towards the aromatic ring, for the two configurations in which two hydrogens were directed towards the ring, rotational effects were somewhat more pronounced, on the order of 0.2 kcal mol^{-1} at distances of 3.8 Å.

For the selected one hydrogen down configuration (**2** in Figure 18), potential energy curves and the ΔCCSD(T) curve are illustrated in Figure 23. The best estimate of the interaction energy is -1.47 kcal mol^{-1} at the estimated CCSD(T)/aug-cc-pVTZ level of theory with an equilibrium inter-fragment separation of 3.8 Å. These results are very similar to the interaction energy of -1.40 kcal mol^{-1} and inter-fragment separation of 3.8 Å found for the methane-benzene complex at the same level of theory, indicating that the hydroxyl substituent has only a minor effect. Note that a single hydroxyl group also had a minor effect in sandwich and T-shaped benzene complexes [102, 104].

4.3.3 Methane-indole complex

For the methane-indole complex, the two aromatic rings of indole necessitated more exploration of geometric binding preferences for the complex. Nine initial configurations were evaluated: methane centered over the six-membered ring, methane centered over the five-membered ring, and methane centered over the bond shared between the five- and six-membered rings, each with one, two, or three hydrogens directed towards the aromatic centers. Of these configurations, the lowest energy configuration centered the methane over the shared bond of indole with one hydrogen pointing towards the center of each ring (**3c**, Figure 18). This configuration, along with the one hydrogen down configurations centered over the five- (**3b**) and six-membered (**3c**) rings (those most analogous to the minimum

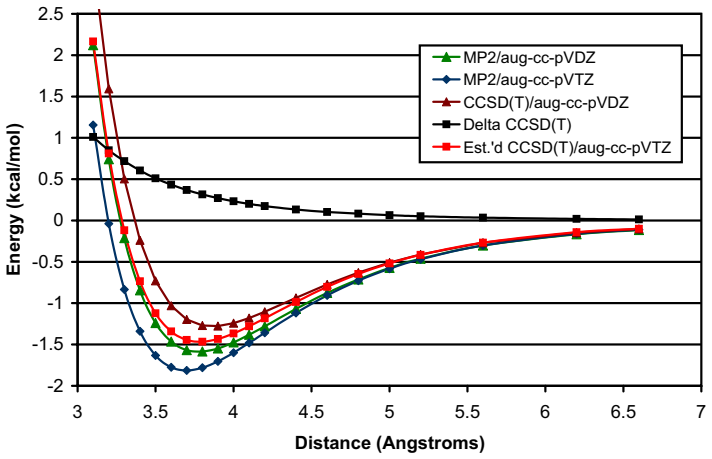


Figure 23: Potential energy curves of the phenol-methane complex.

configurations for methane-benzene and methane-phenol), were chosen for additional analysis.

The effect of rotating the methane hydrogens around the axis containing the methane carbon and the geometric center of the ring (for configurations **3a** and **3b**) or the axis containing the methane carbon and the center of the shared bond (for configuration **3c**) was considered for these three configurations by the procedure described in previous sections. Configuration **3c** was subject to the most significant rotational effects; rotation of the hydrogens of methane around the axis connecting the methane carbon and the center of the shared bond caused a maximum destabilization of 0.4 kcal mol^{-1}, when the hydrogens facing indole were coplanar with the shared bond. Rotational effects were not significant for either of the one hydrogen down methane-indole configurations (**3a** and **3b**).

The potential energy curves as a function of inter-fragment distance for these three indole-methane complex configurations (**3a, 3b, 3c**, Figure 18) are shown in Figures 24-26. The best estimate for the most attractive interaction energy of the indole-methane complex is the estimated CCSD(T)/aug-cc-pVTZ interaction energy for configuration **3c**, -2.08

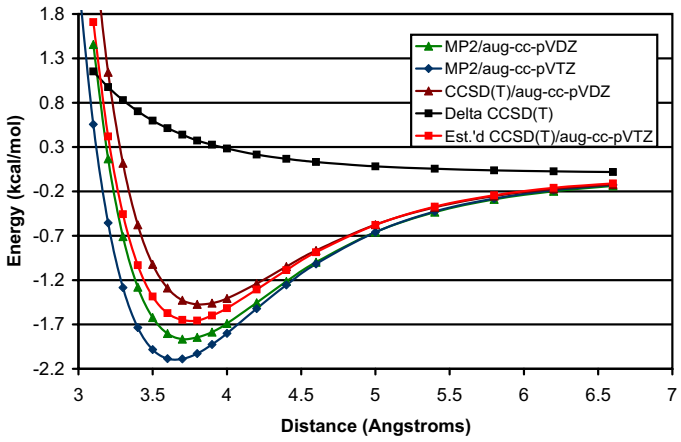


Figure 24: Potential energy curves of the indole-methane complex; configuration **(3a)**: methane centered over the 6-membered aromatic system.

kcal mol^{-1}, with a separation (methane carbon to shared bond) of 3.5 Å. To examine the extent to which this interaction can be considered a sum of two individual C-H/π interactions, the methane-indole complex was divided into a new methane-benzene configuration and a methane-pyrrole complex. The orientation between the methane and the aromatic compound was fixed at the minimum for the methane-indole complex. At the MP2/aug-cc-pVDZ computational level, the total interaction energy for methane-benzene complex (at the indole minimum geometry) was -1.08 kcal mol^{-1} and the methane-pyrrole complex was -0.95 kcal mol^{-1}, giving a total of -2.03 kcal mol^{-1}. At the same computational level and geometry, the methane-indole complex has a total interaction energy of -2.38 kcal mol^{-1}, only slightly larger than the sum of the two separate interactions.

4.3.4 Comparison of complexes

Table 8 shows the equilibrium inter-fragment separation for all five complex configurations determined at several computational levels. In all cases, the (counterpoise-corrected) MP2

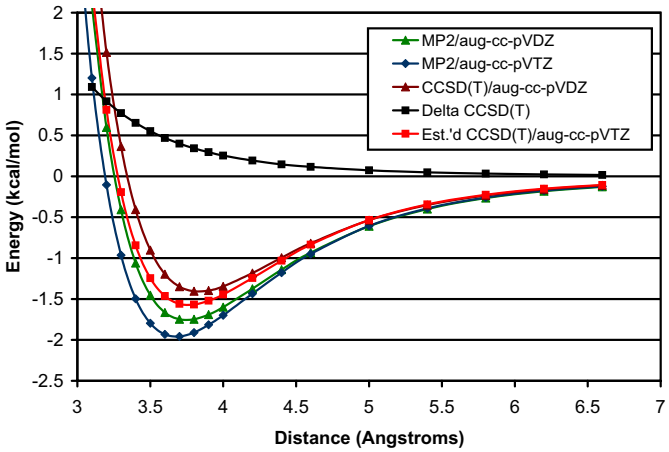


Figure 25: Potential energy curves of the indole-methane complex; configuration **(3b)**: methane centered over the 5-membered aromatic system.

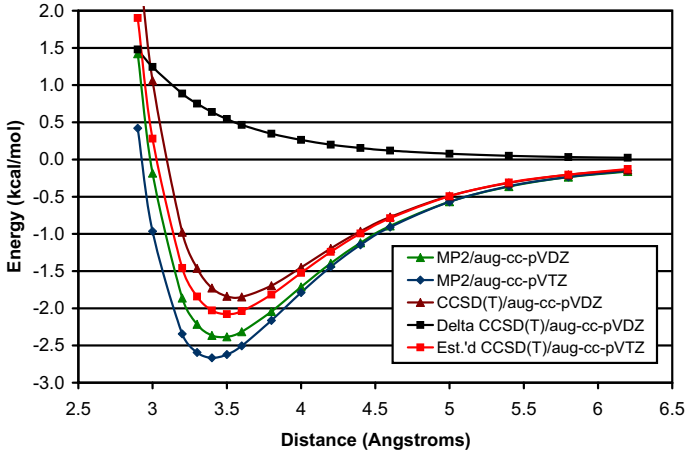


Figure 26: Potential energy curves of the indole-methane complex; configuration **(3c)**: methane is centered over the shared aromatic bond.

Table 8: Equilibrium inter-fragment distances and total interaction energies (in kcal mol^{-1}) for all complex configurations.

	MP2/DZ[a]		MP2/TZ[a]		CCSD(T)/DZ[a]		Est.'d CCSD(T)/TZ[a]	
	R[b]	ΔE_{int}	R[b]	ΔE_{int}	R[b]	ΔE_{int}	R[b]	ΔE_{int}
Methane-Benzene **(1)**	3.8	-1.52	3.7	-1.74	3.9	-1.21	3.8	-1.40
Methane-Phenol **(2)**	3.8	-1.58	3.7	-1.81	3.9	-1.20	3.8	-1.47
Methane-Indole **(3a)**	3.7	-1.87	3.7	-2.09	3.8	-1.47	3.8	-1.66
Methane-Indole **(3b)**	3.7	-1.75	3.7	-1.96	3.8	-1.41	3.8	-1.57
Methane-Indole **(3c)**	3.5	-2.38	3.4	-2.67	3.6	-1.85	3.5	-2.08

[a]Calculations preformed using the aug-cc-pVXZ basis set. [b]Equilibrium monomer separation (using rigid monomers).

interaction energies become more attractive as the basis set is improved from double-ζ to triple-ζ. Comparing the MP2/aug-cc-pVDZ energy to the CCSD(T)/aug-cc-pVDZ results, the more complete description of electron correlation predicts the complexes to be less bound (by about 0.3-0.5 kcal mol^{-1}) and have longer inter-fragment separations (by 0.1 Å). At the estimated CCSD(T)/aug-cc-pVTZ level of theory, the methane-benzene complex is the least bound of all the complexes, with a binding energy of -1.40 kcal mol^{-1}, but the interaction energies for all the configurations which feature one hydrogen down (**1, 2, 3a,** and **3b**) are within 0.20 kcal mol^{-1} of the methane-benzene complex (**1**) at this level of theory. Additionally, all four of these configurations have the same equilibrium inter-fragment separation of 3.8 Å. For these four complexes, the order of increasing stabilization is: **1** < **2** < **3b** < **3a**. At every level of theory considered, the most stabilized complex is the indole-methane complex with one hydrogen directed towards each of the aromatic centers, configuration **3c**.

To provide further insight for the ordering of the configurations, SAPT analysis was performed to divide the total interaction energy into physically significant components. The results of this analysis are presented in Table 9. The similarity of the total interaction energy of the methane-benzene and methane-phenol complexes is reflected in most of the components of the interaction energy. The calculated electrostatic and induction components are almost identical for both complexes, with only slight variances in the exchange and dispersion components. Not surprisingly, the indole-methane complex configuration in

Table 9: Physical components (in kcal mol^{-1}) of total interaction energy determined using SAPT for all complex configurations.

	R	Elst.	Exch.	Ind.	Disp.	SAPT2[a]
Methane-Benzene **(1)**	3.8	-0.898	2.164	-0.255	-2.025	-1.014
Methane-Phenol **(2)**	3.8	-0.898	2.144	-0.254	-2.064	-1.072
Methane-Indole **(3a)**	3.8	-0.893	2.116	-0.291	-2.286	-1.353
Methane-Indole **(3b)**	3.8	-1.165	2.881	-0.344	-2.614	-1.242
Methane-Indole **(3c)**	3.5	-1.349	3.221	-0.334	-3.229	-1.692

[a]All data computed at the cc-pVDZ+ basis using the optimized MP2/aug-cc-pVDZ monomer geometries and the optimum inter-fragment separation as determined by the est.'d CCSD(T)/aug-cc-pVTZ computation.

which one hydrogen is directed towards the six-membered aromatic system (**3a**) also has very similar electrostatic, exchange, and induction contributions. The 0.34 kcal mol^{-1} difference in its total interaction energy (compared to methane-benzene) is primarily due to differing dispersion contributions. However, for the methane-indole complex in which one hydrogen is directed towards the five-membered aromatic system (**3b**), the electrostatic contributions are approximately 0.27 kcal mol^{-1} more stabilizing relative to the other one hydrogen down configurations (**1**, **2**, and **3a**). This is accompanied by a small stabilization (0.09 kcal mol^{-1}) in the induction contribution compared to methane-benzene (**1**). Dispersion is more stabilizing by 0.59 kcal mol^{-1}, but this effect is countered by an additional 0.72 kcal mol^{-1} destabilization in the exchange term.

The most stable of all the complexes considered, the **3c** indole-methane complex, has stabilizing electrostatic (1.35 kcal mol^{-1}) and dispersion (3.23 kcal mol^{-1}) terms which are larger than for any of the other complexes. This configuration has a much shorter equilibrium inter-fragment separation (R=3.5 Å), and shorter separation distances usually lead to more attractive dispersion terms, countered by a larger exchange-repulsion term (in this case 3.22 kcal mol^{-1}, almost completely canceling the dispersion term). The contribution from induction (stabilization of 0.33 kcal mol^{-1}) is similar to that of the other complex configurations considered.

Mulliken population analysis was performed to compare the charge distribution in the

methane-benzene complex versus in the separated monomers. [1] The SCF wavefunction determined using the cc-pVDZ basis set was analyzed (using the population analysis program in MOLPRO [126]) for the methane-benzene complex at an inter-fragment separation of 3.8 Å as well as for the separated complexes at their optimized geometries described above. The most significant difference was found for the charge distribution of methane. For the isolated methane molecule, the hydrogens all had equivalent charges of 0.039 a.u. However, in the methane-benzene complex, the methane hydrogen directed towards the center of the ring took on a greater positive charge (0.078 a.u.) while the other methane hydrogens only had a partial charge of 0.030 a.u. each. These results indicate that the electron distribution in methane polarizes somewhat to reinforce the favorable electrostatic interactions in the complex; this is reflected in the favorable -0.26 kcal mol^{-1} induction term from the SAPT analysis. The population analysis also indicates some transfer of negative electronic charge from methane to benzene, but only a very small amount (0.006 a.u.).

Thus far, the complexes considered have modeled aliphatic C-H/π interactions and have not explored the possibility of aromatic C-H/π contacts, even though these contacts are also prevalent in protein structures [11]. The T-shaped benzene dimer provides a model for such an interaction, in that a hydrogen from the axial benzene interacts with the π cloud of the equatorial benzene. Previous work [103] has determined potential energy surfaces for the T-shaped benzene dimer, using methods similar to those used in this work for the methane-benzene complex. For the T-shaped benzene dimer, the equilibrium C-H/π distance (from the C of the upper benzene to the center of the ring of the lower benzene) is 3.5 Å, and the total interaction determined by adding the MP2/aug-cc-pVTZ energy and a ΔCCSD(T) correction is -2.53 kcal mol^{-1} [103]. Comparing this to the methane-benzene complex at the same computational level, the methane-benzene complex has just over half the binding energy, indicating that a T-shaped benzene dimer may not be as simple as a C-H/π interaction. The results of SAPT analysis of these two systems are shown in Figure 27. Because SAPT analysis is quite dependent on inter-fragment separation, to enable

[1] Although Mulliken analysis can be problematic (e.g. Mulliken charges can be very sensitive to the level of theory), we believe it should suffice for a general discussion of trends.

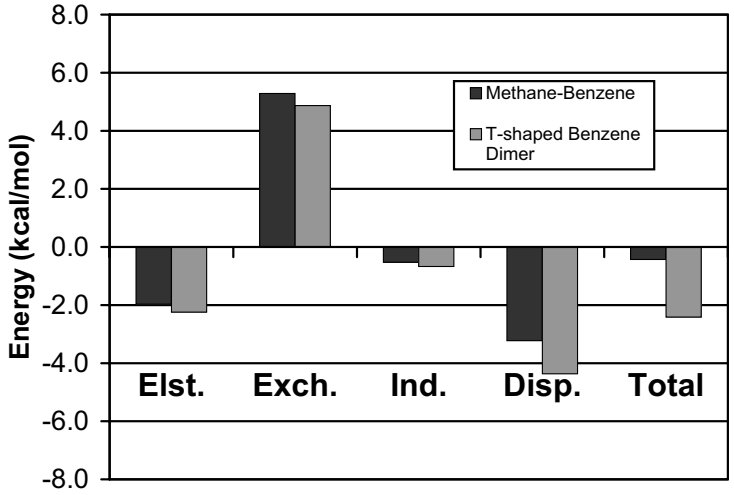


Figure 27: Electrostatic (-1.97, -2.24), exchange-repulsion (5.29, 4.87), induction (-0.53,-0.67), dispersion (-3.22, -4.37), and total interaction energies (-0.43, -2.41) for methane-benzene complex and T-shaped benzene dimer in kcal mol^{-1}; both systems have a CH/π distance of 3.5 Å.

a more direct comparison both monomers were fixed at the T-shaped benzene dimer C-H/π distance of 3.5 Å. The electrostatic, exchange-repulsion, and induction terms for both systems are similar, within 0.5 kcal mol^{-1}. The electrostatic contribution differs by only 0.3 kcal mol^{-1} while the dispersion contributions differ by over 1 kcal mol^{-1}. This suggests that the increased interaction energy of the T-shaped benzene dimer is not primarily caused by the increased acidity of the benzene hydrogen over the methane hydrogen, but rather that an increased dispersion interaction (involving the electrons of the upper π system) and a decreased exchange-repulsion interaction are important in stabilizing the benzene dimer over the methane-benzene complex.

4.4 Conclusions

In this chapter, high-quality potential energy curves were generated for methane-benzene, methane-phenol, and methane-indole complexes as the simplest prototypes noncovalent

C-H/π interactions between protein side-chains. Curves were generated using MP2 and CCSD(T) in conjunction with the aug-cc-pVDZ basis set. By determining the difference between these two curves, the effect of higher electron correlation can be captured in a correction denoted ΔCCSD(T). This correction is then applied to the MP2/aug-cc-pVTZ curve, which gives an accurate estimate of the interaction energy at the robust CCSD(T)/aug-cc-pVTZ level of theory.

For the methane-benzene complex, a two-dimensional potential surface was generated at the CCSD(T)/aug-cc-pVDZ computational level that varied both inter-fragment separation and the angle between the C-H bond of methane and the normal to the plane of benzene. This surface shows that the minimum is found for the configuration in which methane is located directly above the benzene ring. At the best computational level considered in this study, estimated CCSD(T)/CBS, the inter-fragment separation (distance from the methane C to the center of the benzene ring) for the minimum configuration is 3.8 Å and the total interaction energy is -1.454 kcal mol^{-1}. As the inter-fragment separation increases, the preferred angle between the methane carbon and the aromatic ring changes from directly perpendicular to offset. Comparing these results with those from the database study [11] of Brandl et. al., a good correlation between the predicted interaction energies of the potential surface determined in this work and the frequency of C-H/π contacts in crystal structures in the PDB.

The methane-benzene complex is the least bound of the complex configurations considered, but it still lies within 0.20 kcal mol^{-1} of methane-phenol and methane-indole complexes that have similar configurations in which only one hydrogen is directed towards the aromatic system. An indole-methane complex, which features two hydrogens directed towards the aromatic centers, is approximately 0.6 kcal mol^{-1} more stable than the methane-benzene complex. SAPT analysis shows that in complexes where electrostatics are similar (i.e. **1**, **2**, and **3a**), differences in the total interaction energy are caused by differences in the dispersion and induction contributions. SAPT analysis of the methane-benzene complex and the T-shaped benzene dimer indicates that the additional electron density provided by the π system of the upper benzene is important in stabilizing aromatic C-H/π interactions over

aliphatic C-H/π interactions.

The high quality potential energy curves presented here will aid in the analysis of C-H/π interactions in which other steric and geometric constraints prevent equilibrium structures from being attained. This information can also be used to calibrate force fields and to test new density functional theories and other techniques designed to model larger scale systems in which noncovalent interactions are critical.

CHAPTER V

FIRST PRINCIPLES COMPUTATION OF LATTICE ENERGIES OF ORGANIC SOLIDS: THE BENZENE CRYSTAL

5.1 Introduction

Understanding intermolecular interactions is foundational to molecular recognition and crystal engineering. Crystal engineering can capitalize on the understanding of such interactions to design and manipulate crystal properties by making chemical modifications on the molecular level [41, 29, 32]. Often, when crystal structures are predicted for a given molecular structure, multiple stable crystal structures with similar energies can be generated. Accurate computational determination of the lattice energy of such crystal structures would aid in energetically ranking the structures and offer the ability to select structures of a particular energy. Such computations and energy rankings could help identify competing low-energy structures which might complicate synthesis and production of pharmaceutical products [29] and could also aid in the prediction of the resolution behavior of racemic mixtures, by providing a method to rank the energy of the mixture vs. single enantiomer crystal structures [43].

Interest in calculating the lattice energy of crystalline benzene can be found as early as 1966 [5]. Calculations of the lattice energy have generally proceeded by using atom-atom potentials, with parameters fit to experimental observations. Recent work on drug crystals suggests that the lattice energy can be quite sensitive to the chosen parameters [64]; moreover, the need to fit to experimental data to deduce many different atom-atom potentials makes it harder to apply these approaches to a wide variety of systems. Recently, methods have been proposed to take into account intermolecular interactions, rather than just simpler atom-atom interactions [33]. However, the accuracy of these methods is still governed by the quality of the intermolecular parameters and the flexibility of the assumed functional form; when simple model potentials are used, the global minimum crystal structure

is predicted only about a third of the time [29].

Non-empirical models are preferred for their wider applicability and their potential for yielding more accurate results. One successful non-empirical approach is the PIXEL method [38, 39, 40, 32], which is based on the determination of molecular densities and using this information to determine the different physical contributions (coulombic, polarization, dispersion, and repulsion) to the intermolecular interaction energies. Results from the Pixel method have been compared to some first-principles electronic structure calculations which include electron correlation and can preform comparably to second-order perturbation theory (MP2) [40].

To investigate methods which would remove the dependence of lattice energy determination on empirical parameters, Schweizer and Dunitz [95] performed *ab initio* MP2 electronic structure calculations to determine the lattice energy of crystalline benzene and compared these results to those from the Pixel method. The benzene dimer and the methods required to achieve converged results for its interaction energy have been the subject of significant computational effort [105, 106, 53, 49, 117, 116, 115, 119, 48], but using correlated electronic structure methods to determine the lattice energy of crystalline benzene was largely unexplored. Schweizer and Dunitz proposed an additive scheme in which the interaction energy of only the four unique symmetry-related nearest-neighbor dimers is determined. MP2 greatly overestimated the interaction energy of the dimers, and the overestimation grew worse with increasing the size of the basis set and did not provide convergent results. Counterpoise corrections lowered the interaction energy to less than half of the uncorrected MP2 values, indicating that the largest basis set employed, 6-31++G(d,p), is not nearly large enough to approach basis set convergence. These findings are in agreement with other studies documenting the need to use coupled-cluster methods in conjunction with very large basis sets to achieve reliable results for noncovalent interactions between aromatic molecules [106, 103, 115, 118, 49]. In contrast, the Pixel energies converged towards a value of 42.1 kJ mol^{-1} (incorrectly given as 43.8 kJ mol^{-1} in the paper by Schweizer and Dunitz [95] because of an arithmetical error) for the estimated lattice energy with increasing basis set size.

Other than this recent work, ab initio determinations of the lattice energies of crystals have primarily been limited to Hartree-Fock and density functional methods (DFT), which do not always give qualitatively similar results for the lattice parameters and bond distances of the crystals when compared to experimental values [19]. Recent work which adds an empirical van der Waals correction to DFT has shown an improvement in the determination of unit cell parameters (although lattice energies were not reported) [75]. The only ab initio determination of a crystal energy which used highly correlated electronic structure methods such as coupled-cluster theory through perturbative triples [CCSD(T)] [88] computed the electron correlation energies of a series of small LiH crystals and determined the cohesive energy of the crystal by extrapolating these results [67].

In this work, state-of-the-art quantum mechanical methods are used to determine the lattice energy of crystalline benzene with high accuracy. Specifically, the CCSD(T) and MP2 methods are utilized in conjunction with very large basis sets to obtain dimer energies which should provide accurate dimer binding energies to within a few tenths of a kcal mol^{-1} [105], to enable a more accurate determination of the lattice energy of crystalline benzene using the additive system of Schweizer and Dunitz. Going beyond their model, we also consider the effects of including longer-range dimer interactions, as well as three-body interactions among nearest-neighbor trimers. To compare our calculated lattice energy to experimental estimates for the heat of sublimation, corrections must also be included to account for the enthalpy change of the crystal from 0 K to the measurement temperature of the sublimation energy (around 250 K) as well as a zero-point vibrational energy correction to account for the lattice mode vibrations of the crystal. By making these comparisons, it can be demonstrated that state-of-the-art quantum chemistry is capable of computing the lattice energy of organic crystals like benzene to a high accuracy and to provide a definitive methodology for obtaining converged results for the lattice energy of neutral organic crystals.

5.2 Computational details

The coordinates for the benzene dimers were taken from the neutron diffraction crystal structure of Bacon et al. and were not otherwise optimized. These are the coordinates used in recent work by Schweizer and Dunitz [95], and the same coordinates were used in this study for consistency and to enable comparison to the methods used in their study. A more recent neutron diffraction study by Jeffery et al. [54] determined a very accurate structure for deuterated benzene, which reported very similar mean bond lengths to the structure of Bacon et al., but with much smaller uncertainties. However, using the coordinates of this improved structure would have made very little difference in the computation of the lattice energy, as the interaction energy of a typical dimer differs by only 0.01 kcal mol^{-1} for the two structures.

For each dimer, the total counterpoise-corrected interaction energy was determined by MP2 in conjunction with the correlation consistent basis sets augmented with diffuse functions, aug-cc-pVXZ (where X = D, T, and Q), and CCSD(T) with the aug-cc-pVDZ basis set. From the aug-cc-pVTZ and aug-cc-pVQZ results, the MP2 correlation energy was extrapolated to the complete basis set (CBS) limit using the procedure of Halkier et al. [45]. This extrapolation procedure should almost entirely eliminate any basis set incompleteness error from the determination of the dimer interaction energies and thus the lattice energy. To account for additional electron correlation, the counterpoise-corrected CCSD(T) interaction energy was determined using the aug-cc-pVDZ basis set, and a correlation correction term was determined as the difference between the MP2 and CCSD(T) energies determined in the aug-cc-pVDZ basis. This change, denoted ΔCCSD(T), is then added to the MP2/CBS results, giving an estimated CCSD(T)/CBS interaction energy. Previous work [103] indicates that the ΔCCSD(T) correction term is quite insensitive to basis set effect, so that ΔCCSD(T) corrections are probably converged within a few hundredths of a kcal mol^{-1} when the aug-cc-pVDZ basis is used. All computations were performed using MOLPRO [126].

Table 10: Interacting dimer pairs in crystalline benzene.

pair	symmetry operation	N[a]	R[b]
A	a/c glide reflection	4	5.02
B	c/b glide reflection	4	5.81
C	b/a glide reflection	4	5.99
D	±c translation	2	6.81
E	±a translation	2	7.39
F	±b translation	2	9.42
G	±c translation and b/a glide reflection	8	9.07
H	±a translation and c/b glide reflection	8	9.40

[a] Number of symmetry-related pairs involving a given reference. [b] Distance (in Å) between the centers of mass of the two molecules.

Table 11: Interaction energies (in kJ mol^{-1}) for interacting dimers in the first coordination sphere and lattice energy contributions at several computational levels.

pair	N[a]	MP2/DZ[b]	CCSD(T)/DZ[b]	MP2/TZ[b]	MP2/QZ[b]	MP2/CBS	ΔCCSD(T)	Est.'d CCS
A	4	-12.9	-9.8	-14.0	-14.4	-14.6	3.1	-1
B	4	-8.1	-6.6	-8.7	-8.9	-9.1	1.5	-7
C	4	-6.4	-5.2	-6.9	-7.0	-7.1	1.1	-6
D	2	-2.4	-1.9	-2.4	-2.4	-2.4	0.4	-1
Lattice energy contribution		-57.1	-45.1	-61.6	-63.0	-64.0		-5

[a] Number of symmetry-related pairs involving a given reference molecule. [b] Calculations preformed using the aug-cc-pVXZ basis set.

5.3 Results and discussion

5.3.1 Lattice energy determination

The symmetry-related dimers used in the lattice energy determination are described in Table 10, including the distance between the centers of mass of the two benzene molecules. The total interaction energies of each of the symmetry-related dimers (**A, B, C,** and **D**; see Figure 28) of the first coordination sphere are given in Table 11. These dimers are produced by glide-reflection symmetry operations (**A, B,** and **C**) and the c translation operation (**D**). The CCSD(T)/CBS estimate for the lattice energy contribution from these dimers is -52.1 kJ mol^{-1}. By comparison, the computationally inexpensive Pixel method, using MP2/6-31++G(d,p) densities, provides a reasonably good (given the computational cost) estimate of -43.8 kJ mol^{-1} [95].

Around 90 percent of the lattice energy comes from the contributions of these dimers, but

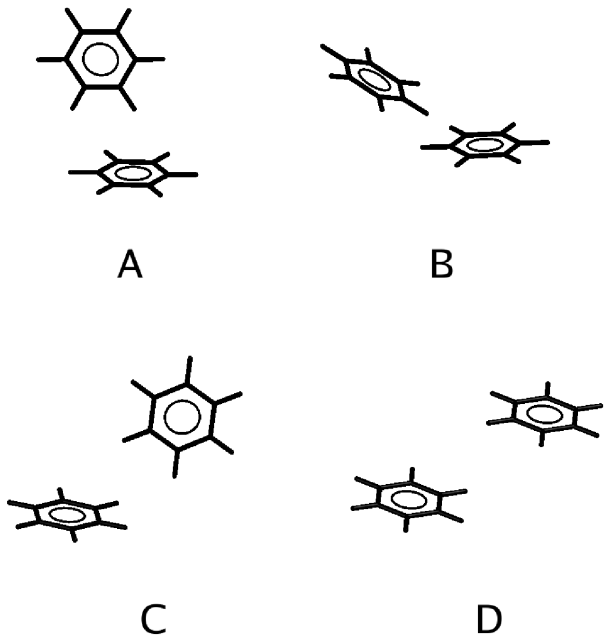


Figure 28: Dimer interactions in the first coordination sphere.

smaller contributions result from interactions outside this first coordination sphere. The largest of these smaller contributions (those that have a total interaction energy greater than 0.25 kJ mol^{-1} at the MP2/aug-cc-pVDZ computational level, and an inter-monomer separation of less than 9.5 Å) come from the dimers produced by the a and b translations, the c translation followed by the b/a glide-reflection, and the a translation followed by the c/b glide-reflection. The interaction energies for these dimers (**E, F, G,** and **H**, respectively; see Figure 29) and their contribution to the lattice energy are summarized in Table 12.

Given that the majority of the lattice energy comes from the interaction energy of the symmetry-related dimers in the first coordination sphere, one might also consider the contributions of the three-body interactions within the first coordination sphere. In the study of Tauer et al. [111], the authors found that the cyclic benzene trimer had a three-body contribution to the interaction energy of over 1 kJ mol^{-1}. However, when the cyclic trimers which would be found in the first coordination sphere for the benzene crystal were considered, the three-body effect was always less than 0.1 kJ mol^{-1}. The benzenes in

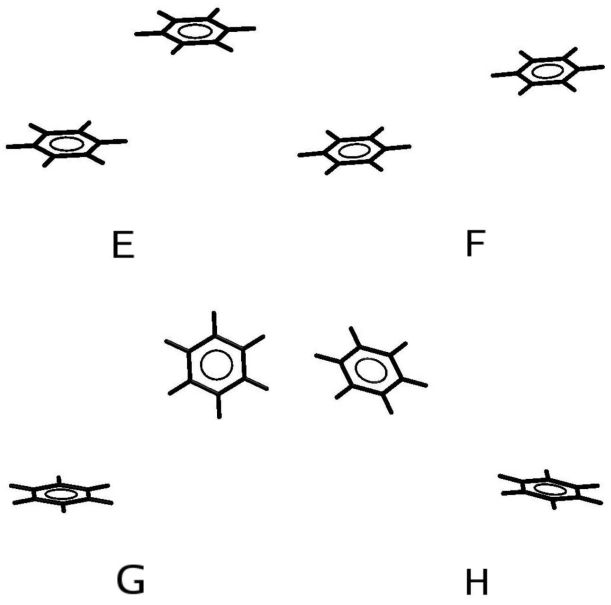


Figure 29: Important dimer interactions beyond the first coordination sphere.

Table 12: Interaction energies (in kJ mol^{-1}) for selected interacting dimers beyond the first coordination sphere lattice energy contributions at several computational levels.

pair	N[a]	MP2/DZ[b]	CCSD(T)/DZ[b]
E	2	-1.4	-1.2
F	2	-0.3	-0.3
G	8	-0.5	-0.4
H	8	-0.4	-0.3
Lattice energy contribution		-5.2	-4.3

[a] Number of symmetry-related pairs involving a given reference molecule. [b] Calculations preformed using the aug-cc-pVXZ basis set.

the crystal are further apart than in the gas-phase configurations of Tauer et al., and the three-body contribution diminishes rapidly with increasing inter-monomer separations.

The total lattice energy is obtained by multiplying the best estimate of the interaction energy for each dimer by the number of symmetry-related pairs involving a given reference molecule (these multiplicities are given in Table 10), summing these products and dividing by 2 (as a result of the counting method [95]). Using the four dimers from the first coordination sphere (**A, B, C,** and **D**; estimated CCSD(T)/CBS results) and the four energetically significant dimers from the second coordination sphere (**E, F, G,** and **H**; CCSD(T)/aug-cc-pVDZ results), our best estimate for the lattice energy of the benzene crystal is -56.4 kJ mol^{-1}.

5.3.2 Enthalpy corrections

To compare the calculated lattice energy to experimental values for the sublimation energy, corrections must be included for the enthalpy changes that would occur between the temperature of the gas phase calculations (0 K) and the measurement temperature of the sublimation energy (around 250 K). The sublimation energy is $\Delta H_{sub} = H_{vapor}$ - $H_{crystal}$. The enthalpies of both phases include the intramolecular electronic energy of the benzene monomers as well as intramolecular vibrational energy contributions. In the gas phase, there are additional translational and rotational enthalpy contributions. In the crystalline phase, there are additional intermolecular (lattice) enthalpy contributions: namely, the intermolecular electronic energy (lattice energy), the zero-point energy of the lattice vibrations, and the finite-temperature (T > 0) contribution of the lattice vibrations.

If the monomer geometry of the benzene molecules were identical in both phases, the intramolecular electronic energy of one mole of benzene molecules would be the same in both phases and would therefore cancel in the computation of the sublimation energy. Jeffrey et al. [54] report a slight deformation from D_{6h} to C_{3v} symmetry in crystalline benzene, and so to examine the effect of this small distortion, the molar intramolecular electronic energy for the 15 K neutron diffraction structure of Jeffrey et al. (the most precise crystal structure

taken for crystalline benzene) was determined and compared to molar intramolecular electronic energy for gas phase benzene. There are inherent difficulties in comparing structures from neutron diffraction studies to gas phase studies due to differences in the quantities measured. The bond lengths determined by neutron diffraction bond lengths are inferred from the difference between the average nuclear positions of the atoms and should thus be compared to similar measurements for the gas phase (r_z values). The r_z values for benzene have been determined experimentally [109] and theoretically [37], with good agreement. Using the r_z values for the C–C and C–H bond lengths of Tamgagawa et al. (1.3976 Å and 1.085 Å, respectively) for the gas phase determination of the molar intramolecular electronic energy of benzene, both phases agree within 0.01 kcal mol^{-1} at the CCSD(T)/aug-cc-pVDZ level. It is also assumed that the internal intramolecular vibrational frequencies are nearly the same in the gas and solid phase, so the enthalpy contribution due to intramolecular vibrations would also cancel in both terms. (The validity of this assumption will be discussed below.) Neglecting the quantities which appear in both phases, the sublimation energy is now given as

$$\Delta H_{sub} = H_{vapor,\ trans\&rot} - (\text{lattice energy} + \text{ZPVE}_{solid,lattice} + H^{finite-T}_{solid,lattice}).$$

The molar enthalpy corrections to the vapor and solid phases can be easily estimated if one assumes that the temperature is sufficient that equipartition of energy applies. In the vapor phase, the translational motions contribute $\frac{5}{2}$RT to the enthalpy correction and the rotational degrees of freedom contribute $\frac{3}{2}$RT, giving a total of 4RT, or 8.3 kJ^{-1} mol^{-1} at 250 K. For the solid phase, the finite temperature enthalpy correction is 6RT (by the the Dulong-Petit approximation), or 12.5 kJ^{-1} mol^{-1} at 250 K. The zero-point contribution of the lattice vibrational modes which would only be present in the solid phase (the $\text{ZPVE}_{solid,lattice}$ term) is described by Nakamura and Miyazawa [74], who calculated the lattice vibrational frequencies for the benzene crystal structure determined by Bacon et al. [4] and determined the frequency distribution of the vibrations. From this distribution, the zero-point energy correction to the sublimation energy was determined as 0.67 kcal mol^{-1}, or 2.8 kJ mol^{-1}. Substituting these values, along with the best estimation of the lattice energy, into the

Table 13: Estimation of the sublimation energy for crystalline benzene.

Calculated lattice energy[a]	
Contribution from the first coordination sphere	
Estimated CCSD(T)/CBS results	-52.1
Contributions for selected dimers	
beyond the first coordination sphere	
CCSD(T)/aug-cc-pVDZ results	-4.3
Total Calculated Lattice Energy	-56.4
Vapor Phase Enthalpy Correction	8.3
Solid Phase Enthalpy Correction	12.5
$ZPVE_{lattice}$	2.8
Sublimation energy	**49.4**
Typical Experimental Values	43-47

[a] All values in kJ mol^{-1}.

equation above gives the best estimate of the sublimation energy, 49.4 kJ mol^{-1}. The results are summarized in Table 13.

To verify that the internal intramolecular vibrational modes of the molecules contribute nearly equivalently to both phases, one could estimate the finite temperature component of the enthalpy of the solid phase and compare it to the 6RT approximation used above. Using heat capacities determined experimentally or with more complete theoretical estimates such as the Debye function, the total finite temperature enthalpy of the solid can be determined by integrating the heat capacity over the appropriate temperature range. Lord and coworkers made such estimates of the heat capacity using the Debye function and calculated the heat capacity of crystalline benzene at 32 discrete temperatures in the range of 0 to 270 K [87]. Using these estimates for the heat capacity, the enthalpy of the crystal is estimated as 14.7 kJ mol^{-1}. The difference between this value and the 6RT estimate (2.2 kJ mol^{-1}) is the finite temperature contribution to the enthalpy from the intramolecular vibrational modes in the solid phase. However, if the total finite temperature enthalpy of the crystal had been used for the solid phase, then the finite temperature correction to the vibrational enthalpy from the intramolecular vibrational modes would have to have been included in the vapor phase as well. This correction would be determined using the vibrational frequencies of a benzene molecule and the usual harmonic oscillator partition function. Using

the frequencies reported by Paige et al. [80], this contribution is 2.3 kJ mol^{-1} at 250 K, almost exactly canceling the difference between the 6RT estimate and the more complete estimation of the T > 0 part of the enthalpy of the solid phase. This indicates that the intramolecular vibrations are, in fact, extremely similar in both phases and if their effects were included, they would appear in both terms and simply cancel.

5.3.3 Comparison to experiment and error analysis

Values for the benzene sublimation energy have been reported from 38.0 to 53.9 kJ mol^{-1} [18] [1], with the majority of values in the 43-47 kJ mol^{-1} range, slightly below the calculated value in this study. Even though the pair interaction energies are each converged to within a few hundredths of a kcal mol^{-1}, this error accumulates in the summation of the lattice energy. The largest sources of error in the dimer interaction energies are basis set incompleteness, higher-order electron correlation, and correlation of core electrons.

The aug-cc-pVTZ/aug-cc-pVQZ extrapolation of the MP2 correlation energy should provide results nearly converged to the CBS limit and nearly eliminate errors associated with the incompleteness of the one-particle basis set. The approximate size of any remaining basis set error can be estimated by comparing the MP2/aug-cc-pVQZ and MP2/CBS interaction energies for the dimers in the first coordination sphere (**A, B, C,** and **D**). The difference in these interaction energies is the largest for **A** and is 0.25 kJ mol^{-1}. The remaining basis set error in the interaction energy is very likely less than this value. Additionally, the basis set error diminishes rapidly as the inter-monomer separation increases and for **D**, the difference in the MP2/aug-cc-pVQZ and MP2/CBS estimates for the interaction energy is less than 0.01 kJ mol^{-1} and is completely negligible for **E, F, G,** and **H**. Estimating the remaining basis set error as the difference between the MP2/aug-cc-pVQZ and MP2/CBS interaction energies for the dimers involved in the first coordination sphere and propagating this error in the lattice energy calculation, the error introduced to the lattice energy from remaining basis set incompleteness is at most -1.0 kJ mol^{-1}.

The importance of higher-order electron correlation is evident by the size of the ΔCCSD(T)

[1] For an online compilation of sublimation energies, see the NIST webbok at http://webbook.nist.gov/cgi/cbook.cgi?ID=C71432&Units=SI&Mask=4 and references given there.

correction used in the determination of the lattice energy, and the contribution of the triple excitations is essential in accurately determining the interaction energy of noncovalent systems. Given the importance of the triple excitations, it is certainly possible that even higher-order contributions to the electron correlation, such as quadruple or pentuple excitations, may make small, but not insignificant contributions to the interaction energies. Hopkins and Tschumper investigated the importance of quadruple excitations on the interaction energy of several small dimers [50]. For their test set of π-π interacting dimers, the contribution of the quadruple excitations is between 5% and 27% that of the triple excitations. Using this guideline, the contribution of the quadruple excitations for each of the eight dimers was taken as these percentages of the difference between the CCSD(T)/aug-cc-pVDZ and CCSD/aug-cc-pVDZ interaction energies. (While there are certainly contributions from pentuple excitations and beyond, they will be dwarfed by the contributions from the quadruple excitations.) From these estimations, an error in the lattice energy due to contributions from higher-order electron correlation was obtained which was 0.6 (using the 5% estimate) to 3.2 kJ mol^{-1} (using the 27% estimate).

In all the computations using the aug-cc-pVXZ basis sets, all the core orbitals were doubly occupied; that is, the frozen-core approximation was utilized. To estimate the effect of removing this restriction, MP2 interaction energies were determined using the double-ζ core-valence basis set, aug-cc-pCVDZ [131], for dimer **A**. The interaction energy was determined within this basis set using the frozen-core approximation and again allowing the core electrons to be correlated. The difference between these two interaction energies was only 0.06 kJ mol^{-1}. As discussed above in regards to basis set incompleteness, as the total interaction energy of the complex decreases, so does the size of the error associated with the dimer's interaction energy. Since dimer **A** has the greatest interaction energy of the dimers considered and the core correlation energy for the remaining dimers will be even smaller for the other dimers, the core correlation is likely not a significant source of error in the determination of the interaction energies of these systems and should not introduce a sizable error in the determination of the lattice energy.

Any additional sources of error, such as errors introduced by the Born-Oppenheimer

approximation or relativistic effects, are much smaller than the sources of errors just discussed. The effects of higher-order electron correlation are certainly the largest source of error and give a less bound estimate of the lattice energy and a lower value for the sublimation energy. These effects would be partially countered by the error due to basis set incompleteness, which would give a more bound estimate of the lattice energy and a larger sublimation energy. Using these estimates of -1.0 to 0 kJ mol^{-1} for the basis set incompleteness error and 0.6 to 3.2 kJ mol^{-1} for the higher-order correlation error, estimates of -53.2 to -56.8 kJ mol^{-1} for the lattice energy, or 46.2 to 49.8 kJ mol^{-1} for the sublimation energy were obtained.

Taking into account the error analysis of the calculated sublimation energy, this estimate of the sublimation energy for benzene is likely within "chemical accuracy" (within 1 kcal mol^{-1}) of typical experimental values. Using, for instance, the most recent value included in the NIST compilation as a benchmark (45.2 kJ mol^{-1}) [107], a "chemical accuracy" estimate could range from 41.0 kJ mol^{-1} to 49.4 kJ mol^{-1}, encompassing almost all the entire range of theoretical values predicted by this study. The computational rigor of the methods required to achieve this result underscores the need to use highly converged electronic structure methods to make high-accuracy ab initio determinations of sublimation energies.

5.4 Conclusions

The lattice energy of crystalline benzene has been determined using highly correlated electronic structure methods and large augmented basis sets and has been extrapolated to the CCSD(T) complete basis set limit. This work extends previous work on ab initio lattice energy determination in several important ways. Correlated methods beyond second-order perturbation theory have been used to more accurately determine the interaction energy of the dimeric interactions involved in the first coordination sphere for a reference benzene molecule. The size of the three-body interactions in the first coordination sphere have been investigated and show that these interactions likely make negligible contributions to the lattice energy. However, longer range dimeric interactions (beyond the first coordination sphere) account for almost 10% of the total lattice energy and should not be neglected if

one hopes to make a high-accuracy determination of the lattice energy.

Using converged methods is particularly important since even small systematic errors (on the order of a few hundredths of a kcal mol^{-1} in these calculations) result in larger errors in the determination of the lattice energy because they accumulate in the addition of all the pair energies. Sources of such systematic errors were discussed and estimates for the sizes of these errors were included to estimate the error bars on the calculated sublimation energy. Including enthalpy corrections, the sublimation energy of benzene was estimated to be 46.2-49.8 mol^{-1} (with a best estimate of 49.4 kJ mol^{-1}), compared to typical experimental values of 43-47 kJ mol^{-1}. These computations demonstrate that the lattice energy can be accurately determined (to around 1 kcal mol^{-1}) for neutral organic molecular crystals using converged ab initio electronic structure methods and establish a general methodology to make such high-accuracy determinations.

The highly accurate determination of lattice energies provides a new tool for the crystal engineer to energetically rank and compare competing crystal structures. Furthermore, it should be possible to directly obtain the most thermodynamically stable crystal structures by minimizing the lattice energy with respect to the crystal geometry, employing techniques described here or judicious approximations of them. Clearly the ability to predict the structures and energetics of crystals to a high degree of accuracy would be of great utility in crystal design.

CHAPTER VI

CONCLUSIONS AND OUTLOOK

In this thesis, I have discussed how noncovalent interactions in complex chemical systems can be studied using small model systems with correlated electronic structure theory. The overarching goal of the work was to demonstrate that energetic and structural insight gained from computations on small models systems has relevance and predictive capability in large systems. The use of small model systems enables the isolation and examination of specific noncovalent interactions such that they can be characterized independently of other environmental effects of the system, but this approach must be complemented by comparisons to large chemical systems to insure validity of the model systems. Several types of noncovalent interactions have been examined in a variety of biophysical and chemical systems.

6.1 Substituent effects in π stacking

6.1.1 Major findings

Substituent effects in π stacking were explored by making a series of substitutions on a benzene dimer complex to ascertain the effect of multiple substituents on π stacking. For sandwich (fully co-facial) configurations, the substituent effect was found to be linearly additive. For T-shaped configurations, a more complex model was required which took into account the effect of the substitution on the electrostatic and dispersion components of the interaction energy, as well as direct interactions between the substituent and the other aromatic ring. The additivity of substituent effects in sandwich configurations also counters assertions that substituent effects are governed solely by electrostatic control, as the differential dispersion contributions accumulate with multiple substituents and give molecules with very different electrostatic potentials very similar attractions to benzene molecules.

6.1.2 Outlook

Substituent effects in aromatic interactions continue to be a topic of significant experimental and theoretical investigation. Ongoing work in the area includes examining to what extent the electron density in the π cloud is modified by the substituents and to what extent substituent effects are actually just direct interactions between the substituents and the other molecules in the complex. In this work, the importance of direct substituent interactions in predicting interaction energies was highlighted, and additional work in this area suggets that the influence of these effects may extremely significant in sandwich configurations [127]. Additional work to examine the effect of direct substituent interactions for the many possible configurations of arene-arene interactions is needed.

6.2 S/π interactions

6.2.1 Major findings

The optimum geometry of sulfur/aromatic interactions in protein structures was evaluated using the H_2S-benzene complex. A constrained potential energy surface scan was performed using a lower-level computational method to identify local minima on the potential energy surface. From these approximate values, high-level coupled-cluster calculations were utilized to determine accurate interaction energies for the complexes. The optimum geometry for a S/π interaction occurs when the sulfur is directly above the center of the aromatic ring at a separation distance of 3.8 Å. The optimum in-plane geometry is found at a longer separation of 5.5 Å. To validate the use of this particular model system as a general description of S/π, over 700 structures from the PDB were then analyzed, and the the interaction geometry for any S atom located with 12 angstroms of the center of an aromatic systems was determined. A correspondence was found between the geometries predicted by the quantum mechanical calculations and the interaction geometries that appeared most frequently in the PDB analysis.

6.2.2 Outlook

The potential energy surface of the H_2S benzene complex is very flat, particularly in the angular space centered around the point directly above the center of the aromatic ring. Thus, the global minimum configuration can be difficult to identify because a full gradient-based optimization procedure would minimize extremely slowly. One potential solution to this problem is to perform the optimization in a set of internal coordinates which explicity included so called "interfragment" coordinates that connected the fragments in the complex, rather than only the internal coordinates which would be connected by covalent bonds. Such a procedure would aid in the indentification of true stationary points of the potential energy surface which would enable an accurate frequency analysis on noncovalent complexes.

6.3 C-H/π interactions

6.3.1 Major findings

C-H/π interactions were examined through several complexes which involved methane and aromatic compounds (benzene, phenol, and indole) which represented the aromatic amino acids. The fundamental C-H/π interaction was found to be relatively insensitive to the type of aromatic ring involved in the interaction and thus, a general C-H/π interaction could be modeled with a five- or six-membered aromatic ring. The general C-H/π interaction has an preferred interaction distance at around 3.8 Åwith an interaction energy of -1.4 to -1.6 kcal mol^{-1}.

6.3.2 Outlook

The general strategy of using quantum mechanical computations for model systems to identify optimal interaction geometries for noncovalent interactions and then analyzing crystal structures to determine if these interactions are seen with greater frequency could easily be applied to C-H/π type systems. In fact, the existing code created for the S/π PDB analysis is purposefully structured to enable easy expansion such than other types of interactions could be considered. Some limited database analyses have been reported in the literature to examine C-H/π interactions (and were discussed in Chapter 4), but the majority of these

analyses were flawed in their normalization procedure or scope, so additional work in this area is certainly still viable.

6.4 Lattice energy determination for small neutral organic crystals

6.4.1 Major findings

This work has established a definitive methodology to computationally determine the lattice energy of small, neutral organic crystals using an additive scheme of interaction energies of individual interacting dimers within the crystal. This enables the accurate determination of energetic information from structural information about the crystal. While three-body effects do not contribute significantly to the overall lattice energy of the crystal, some rather distant two-body interactions can make significant contributions. To compare calculated lattice energies to experimentally determined sublimation energies, enthalpy corrections that take into account the translational and rotational enthalpy of the vapor and the zero-point energy of the lattice vibrations must be included. With such corrections, a "chemically accurate" (within 1 kcal mol^{-1}) determination of the sublimation energy of crystalline benzene has been made.

6.4.2 Outlook

Other investigations which explored computational methods to determine the lattice energy of crystals have called into question the negligence of three-body effects in determining the overall lattice energy of a crystal [84, 83]. The magnitude of these interactions could be better estimated by computational methods which include higher-order electron correlation than were used in the evaluation of the three-body and would give a better idea of the error incurred by neglecting these interactions. Additionally, even if the magnitude of any individual three-body interaction is quite small, the multiplicity of these interactions in the lattice energy summation should also be considered in determining the overall importance of three-body interactions.

Perhaps the ultimate goal to utilize lattice energy computations in crystal engineering would be to predict the crystal structure of a material starting from only the molecular structure of the monomers in the material thorough an optimization of the lattice energy.

Such an undertaking would involve significant scientific and technical progress forward from this work, as an automated, efficient way to determine the individual interactions and sum them into a lattice energy would have to be developed. Additionally, a method to minimize some type of unified function which described all types of interactions in the crystal would be needed to reconcile how to adjust multiple coupled parameters during the optimization procedure. At the present time, the state of the art in crystal structure optimization uses periodic codes which rely on simple (often Lennard-Jones type) functional forms to describe the interactions between the individual molecules of the crystal. This work has shown that such simple descriptions of the interaction are not accurate, since highly correlated electronic structure methods were required to make a chemically accurate determination of the lattice energy. If such simple potentials are to be used, they could be improved by reparametrization to accurate quantum mechanical data such as that presented in this work.

6.5 Computational discovery from small to big

Computational simulations have become accepted as the "third mode of discovery, along with experimentation and theory" in the advancement of scientific knowledge and engineering practice (from the Strategic Plan of the Office of Science, U.S. Department of Energy, 1999). Computational methods used as a discovery mode in chemistry can span a full range of system types, sizes, and conditions. The work in this thesis has capitalized on theory and computation to gain unique insight about noncovalent interactions in chemical systems in a way that bridges accurate computational methods used to characterize small systems to large scale chemical systems. This "small to big" methodology is the framework for bottom-up development of chemical systems and is of ever increasing importance in the development of molecular engineering. Creating this bridge is a vital step to understanding how molecular properties can be utilized to solve chemical problems.

REFERENCES

[1] Adams, H., Carver, F. J., Hunter, C. A., Morales, J. C., and Seward, E. M., "Chemical double-mutant cycles for the measurement of weak intermolecular interactions: Edge-to-face aromatic interactions," *Angew. Chem. Int. Ed. Engl.*, vol. 35, no. 13/14, pp. 1542–1544, 1996.

[2] Amicangelo, J. C., Gung, B. W., Irwin, D. G., and Romano, N. C., "Ab initio study of substituent effects in the interactions of dimethyl ether with aromatic rings," *Phys. Chem. Chem. Phys.*, vol. 10, pp. 2695–2705, 2008.

[3] Arunan, 2006. E. Arunan, personal communication.

[4] Bacon, G. E., Curry, N. A., and Wilson, S. A., "Crystallographic study of solid benzene by neutron diffraction," *Proc. R. Soc. London Ser. A*, vol. 279, pp. 98–100, 1964.

[5] Banerjee, K. and Salem, L., "Forces in benzene crystal. i. lattice energy of crystalline benzene," *Mol. Phys.*, vol. 11, pp. 405–420, 1966.

[6] Bartlett, R. J., "Many-body perturbation theory and coupled cluster theory for electron correlation in molecules," *Ann. Rev. Phys. Chem.*, vol. 32, pp. 359–401, 1981.

[7] Bartlett, R. J. and Purvis, G. D., "Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem," *Int. J. Quantum Chem.*, vol. 14, pp. 561–581, 1978.

[8] Becke, A. D. and Johnson, E. R., "A density-functional model of the dispersion interaction," *J. Chem. Phys.*, vol. 123, p. 154101, 2005.

[9] Beg, S., Waggoner, K., Ahmad, Y., Watt, M., and Lewis, M., "Predicting face-to-face arene-arene binding energies," *Chem. Phys. Lett.*, vol. 455, pp. 98–102, 2008.

[10] Boys, S. F. and Bernardi, F., "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors," *Mol. Phys.*, vol. 19, no. 4, pp. 553–566, 1970.

[11] Brandl, M., Weiss, M. S., Jabs, A., Suhnel, J., and Hilgenfeld, R., "Ch/π interactions in proteins," *J. Mol. Biol.*, vol. 307, pp. 357–377, 2001.

[12] 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., and 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. See: http://www.physics.udel.edu/∼szalewic/SAPT/SAPT.html.

[13] Burley, S. K. and Petsko, G. A., "Aromatic-aromatic interaction: A mechanism of protein structure stabilization," *Science*, vol. 229, pp. 23–28, 1985.

[14] Carroll, W. R., Pellechia, P., and Shimizu, K. D., "A rigid molecular balance for measuring face-to-face arene-arene interactions," *Org. Lett.*, vol. 10, pp. 3547–3550, 2008.

[15] Carter, P. J., Winter, G., and Fersht, A. R., "The use of double mutants to detect structural changes in the active site of the tyrosyl-transfer rns-synthetase (bacillus-stearothermophilus)," *Cell*, vol. 38, pp. 835–840, 1984.

[16] Carver, F. J., Hunter, C. A., Livingstone, D. J., McCabe, J. F., and Seward, E. M., "Substituent effects on edge-to-face aromatic interactions," *Chem. Eur. J.*, vol. 8, no. 13, pp. 2848–2859, 2002.

[17] Cheney, B. V., Schulz, M. W., and Cheney, J., "Complexes of benzene with formamide and methanethiol as models for interactions of protein substructures," *Biochim. Biophys. Acta.*, vol. 996, pp. 116–124, 1989.

[18] Chickos, W. E. and Jr., W. E. A., "Enthalpies of sublimation of organic and organometallic compounds," *J. Phys. Chem. Ref. Data*, vol. 31, pp. 537–698, 2002.

[19] Civalleri, B., Doll, K., and Zicovich-Wilson, C. M., "Ab initio investigation of structure and cohesive energy of crystalline urea," *J. Phys. Chem. B*, vol. 111, pp. 26–33, 2007.

[20] Claessens, C. G. and Stoddart, J. F., "π-π interactions in self-assembly," *J. Phys. Org. Chem.*, vol. 10, pp. 254–272, 1997.

[21] Cockroft, S. L. and Hunter, C. A., "Desolvation tips the balance: solvent effects on aromatic interactions," *Chem. Commun.*, vol. 36, pp. 3806–3808, 2006.

[22] Cockroft, S. L. and Hunter, C. A., "Chemical double-mutant cycles: dissecting non-covalent interactions," *Chem. Soc. Rev.*, vol. 36, pp. 172–188, 2007.

[23] Cozzi, F., Cinquini, M., Annuziata, R., Dwyer, T., and Siegel, J. S., "Polar/π interactions between stacked aryls in 1,8-diarylnaphthalenes," *J. Am. Chem. Soc.*, vol. 114, no. 14, pp. 5729–5733, 1992.

[24] Cozzi, F., Cinquini, M., Annuziata, R., and Siegel, J. S., "Dominance of polar/π over charge-transfer effects in stacked phenyl interactions," *J. Am. Chem. Soc.*, vol. 115, pp. 5330–5331, 1993.

[25] Cozzi, F., Ponzini, F., Annunziata, R., Cinquini, M., and Siegel, J. S., "Polar interactions between stacked π systems in fluorinated 1,8-diarylnaphthalenes: Importance of quadrupole moments in molecular recognition," *Angew. Chem. Int. Ed. Engl.*, vol. 34, no. 9, pp. 1019–1020, 1995.

[26] Crawford, T. D. and Schaefer, H. F., "An introduction to coupled cluster theory for computational chemists," in *Reviews in Computational Chemistry* (Lipkowitz, K. B. and Boyd, D. B., eds.), vol. 14, pp. 33–136, New York: VCH Publishers, 2000.

[27] Dahl, T., "The nature of stacking interactions between organic molecules elucidated by analysis of crystal structures," *Acta Chem. Scand.*, vol. 48, pp. 95–106, 1994.

[28] Daniel, J. M., Friess, S. D., Rajagopalan, S., Wendt, S., and Zenobi, R., "Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry," *Int. J. Mass Spec.*, vol. 216, pp. 1–27, 2002.

[29] Day, G. M., Chisholm, J., Shan, N., Motherwell, W. D. S., and Jones, W., "Assessment of lattice energy minimization for the prediction of molecular organic crystal structures," *Crystal Growth & Design*, vol. 4, pp. 1327–1340, 2004.

[30] Dion, M., Rydberg, H., Schröder, E., Langreth, D. C., and Lundqvist, B. I., "van der waals density functional for general geometries," *Phys. Rev. Letters*, vol. 92, no. 24, p. 246401, 2004.

[31] Duan, G. L., Smith, V. H., and Weaver, D. F., "Characterization of aromatic-thiol pi-type hydrogen bonding and phenylalanine-cysteine side chain interactions through ab initio calculations and protein database analyses," *Mol. Phys.*, vol. 99, pp. 1689–1699, 2001.

[32] Dunitz, J. D. and Gavezzotti, A., "Molecular recognition in organic crystals: Directed intermolecular bonds or nonlocalized bonding?," *Angew. Chem., Int. Ed.*, vol. 44, pp. 1766–1787, 2005.

[33] Dunitz, J. D. and Gavezzotti, A., "Toward a quantitative description of crystal packing in terms of molecular pairs: Application to the hexamorphic crystal system, 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile," *Crystal Growth & Design*, vol. 5, pp. 2180–2189, 2005.

[34] Edwards, T. H., Moncur, N. K., and Snyder, L. E., "Ground-state molecular constants of hydrogen sulfide," *J. Chem. Phys.*, vol. 46, p. 2139, 1967.

[35] Elstner, M., Hobza, P., Frauenheim, T., Suhai, S., and Kaxiras, E., "Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment," *J. Chem. Phys.*, vol. 114, pp. 5149–5155, 2001.

[36] Feller, D. J., "Application of systematic sequences of wave-functions to the water dimer," *J. Chem. Phys.*, vol. 96, p. 6104, 1992.

[37] Gauss, J. and Stanton, J. F., "The equilibrium structure of benzene," *J. Phys. Chem. A*, vol. 104, pp. 2865–2868, 2000.

[38] Gavezzotti, A., "Calculation of intermolecular interaction energies by direct numerical integration over electron densities. 1. electrostatic and polarization energies in molecular crystals," *J. Phys. Chem. B*, vol. 106, pp. 4145–4154, 2002.

[39] Gavezzotti, A., "Calculation of intermolecular interaction energies by direct numerical integration over electron densities. 2. an improved polarization model and the evaluation of dispersion and repulsion energies," *J. Phys. Chem. B*, vol. 107, pp. 2344–2353, 2003.

[40] Gavezzotti, A., "Quantitative ranking of crystal packing models by systematic calculations on potential energies and vibrational amplitudes of molecular dimers," *J. Chem. Theory Comput.*, vol. 1, pp. 834–840, 2005.

[41] Gavezzotti, A., *Molecular Aggregation: Structure Analysis and molecular simulation of crystals and liquids*. New York: Oxford University Press, 2007.

[42] Gonzalez, C. and Lim, E. C., "Evaluation of the hartree-fock dispersion (hfd) model as a practical tool for probing intermolecular potentials of small aromatic clusters: Comparison of the hfd and mp2 intermolecular potentials," *J. Phys. Chem. A*, vol. 107, pp. 10105–10110, 2003.

[43] Gourlay, M. D., Kendrick, J., and Leusen, F. J. J., "Rationalization of racemate resolution: Predicting spontaneous resolution through crystal structure prediction," *Crystal Growth & Design*, vol. 7, pp. 56–63, 2007.

[44] Grimme, S., "Accurate description of van der waals complexes by density functional theory including empirical corrections," *J. Comput. Chem.*, vol. 25, pp. 1463–1473, 2004.

[45] Halkier, A., Helgaker, T., Jørgensen, P., Klopper, W., Koch, H., Olsen, J., and Wilson, A. K., "Basis-set convergence in correlated calculations on Ne, N_2, and H_2O," *Chem. Phys. Lett.*, vol. 286, pp. 243–252, 1998.

[46] Halkier, A., Klopper, W., Helgaker, T., Jørgensen, P., and Taylor, P. R., "Basis set convergence of the interaction energy of hydrogen-bonded complexes," *J. Chem. Phys.*, vol. 111, pp. 9157–9167, 1999.

[47] Hesselmann, A., Jansen, G., and Schütz, M., "Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies," *J. Chem. Phys.*, vol. 122, p. 014103, 2005.

[48] Hobza, P., Selzle, H. L., and Schlag, E. W., "Potential energy surface of the benzene dimer: Ab initio theoretical study," *J. Am. Chem. Soc.*, vol. 116, pp. 3500–3506, 1994.

[49] Hobza, P., Selzle, H. L., and Schlag, E. W., "Potential energy surface for the benzene dimer. results of *ab initio* CCSD(T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced," *J. Phys. Chem.*, vol. 100, pp. 18790–18794, 1996.

[50] Hopkins, B. W. and Tschumper, G. S., "Ab initio studies of $\pi \cdots \pi$ interactions: The effects of quadruple excitations," *J. Phys. Chem. A*, vol. 108, no. 15, pp. 2941–2948, 2004.

[51] Hunter, C. A. and Sanders, J. K. M., "The nature of π-π Interactions," *J. Am. Chem. Soc.*, vol. 112, no. 14, pp. 5525–5534, 1990.

[52] Hunter, C. A., Singh, J., and Thornton, J. M., "π-π interactions: the geometry and energetics of phenylalanine-phenylalanine interactions in proteins," *J. Mol. Biol.*, vol. 218, pp. 837–846, 1991.

[53] Jaffe, R. L. and Smith, G. D., "A quantum chemistry study of benzene dimer," *J. Chem. Phys.*, vol. 105, pp. 2780–2788, 1996.

[54] Jeffrey, G. A., Ruble, J. R., McMullan, R. K., and Pople, J. A., "The crystal structure of deuterated benzene," *Proc. R. Soc. London, Ser. A*, vol. 414, pp. 47–57, 1987.

[55] Jensen, F., *Introduction to Computational Chemistry*. Chichester: Wiley, 1999.

[56] Jeziorski, B., Moszynski, R., and Szalewicz, K., "Perturbation theory approach to intermolecular potential energy surfaces of van der waals complexes," *Chem. Rev.*, vol. 94, pp. 1887–1930, 1994.

[57] Kendall, R. A., Dunning, T. H., and Harrison, R. J., "Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions," *J. Chem. Phys.*, vol. 96, pp. 6796–6806, 1992.

[58] Kim, E., Paliwal, S., and Wilcox, C. S., "Measurements of molecular electrostatic field effects in edge-to-face aromatic interactions and CH-π interactions with implications for protein folding and molecular recognition," *J. Am. Chem. Soc.*, vol. 120, pp. 11192–11193, 1998.

[59] Kong, J., White, C. A., Krylov, A. I., Sherrill, 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., and Pople, J. A., "Q-Chem 2.0: A high performance ab initio electronic structure program package," *J. Comp. Chem.*, vol. 21, pp. 1532–1548, 2000.

[60] Kumar, A., Elstner, M., and Suhai, S., "SCC-DFTB-D study of intercalating carcinogens: Benzo(a)pyrene and its metabolites complexed with the G-C base pair," *Int. J. Quantum Chem.*, vol. 95, pp. 44–59, 2003.

[61] Lee, E. C., Hong, B. H., Lee, J. Y., Kim, J. C., Kim, D., Kim, Y., Tarakeshwar, P., and Kim, K. S., "Substituent effects on edge-to-face aromatic interactions," *J. Am. Chem. Soc.*, vol. 127, pp. 4530–4537, 2005.

[62] Lee, E. C., Kim, D., Jurečka, P., Tarakeshwar, P., Hobza, P., and Kim, K. S., "Understanding of assembly phenomena by aromatic-aromatic interacitons: Benzene dimer and the substituted systems," *J. Phys. Chem. A*, vol. 111, pp. 3446–3457, 2007.

[63] Lehn, J.-M., *Supramolecular Chemistry: Concepts and Perspectives*. New York: VCH, 1995.

[64] Li, T. and Feng, S., "Empirically augmented density functional theory for predicting lattice energies of aspirin, acetaminophen polymorphs, and ibuprofen homochiral and racemic crystals," *Pharma. Res.*, vol. 23, pp. 2326–2332, 2006.

[65] Lide, D. R., ed., *CRC Handbook of Chemistry and Physics*. Boca Raton, FL: CRC Press, 77th ed., 1997.

[66] Macias, A. T. and A. D. MacKerell, J., "Ch/π interactions involving aromatic amino acids: Refinement of the charmm tryptophan force field," *J. Comput. Chem.*, vol. 26, pp. 1452–1463, 2005.

[67] Manby, F. R., Alfe, D., and Gillan, M. J., "Extension of molecular electronic structure methods to the solid state: Computation of the cohesive energy of lithium hydride," *Phys. Chem. Chem. Phys.*, vol. 8, pp. 5178–5180, 2006.

[68] Mei, X. and Wolf, C., "Highly congested nondistorted diheteroarylnaphthalenes: Model compounds for the investigation of intramolecular pi-stacking interactions," *J. Org. Chem.*, vol. 70, p. 2299, 2005.

[69] Meyer, E. A., Castellano, R. K., and Diederich, F., "Interactions with aromatic rings in chemical and biological recognition," *Angew. Chem. Int. Ed. Engl.*, vol. 42, no. 11, pp. 1210–1250, 2003.

[70] Mignon, P., Loverix, S., and Geerlings, P., "Interplay between pi-pi interactions and the h-bonding ability of aromatic nitrogen bases," *Chem. Phys. Lett.*, vol. 401, pp. 40–46, 2005.

[71] Morgan, R. S., Tatsh, C. E., Gushard, R. H., Mcadon, J. M., and Warme, P. K., "Chains of alternating sulfur and pi-bonded atoms in 8 small proteins," *Int. J. Pept. Prot. Res.*, vol. 11, pp. 209–217, 1978.

[72] Muraki, M., "The importance of ch/π interactions to the function of carbohydrate binding proteins," *Protein and Peptide Letters*, vol. 9, pp. 195–209, 2002.

[73] Nakamura, K. and Houk, K. N., "Theoretical studies of the wilcox molecular torsion balance. is the edge-to-face aromatic interaction important?," *Org. Lett.*, vol. 1, no. 13, pp. 2049–2051, 1999.

[74] Nakamura, M. and Miyazawa, T., "Frequency distribution, zero-point energy, specific heat, and neutron scattering of the benzene crystal," *J. Chem. Phys.*, vol. 51, pp. 3146–3147, 1969.

[75] Neumann, M. and Perrin, M., "Energy ranking of molecular crystals using density functional theory calculations and an empirical van der waals correction," *J. Phys. Chem. B*, vol. 109, pp. 15531–15541, 2005.

[76] Nishio, M., "Ch/π hydrogen bonds in crystals," *Cryst. Eng. Comm.*, vol. 6, pp. 130–158, 2004.

[77] Nishio, M., Hirota, M., and Umezawa, Y., *The CH/π Interaction*. New York: Wiley-VCH, 1998.

[78] Nishio, M., Umezawa, Y., Hirota, M., and Takeuchi, Y., "The ch/π interaction - significance in molecular recognition," *Tetrahedron*, vol. 51, pp. 8665–8701, 1995.

[79] Obst, U., Banner, D. W., Weber, L., and Diederich, F., "Molecular recognition at the thrombin active site: Structure-based design and synthesis of potent and selective thrombin inhibitors and the x-ray crystal structures of two thrombin-inhibitor complexes," *Chem. Biol.*, vol. 4, pp. 287–295, 1997.

[80] Page, R. H., Shen, Y. R., and Lee, Y. T., "Infrared-ultraviolet double resonance studies of benzene molecules," *J. Chem. Phys.*, vol. 88, pp. 5362–5376, 1988.

[81] Paliwal, S., Geib, S., and Wilcox, C. S., "Molecular torsion balance for weak molecular recognition forces. effects of "tilted-t" edge-to-face aromatic interactions on conformational selection and solid-state structure," *J. Am. Chem. Soc.*, vol. 116, pp. 4497–4498, 1994.

[82] Podeszwa, R. and Szalewicz, K., "Accurate interaction energies for argon, krypton, and benzene dimers from perturbation theory based on the kohn-sham model," *Chem. Phys. Lett.*, vol. 412, pp. 488–493, 2005.

[83] Podeszwa, R., "Comment on "beyond the benzene dimer: An investigation of the additivity of π-π interactions"," *J. Phys. Chem. A*, vol. 112, pp. 8884–8885, 2008.

[84] Podeszwa, R., Ric, B. M., and Szalewicz, K., "Predicting structure of molecular crystals from first principles," *Phys. Rev. Lett.*, vol. 101, p. 115503, 2008.

[85] Pople, J. A., Krishnan, R., Schlegel, H. B., and Binkley, J. S., "Electron correlation theories and their application to the study of simple reaction potential surfaces," *Int. J. Quantum Chem.*, vol. 14, pp. 545–560, 1978.

[86] Purvis, G. D. and Bartlett, R. J., "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples," *J. Chem. Phys.*, vol. 76, pp. 1910–1918, 1982.

[87] R. C. Lord, J., Ahlberg, J. E., and Andrews, D. H., "Calculation of heat capacity curves of crystalline benzene and benzene-d_6," *J. Chem. Phys.*, vol. 5, pp. 649–654, 1937.

[88] Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M., "A 5th-order perturbation comparison of electron correlation theories," *Chem. Phys. Lett.*, vol. 157, pp. 479–483, 1989.

[89] Rashkin, M. J. and Waters, M. L., "Unexpected substituent effects in offset π-π stacked interactions in water," *J. Am. Chem. Soc.*, vol. 124, no. 9, pp. 1860–1861, 2002.

[90] Reid, K. S. C., Lindley, P. F., and Thornton, J., "Sulfur-aromatic interactions in proteins," *FEBS Lett.*, vol. 190, pp. 209–213, 1985.

[91] Ribas, J., Cubero, E., Luque, F. J., and Orozco, M., "Theoretical study of Alkyl-π and Aryl-π interactions. reconciling theory and experiment," *J. Org. Chem.*, vol. 67, pp. 7057–7065, 2002.

[92] Riley, K. E. and Merz, K. M., "Effects of fluorine substitution on the edge-to-face interaction of the benzene dimer," *J. Phys. Chem. B*, vol. 109, pp. 17752–17756, 2005.

[93] Ringer, A. L., Figgs, M. S., Sinnokrot, M. O., and Sherrill, C. D., "Aliphatic C–H/π interactions: Methane-benzene, methane-phenol, and methane-indole complexes," *J. Phys. Chem. A*, vol. 110, pp. 10822–10828, 2006.

[94] Saenger, W., *Principles of Nucleic Acid Structure*. New York: Springer-Verlag, 1984.

[95] Schweizer, W. B. and Dunitz, J. D., "Quantum mechanical calculations for benzene dimer energies: Present problems and future challenges," *J. Chem. Theory Comput.*, vol. 2, pp. 288–291, 2006.

[96] Scuseria, G. E., Janssen, C. L., and Schaefer, H. F., "An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations," *J. Chem. Phys.*, vol. 89, p. 7382, 1988.

[97] Scuseria, G. E. and Lee, T. J., "Comparison of coupled-cluster methods which include the effects of connected triple excitations," *J. Chem. Phys.*, vol. 93, p. 5851, 1990.

[98] Scuseria, G. E., Scheiner, A. C., Lee, T. J., Rice, J. E., and Schaefer, H. F., "The closed-shell coupled sluster single and double excitation (CCSD) model for the description of electron correlation. a comparison with configuration interaction (CISD) results," *J. Chem. Phys.*, vol. 86, p. 2881, 1987.

[99] Shankar, R., *Principles of Quantum Mechanics*. New York: Springer, 1980.

[100] Shibasaki, K., Fujii, A., Mikami, N., and Tsuzuki, S., "Magnitude of the ch/π interaction in the gas phase: Experimental and theoretical determination of the accurate interaction energy in benzene-methane," *J. Phys. Chem. A*, vol. 110, pp. 4397–4404, 2006.

[101] Shimohigashi, Y., Nose, T., Yamauchi, Y., and Maeda, I., "Design of serine protease inhibitors with conformation restricted by amino acid sidechain-sidechain ch/π interaction," *Biopolym.*, vol. 51, pp. 9–17, 1999.

[102] Sinnokrot, M. O. and Sherrill, C. D., "Unexpected substituent effects in face-to-face π-stacking interactions," *J. Phys. Chem. A*, vol. 107, pp. 8377–8379, 2003.

[103] Sinnokrot, M. O. and Sherrill, C. D., "Highly accurate coupled cluster potential energy curves for benzene dimer: The sandwich, t-shaped, and parallel-displaced configurations," *J. Phys. Chem. A*, vol. 108, no. 46, pp. 10200–10207, 2004.

[104] Sinnokrot, M. O. and Sherrill, C. D., "Substituent effects in π-π interactions: Sandwich and t-shaped configurations," *J. Am. Chem. Soc.*, vol. 126, pp. 7690–7697, 2004.

[105] Sinnokrot, M. O. and Sherrill, C. D., "High-accuracy quantum mechanical studies of π-π interactions in benzene dimers," *J. Phys. Chem. A*, vol. 110, pp. 10656–10668, 2006.

[106] Sinnokrot, M. O., Valeev, E. F., and Sherrill, C. D., "Estimates of the ab initio limit for π-π interactions: The benzene dimer," *J. Am. Chem. Soc.*, vol. 124, pp. 10887–10893, 2002.

[107] Stephenson, R. and Malanowski, S., *Handbook of the Thermodynamics of Organic Compounds*. New York: Elsevier, 1987.

[108] Suresh, C. H., Alexander, P., Vijayalakshmi, K. P., Sajith, P. K., and Gadre, S. R., "Use of molecular electrostatic potential for quantitative assessment of inductive effect," *Phys. Chem. Chem. Phys.*, vol. 10, pp. 6492–6499, 2008.

[109] Tamagawa, K., Iijima, T., and Kimura, M., "Molecular structure of benzene," *J. Mol. Structure*, vol. 30, pp. 243–253, 1976.

[110] Tauer, T. P., Derrick, M. E., and Sherrill, C. D., "Estimates of the ab initio limit for sulfur-π interactions: The H_2S-Benzene dimer," *J. Phys. Chem. A*, vol. 109, no. 1, pp. 191–196, 2005.

[111] Tauer, T. P. and Sherrill, C. D., "Beyond the benzene dimer: An investigation of the additivity of π-π interactions," *J. Phys. Chem. A*, vol. 109, pp. 10475–10478, 2005.

[112] Toth, G., Murphy, R. F., and Lovas, S., "Stabilization of local structures by /*pi*-ch and aromatic-backbone amide interactions involving prolyl and aromatic residues," *Prot. Eng.*, vol. 14, pp. 543–547, 2001.

[113] Tsuzuki, S., Honda, K., Uchimaru, T., and Mikami, M., "Ab initio calculations of structures and interaction energies of toluene dimers including CCSD(T) level electron correlation correction," *J. Chem. Phys.*, vol. 122, p. 144323, 2005.

[114] Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., and Tanabe, K., "The magnitude of the ch/π interaction between benzene and some model hydrocarbons," *J. Am. Chem. Soc.*, vol. 122, pp. 3746–3753, 2000.

[115] Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., and Tanabe, K., "Origin of attraction and directionality of the π-π interaction: Model chemistry calculations of benzene dimer interaction," *J. Am. Chem. Soc.*, vol. 124, no. 1, pp. 104–112, 2002.

[116] Tsuzuki, S. and Lüthi, H. P., "Interaction energies of van der Waals and hydrogen bonded systems calculated using density functional theory: Assessing the PW91 model," *J. Chem. Phys.*, vol. 114, no. 9, pp. 3949–3957, 2001.

[117] Tsuzuki, S., Uchimaru, T., Matsumura, K., Mikami, M., and Tanabe, K., "Effects of the higher electron correlation correction on the calculated intermolecular interaction energies of benzene and naphthalene dimers: Comparison between MP2 and CCSD(T) calculations," *Chem. Phys. Lett.*, vol. 319, pp. 547–554, 2000.

[118] Tsuzuki, S., Uchimaru, T., Sugawara, K., and Mikami, M., "Energy profile of the interconversion path between t-shape and slipped-parallel benzene dimers," *J. Chem. Phys.*, vol. 117, no. 24, pp. 11216–11221, 2002.

[119] Tsuzuki, S., Uchimaru, T., and Tanabe, K., "Basis set effects on the intermolecular intaraction of hydrocarbon molecules obtained by ab initio molecular orbital method: Evaluation of dispersion energy," *J. Mol. Struct. THEOCHEM*, vol. 307, pp. 107–118, 1994.

[120] Umezawa, K., Kawakami, M., and Watanabe, T., "Molecular design and biological activities of protein-tyrosine phosphatase inhibitors," *Pharm. Thera.*, vol. 99, pp. 15–24, 2003.

[121] Umezawa, Y., Tsuboyama, S., Takahashi, H., and Nishio, J. U. M., "Ch/π interaction in the conformation of peptides. a database study," *Bioorg. Med. Chem.*, vol. 7, pp. 2021–2026, 1999.

[122] Vanspeybrouck, W., Herrebout, W. A., van der Veken, B. J., and Perutz, R. N., "Direct measurement of the stability of the supermolecular synthon benzene-hexafluorobenzene," *J. Phys. Chem. B*, vol. 107, pp. 13855–13861, 2003.

[123] von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U., and Sebastiani, D., "Optimization of effective atom centered potentials for london dispersion forces in density functional theory," *Phys. Rev. Letters*, vol. 93, no. 15, p. 153004, 2004.

[124] von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U., and Sebastiani, D., "Performance of optimized atom-centered potentials for weakly bonded systems using density functional theory," *Phys. Rev. B*, vol. 71, p. 195119, 2005.

[125] Watanabe, T., Suzuki, T., Umezawa, Y., Takeuchi, T., Otsuka, M., and Umezawa, K., "Structure-activity relationship and rational design of 3.4-dephostatin derivatives as protein tryosine phosphatase inhibitors," *Tetrahedron*, vol. 56, pp. 741–752, 2000.

[126] Werner, H. J. and Knowles, P. J. MOLPRO, a package of ab initio programs designed by H.J. Werner and P.J. Knowles, version 2002.1, R.D. Amos, A. Bernhardsson, A. Benning, P. Celani, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P.J. Knowles, T. Korona, R. Lindh, A.W. Lloyd, S.J. McNicholas, F.R. Manby, W. Meyer, M.E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Sch tz, U. Schumann, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, and H.J. Werner.

[127] Wheeler, S. E. and Houk, K. N., "Substituent effects in the benzene dimer are due to direct interactions of the substituents with the unsubstituted benzene," *J. Am. Chem. Soc.*, vol. 130, pp. 10854–10855, 2008.

[128] Williams, H. L., Szalewicz, K., Jeziorski, B., Moszynski, R., and Rybak, S., "Symmetry-adapted perturbation theory calculation of the Ar-H_2 intermolecular potential energy surface," *J. Chem. Phys.*, vol. 98, pp. 1279–1292, 1993.

[129] Williams, J. H., Cockcroft, J. K., and Fitch, A. N., "Structure of the lowest temperature phase solid benzene-hexafluorobenzene adduct," *Angew. Chem. Int. Ed. Engl.*, vol. 31, pp. 1655–1657, 1992.

[130] Williams, V. E. and Lemieux, R. P., "Role of dispersion and electrostatic forces on solute-solvent interactions in a nematic liquid crystal phase," *J. Am. Chem. Soc.*, vol. 120, pp. 11311–11315, 1998.

[131] Woon, D. E. and Dunning, T. H., "Gaussian basis sets for use in correlated molecular calculations. v. core-valence basis sets for boron through neon," *J. Chem. Phys.*, vol. 103, p. 4572, 1995.

[132] Wu, Q. and Yang, W., "Empirical correction to density functional theory for van der waals interactions," *J. Chem. Phys.*, vol. 116, pp. 515–524, 2002.

[133] Wu, X., Vargas, M. C., Nayak, S., Lotrich, V., and Scoles, G., "Towards extending the applicability of density functional theory to weakly bound systems," *J. Chem. Phys.*, vol. 115, pp. 8748–8757, 2001.

[134] Zauhar, R. J., Colbert, C. L., Morgan, R. S., and Welsh, W. J., "Evidence for a strong sulfur-aromatic interaction derived from crystallographic data," *Biopolymers*, vol. 53, pp. 233–248, 2000.

[135] Zhao, Y., Schultz, N. E., and Truhlar, D. G., "Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions," *J. Chem. Theory Comput.*, vol. 2, pp. 364–382, 2006.

[136] Zhao, Y. and Truhlar, D. G., "Multicoefficient extrapolated density functional theory studies of $\pi \cdots \pi$ interactions: The benzene dimer," *J. Phys. Chem. A*, vol. 109, pp. 4209–4212, 2005.

[137] Zimmerli, U., Parrinello, M., and Koumoutsakos, P., "Dispersion corrections to density functionals for water aromatic interactions," *J. Chem. Phys.*, vol. 120, pp. 2693–2699, 2004.

VITA

Ashley L. Ringer was born in Mississippi, where she lived for the next 22 years. Despite her continually mischievous behavior throughout childhood, her parents managed not to kill her, and she eventually graduated from Florence High School in Florence, Mississippi in 2000. She then attended Mississippi College, where she worked under Prof. David H. Magers, graduating *summa cum laude* with a B.S. in Chemistry in 2004. Then, despite having never lived in a town with more than two major intersections, she felt it appropriate to move to Atlanta, Georgia to pursue a Ph.D. with Prof. C. David Sherrill in the School of Chemistry and Biochemistry at the Georgia Institute of Technology.