

University of Mississippi

eGrove

Electronic Theses and Dissertations

Graduate School

2011

Theoretical Characterization of Non-Covalent Weakly Bound Clusters Through the Application of Sophisticated Computational Quantum Chemistry Methodologies and the Development of Integrated Fragmentation Techniques

Desiree Bates

Follow this and additional works at: <https://egrove.olemiss.edu/etd>

 Part of the [Physical Chemistry Commons](#)

Recommended Citation

Bates, Desiree, "Theoretical Characterization of Non-Covalent Weakly Bound Clusters Through the Application of Sophisticated Computational Quantum Chemistry Methodologies and the Development of Integrated Fragmentation Techniques" (2011). *Electronic Theses and Dissertations*. 48.
<https://egrove.olemiss.edu/etd/48>

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

THEORETICAL CHARACTERIZATION OF NON-COVALENT WEAKLY BOUND CLUSTERS
THROUGH THE APPLICATION OF SOPHISTICATED COMPUTATIONAL QUANTUM
CHEMISTRY METHODOLOGIES AND THE DEVELOPMENT OF INTEGRATED
FRAGMENTATION TECHNIQUES

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Mississippi

DESIREE M. BATES

March 2011

© 2011

Desiree M. Bates

ALL RIGHTS RESERVED

ABSTRACT

Non-covalent, weakly bound clusters have been and remain of significant interest to many researchers. However, with computational studies, accurate description of these interactions requires sophisticated electronic structure methods employing large basis sets. This methodology becomes extremely computationally demanding as the size of the system increases. This work presents benchmark data and explores methods for obtaining highly accurate *ab initio* results for larger systems at greatly reduced computational costs. CCSD(T) complete basis set limit interaction energies are presented for a variety of parallel-slipped $\pi \cdots \pi$ dimers and low-lying isomers of $(\text{H}_2\text{O})_6$. The calibration of a 2-body:Many-body fragmentation method for computing interaction energies of several $(\text{H}_2\text{O})_n$ clusters with n ranging from 3-10 is performed. As a result, 2-body:Many-body QM:QM approach is extended to a 3-body:Many-body technique. In addition to calculating the energetics, the 2-body:Many-body fragmentation method, which is cast within the ONIOM framework, is used for the determination of Cartesian analytic gradients for the purpose of geometry optimizations.

DEDICATION

For Nolan; I never knew true motivation until he came into my life.

ACKNOWLEDGEMENTS

I would like to thank all, past and present, Tschumper Research Group Members. In particular, Adel ElSohly and Dr. Brian Hopkins for all their helpful discussions. I would especially like to thank Dr. Gregory S. Tschumper for his continuous guidance, mentoring and teaching. I would also like to thank Dr. Nathan I. Hammer, Dr. Jeanne Franz, and Dr. Thomas Nalli for their professional support throughout my graduate career.

Finally, I must also thank my husband, Paul. Without his support and constant motivation throughout the years, I never would have finished what I started. He believed in me from start to finish.

This work was financially supported by the National Science Foundation (EPS-01326618, EPS-0556308, EPS-0903787, CHE-051067 and CHE-0957317). Many of the calculations reported here were performed with Mississippi Center for Supercomputing Research cluster resources.

CONTENTS

CHAPTER

1	Introduction	1
1.1	Non-covalent Interactions	1
1.1.1	Scaling Problem: Basis Sets and Electronic Structure Methods	2
1.1.2	Basis Set Incompleteness Error and Basis Set Superposition Error	3
1.1.3	Basis Set Extrapolation Techniques and Explicitly Correlated Methods	4
1.1.4	Methods for Studying Large Non-Covalent Clusters	5
2	Probing the effects of heterogeneity on delocalized $\pi \cdots \pi$ interaction energies	9
2.1	Abstract	9
2.2	Introduction	10
2.3	Computational Details	12
2.4	Results	15
2.5	Conclusions	20
2.6	Acknowledgements	21

3	CCSD(T) Complete Basis Set Limit Relative Energies for Low-lying Water	
	Hexamer Structures	22
3.1	Abstract	22
3.2	Introduction	23
3.3	Computational Details	26
3.4	Results	29
	3.4.1 Relative Electronic Energies	29
	3.4.2 ZPVE Inclusive Relative Energies	31
3.5	Conclusions	34
3.6	Acknowledgments	36
3.7	Note Added in Proof	36
4	Development of a 3-body:Many-body Integrated Fragmentation Method	
	for Weakly Bound Clusters and Application to Water Clusters (H₂O)_n	
	n=3-9	37
4.1	Abstract	37
4.2	Introduction	38
4.3	Theoretical Methods	39
4.4	Computational Methods	43
4.5	Results and Discussion	45
	4.5.1 2-body:Many-body Approximation	45
	4.5.2 3-body:Many-body Approximation	48

4.6	Conclusions	50
4.7	Acknowledgements	52
5	Efficient and Accurate Methods for the Geometry Optimization of Water Clusters: Application of Analytic Gradients for the 2-body:Many-Body QM:QM Fragmentation Method to $(\text{H}_2\text{O})_n$, $n = 3 - 10$	53
5.1	Abstract	53
5.2	Introduction	54
5.3	Theoretical Background	55
5.4	Computational Methods	58
5.5	Results and discussion	59
5.6	Conclusions	72
5.7	Acknowledgements	73
6	Conclusions	74
	Bibliography	77
7	Appendix	97
7.1	Structure of various water clusters $(\text{H}_2\text{O})_n$ $n=3-10$	98
7.2	Supporting Information for CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures	107

LIST OF TABLES

2.1	MP2 interaction energies in kcal mol ⁻¹ without CP corrections.....	16
2.2	Corrections to E_{int} from higher-order correlation effects obtained from CCSD(T) and MP2 calculations with the haDZ basis sets in kcal mol ⁻¹ .	16
2.3	MP2 and CCSD(T) CBS limit interaction energies in kcal mol ⁻¹	17
2.4	SCS-MP2 and SCSN-MP2 interaction energies in kcal mol ⁻¹	18
2.5	Contribution to the interaction energy from SAPT2 calculations	19
3.1	MP2 and CCSD(T) CBS limits for 8 low-lying (H ₂ O) ₆ isomers.....	30
3.2	Harmonic ZPVE corrections for 8 low-lying (H ₂ O) ₆ isomers.....	32
4.1	MP2 and CCSD(T) obtained with haTZ basis set interaction energies for low-lying(H ₂ O) _n clusters, $n=3-10$	47
4.2	Errors associated with n -body QM:QM methods for low-lying (H ₂ O) _n clusters $n=3-10$	49
5.1	RMS deviations for optimized structures relative to the MP2/haTZ structures.	62
5.2	Average and Maximum RMS deviations for various optimized structures relative to the MP2/haTZ optimized structures for low-lying (H ₂ O) _n clusters $n=3-10$	66

5.3	Errors associated with MP2/haTZ energies performed on various structures relative to the MP2/haTZ//MP2/haTZ values.....	68
5.4	Average and maximum errors for MP2/haTZ energies performed on various optimized structures relative to the MP2/haTZ/MP2/haTZ values.	72

LIST OF FIGURES

2.1	6 dimers used to study the effect of heterogeneity on $\pi \cdots \pi$ dimers	13
2.2	Structures of 8 low-lying $(\text{H}_2\text{O})_6$ isomers.....	25

CHAPTER 1

INTRODUCTION

1.1 NON-COVALENT INTERACTIONS

It is hard to grasp how much of an impact non-covalent interactions have in supramolecular chemistry. Life, as we know it, would not exist without non-covalent interactions. The structure and function of biomolecules such as DNA, RNA and proteins, which are essential for life, are all governed by non-covalent interactions. For example, the first step in HIV infection is the formation of a non-covalent interaction between a viral envelope and cellular receptor of a protein.^{1,2} In addition to the biological impact, non-covalent interactions impact solvation, condensation, catalysis, assembly of nanomaterials, molecular recognition and absorption, distribution, metabolism, and excretion of pharmaceuticals in the body.

Understanding and characterizing non-covalent interactions, in particular hydrogen bonding and stacking interactions for weakly bound clusters, is the purpose for the research presented in this dissertation. Both of these motifs are found within DNA, RNA and protein-ligand interactions. Though both interactions can simultaneously exist in a single complex, the origins of hydrogen bonding and stacking interactions are different. Hydrogen bonding is dominated by electrostatic and charge-transfer interactions whereas stacking interactions are a result of London dispersion forces. Although electrostatic interactions are the largest

attractive force for non-covalent interactions, correct description of London dispersion is still necessary for highly accurate results. Unfortunately, dispersion is also a direct result of electron correlation and therefore is a challenge for computational chemists.³⁻⁶

Because of the importance of including electron correlation for non-covalent interactions, it is essential to determine an electronic structure method that routinely produces accurate results for these interactions. Coupled-cluster (CC) theory, in particular the coupled-cluster method which includes single and double electron excitations iteratively along with a perturbative approximation of connected triple excitations, CCSD(T), is the “gold standard” for computational methods.⁷⁻⁹

Besides choosing the correct electronic structure method for studying non-covalent clusters, it is equally important to employ an appropriate basis set. An ideal basis set would include an infinite number of basis functions. However, an infinite number of basis functions is not a realistic option for computational chemist. Therefore, within quantum chemistry it is a daunting and ongoing task to determine the best basis set for a given situation. Intuitively, one would think the larger of a basis set, the better. However, this is only the case for basis sets which are designed to systematically converge to the CBS limit. For example, Dunning developed a family of basis sets known as correlation-consistent basis sets.¹⁰⁻¹²

1.1.1 SCALING PROBLEM: BASIS SETS AND ELECTRONIC STRUCTURE METHODS

Seeking the best result the obvious approach is to use the best method available (generally, CCSD(T)) in the largest basis set possible. The problem of polynomial scaling makes this approach impossible. The computational cost (memory, disk space, and CPU time) of

CCSD(T) scale with the size of the system to the 7th power. In the context of intermolecular interactions, this means that if a CCSD(T) calculation on a single molecule takes an hour, a similar calculation on two molecules would take $1 \text{ hour} \times 2^7 = 128 \text{ hours}$, and a similar calculation on 6 molecules would take $1 \text{ hour} \times 6^7 = 32 \text{ years}$.

1.1.2 BASIS SET INCOMPLETENESS ERROR AND BASIS SET SUPERPOSITION ERROR

There is, of course, no way to truly employ a basis set containing an infinite number of basis functions. The error associated with not using an infinite number of basis functions is referred to as basis set incompleteness error (BSIE). BSIE is present in every calculation. Various extrapolation techniques outlined in the next section are used to measure and correct for BSIE.

Another error that arises from the use of incomplete basis sets is basis set superposition error (BSSE).^{13,14} BSSE arises from inconsistencies in a basis sets. One form of BSSE can be explained by considering the different stereoisomers of the ethene molecule. Take for example the gauche and eclipsed isomers. Because the functions are centered on the nuclei, the location of basis functions varies between the eclipsed and gauche conformations of ethene. This leads to an inconsistency. Another example is comparing a molecule optimized with two different methods that employed the same basis set. The nuclei for both optimized geometries will differ and therefore the position of basis functions will also differ between both geometries.

A second form of BSSE arises from comparing the energy of a cluster to the energies of the fragments (e. g., whenever computing dissociation or interaction energies). Some basis

functions present in the cluster computation are missing for the isolated fragments causing an inconsistency in basis sets. The counterpoise correction introduced by Boys and Bernardi is a technique to remedy BSSE.¹⁵ Readers interested in a more detailed treatment of the performance of CP corrections for weakly bound systems are directed to Reference 16 and References within.

1.1.3 BASIS SET EXTRAPOLATION TECHNIQUES AND EXPLICITLY CORRELATED METHODS

Because the use of a basis set with an infinite number of basis functions is impossible, many researchers attempt to approach the same result, the complete basis set (CBS) limit, by various techniques. Again, because correlation-consistent basis sets are constructed in a way to systematically converge to the CBS limit, it is possible to extrapolate to the CBS limit. Feller developed a three-parameter exponential function that accurately predicts the Hartree-Fock (HF) energy at the CBS limit.^{17,18} Additional techniques were developed to obtain CBS limits for correlation energy. Hartree-Fock energy is defined as the difference between the total energy and the correlation energy. Therefore, extrapolation techniques designed for correlation energy should not be applied to the total energy. A simple two-parameter formula introduced by Hekgaker and coworkers is one of the most popular techniques.¹⁹

Besides extrapolation techniques, other more recent approaches have been introduced to obtain the CBS limit. Explicitly correlated R12 and F12 methods are one way to expedite the convergence to the CBS limit.²⁰⁻²⁷ Because of the slow convergence with respect to the Coulomb hole around an electron, interelectronic distances (explicit distance between two

electrons) are included in the wavefunction for explicitly correlated methods. This idea was first introduced by Hylleraas in 1929.²⁸ However, due to the increase in computational cost associated with these methods their routine application to large systems is unrealistic.

1.1.4 METHODS FOR STUDYING LARGE NON-COVALENT CLUSTERS

In a response to the scaling problem, a number of techniques have been developed for studying non-covalent clusters. One family of methods works by introducing approximations into coupled cluster methods. Fragment molecular orbital²⁹⁻³² (FMO) and divide-and-conquer methods³³⁻³⁵ achieve near linear scaling. These methods either perform the actual calculation on fragments while embedded in an electrostatic field of the entire system (FMO) or are based on a simple sampling algorithm (divide-and-conquer). In 2005, Hirata and coworkers improved on the FMO idea and obtained errors less than 0.0001% for total energies of water clusters.³⁶

Another approach to obtaining accurate interaction energies is to improve the overall performance of less demanding methods such as MP2. For stacking interactions such as $\pi \cdots \pi$ interactions, MP2 is known for its large overestimation of the interaction energy. For example, in the parallel-displaced configuration of the benzene dimer the correction for higher-order correlation is (i. e. that part not captured by MP2) 2.18 kcal mol⁻¹. This is 58% of the total CCSD(T)/CBS limit interaction energy! Because of this large discrepancy, Spin-Component Scaled methods have been developed in an attempt to improve the accuracy of MP2 for $\pi \cdots \pi$ interactions.³⁷ Because Hartree-Fock (HF) contains no correlation of electrons with opposite spin, Grimme determined scaling factors for the MP2 correlation

energy to compensate for the lacking correlation within HF. Others have attempted to optimize these scaling factors for different types of interactions such as Spin-Component Scaled method for weak and stacking interactions³⁸ and Scaled Opposite-Spin.³⁹

The N-body decomposition approach introduced by Jordan and Christie permits calculation of accurate interaction energies for water clusters containing as many as 50 monomers.⁴⁰ With this method, the N-body expansion is truncated to contain up to 4-body terms. For the 1- and 2-body terms large basis sets are used; however, additional speed up is reached when small basis sets are applied to the 3- and 4-body terms.

Similarly, Truhlar’s group developed a method based on a many-body expansion.⁴¹ The Electrostatic Embedded Many-Body Expansion (EE-MB) method divides a cluster into fragments composed of monomers, dimers and trimers. Each fragment is computed with an accurate electronic structure method while embedded in a field of point charges acting as the electrostatic field of the other fragments in the cluster. This method yields interaction energies with an error 0.4% of the high-level calculation on the entire cluster for a low-lying water pentamer structure.

The effective fragment potential method (EFP1 and EFP2) was developed by Gordon and coworkers.^{42–46} Here the interaction energy is composed into five terms. These terms can be defined as either short- or long-range interactions. Short-range interactions (exchange-repulsion and charge transfer) decay exponentially, $V \frac{1}{R}$ whereas long-range interactions (coulomb, induction and dispersion) are dependent on distance and decay as $(1/R)^n$, where n is 2 for charge-charge interaction, 4 for charge-induced dipole interaction, 5 dipole-

induced dipole interaction and 6 for induced dipole-induced dipole interaction. Within the EFP2 method, the coulomb and induction terms are treated by classical approximations at long range, while exponential damping functions are used to obtain correct behavior at short-range. Exchange-repulsion interaction is expressed as a 2-body interaction and is derived as an expansion. The overlap expansion then uses a frozen localized molecular orbital treatment on each monomer. Results show the expansion can be truncated at the quadratic term and can employ much larger basis sets with minimal cost. However, implementation of the EFP2 method with quantum mechanical methods is still being developed. For more information on EFP or FMO methods readers are directed to the outstanding review article by Gordon et. al.⁴⁶

Jiang, Ma, and Li developed a method for linearly scaled coupled-cluster calculations named clusters in molecules.^{47,48} This method employs localized molecular orbitals and limits the double excitations with a spacial threshold; with both approximations made, this method scales linearly and can be applied to large molecular systems. For a wide range of molecules, this method recovers 98.5% of the total CCSD correlation energy. However, this method has not been expanded to include triple excitations which are necessary to reproduce the “gold standard”, CCSD(T).

Many-body decomposition schemes are an alternative to the previous approaches. For a detailed explanation of the many-body approach see Chapters 4 and 5. Tschumper and co-workers have shown that for weakly bound systems a 2-body:Many-body, can reproduce CCSD(T) interactions energies within 1%. This accuracy is also achieved with a fraction of

the computational cost of the full CCSD(T) calculation.⁴⁹ With this method, the largest CCSD(T) calculation is a dimer. Because this approach is implemented within the ONIOM formalism, (i. e. using a simple linear combination) it is extensible to first and second analytic derivatives.⁵⁰ Additional information about the implementation of analytic derivatives within the 2-body:Many-body technique can be found in Chapter 5.

Gregory Beran has developed a similar technique to the 2-body:Many-body approach.⁵¹ With this method all 1- and 2-body terms are calculated quantum mechanically whereas the higher-order terms are treated classically. A polarizable force field such as AMOEBA is used to calculate higher-order n-body effects, where $n \geq 3$. In addition, the Beran work also examined the effects of embedding the quantum mechanical 1- and 2-body calculations in a electrostatic field representing the other fragments. The results from embedding charges were very dependent on the choice of basis set used to determine the charges. In 2010, Beran improved his approach by parameterizing the force field used.⁵² Unfortunately, with the hybrid many-body interaction approach the force field has to be reparametrized for every different type of molecule.

CHAPTER 2

PROBING THE EFFECTS OF HETEROGENEITY ON DELOCALIZED $\pi \cdots \pi$ INTERACTION ENERGIES

2.1 ABSTRACT

Dimers composed of benzene (Bz), 1,3,5-triazine (Tz), cyanogen (Cy) and diacetylene (Di) are used to examine the effects of heterogeneity at the molecular level and at the cluster level on $\pi \cdots \pi$ stacking energies. The MP2 complete basis set (CBS) limits for the interaction energies (E_{int}) of these model systems were determined with extrapolation techniques designed for correlation consistent basis sets. CCSD(T) calculations were used to correct for higher-order correlation effects ($\delta E_{\text{MP2}}^{\text{CCSD(T)}}$) which were as large as $+2.81 \text{ kcal mol}^{-1}$. The introduction of nitrogen atoms into the parallel-slipped dimers of the aforementioned molecules causes significant changes to E_{int} . The CCSD(T)/CBS E_{int} for Di/Cy is $-2.47 \text{ kcal mol}^{-1}$ which is substantially larger than either Cy/Cy ($-1.69 \text{ kcal mol}^{-1}$) or Di/Di ($-1.42 \text{ kcal mol}^{-1}$). Similarly, the heteroaromatic Bz/Tz dimer has an E_{int} of $-3.75 \text{ kcal mol}^{-1}$ which is much larger than either Tz/Tz ($-3.03 \text{ kcal mol}^{-1}$) or Bz/Bz ($-2.78 \text{ kcal mol}^{-1}$). Symmetry-adapted perturbation theory calculations reveal a correlation between the electrostatic component of E_{int} and the large increase in the interaction energy for the mixed dimers. However, all components (exchange, inductions, dispersion) must be consid-

ered to rationalize the observed trend. Another significant conclusion of this work is that basis set superposition error has a negligible impact on the popular $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ correction, which indicates that counterpoise corrections are not necessary when computing higher-order correlation effects on E_{int} . Spin component scaled MP2 (SCS-MP2 and SCSN-MP2) calculations with a correlation consistent triple- ζ basis set reproduce the trends in the interaction energies despite overestimating the CCSD(T)/CBS E_{int} of Bz/Tz by 20–30%.

2.2 INTRODUCTION

Weak intermolecular forces play a significant role in biological systems. Specifically, $\pi \cdots \pi$ stacking interactions of aromatic systems are of particular importance. They largely contribute to essential biological systems such as DNA base pair stacking,^{53–56} protein-ligand interactions,^{57–61} and adenosine 5'-triphosphate (ATP) recognition.^{62,63} These $\pi \cdots \pi$ stacking interactions are difficult to isolate and study since they generally consist of a small aromatic system interacting with the side chain of a much larger molecular system (containing many atoms) which are currently too large for study of high accuracy quantum chemical levels of theory. Instead, many smaller prototypes have been studied in order to learn more about the nature of $\pi \cdots \pi$ stacking interactions.^{61,64–66} By far the most widely studied molecule for modeling $\pi \cdots \pi$ interactions is the benzene dimer (Bz/Bz).^{67–83} It was recently noted that the delocalized $\pi \cdots \pi$ dimer consisting of two diacetylene molecules, $(\text{H} - \text{C} \equiv \text{C} - \text{C} \equiv \text{C} - \text{H})_2$ or Di/Di, behaves very much like the benzene dimer and can also serve as a useful prototype for π -type interactions.⁸⁴

While Bz/Bz and Di/Di are extremely useful models, π -type interactions in real systems such as biomolecules and nanomaterials tend to be more intricate due to substituent effects (e.g., aromatic amino acids) and the presence of heteroatoms (e.g., nucleic acid bases). Substituent effects have already been examined in detail.^{85–88} One of the most significant outcomes of these studies is that simple electrostatic arguments based on Hunter-Sanders rules cannot account for the relative stability of monosubstituted benzene dimers where “Dispersion and exchange-repulsion are more important than electrostatics in determining the total binding energies”.⁸⁶ A few studies of heteroaromatic $\pi \cdots \pi$ systems have appeared in the literature.^{9,89–91} Unfortunately, the effect of heteroatoms on $\pi \cdots \pi$ stacking interactions was not the focus of these investigations.

Unlike the Bz/Bz dimer, relatively few studies have examined the Tz/Tz or Bz/Tz dimers. Eleven arrangements of the Bz/Tz dimer have been studied by Massera *et al.* including T-shaped, stacked and parallel-slipped structures in an effort to obtain a description of the Bz/Tz potential energy surface.⁹² They found the parallel-slipped structure to be the most stable Bz/Tz configuration with an interaction energy of $-5.28 \text{ kcal mol}^{-1}$ at the MP2/6-311++G(3df,p) level of theory. Šponer and Hobza provided one of the earliest studies of the Tz/Tz dimer with correlated *ab initio* methods.^{93,94} Dispersion-corrected density functional theory (DFT-D) and quantum Monte Carlo (QMC) calculations confirm that the “anti” Tz/Tz structure used by Šponer and Hobza is the most stable configuration⁹⁵.

To examine the effects of heteroatoms on $\pi \cdots \pi$ stacking within heterogenous dimers, the interaction energies of 3 parallel slipped dimers composed of 1,3,5-triazine and benzene

(Tz/Tz, Bz/Bz and Bz/Tz, bottom of Figure 1) are compared at the CCSD(T) complete basis set (CBS) limit. For a study of this nature, Tz has an advantage over other N-substituted benzenes because it is a highly symmetric and non-polar molecule (similar to Bz). The analogous structures for the smaller $\pi \cdots \pi$ prototypes composed of diacetylene and cyanogen, $\text{N} \equiv \text{C} - \text{C} \equiv \text{N}$, are also examined (Cy/Cy, Di/Di and Di/Cy, top of Figure 1). Spin component scaled MP2 (SCS-MP2 and SCSN-MP2) interaction energies are also computed in an effort to identify a less demanding computational approach that can be used to more thoroughly characterize the Tz/Tz and Bz/Tz systems as well as dimers composed of other N-substituted benzenes.

2.3 COMPUTATIONAL DETAILS

The Di/Cy, Tz/Tz, and Bz/Tz parallel-slipped dimers shown in Figure 1 were optimized with C_s symmetry using second-order Møller-Plesset perturbation theory (MP2) in conjunction with a double- ζ basis set including diffuse and polarization functions on all atoms (DZP++).⁹⁶⁻⁹⁸ The corresponding Bz and Tz monomers were optimized at this same level of theory with D_{6h} and D_{3h} symmetry, respectively. The geometries of the Di/Di and Cy/Cy dimers as well as the Di and Cy monomers were obtained from literature.⁸⁴ Bz/Bz interaction energies were taken from References 76 and 88. All computations were carried out with the Gaussian 03,⁹⁹ PSI3,¹⁰⁰ MPQC 2.3.1¹⁰¹⁻¹⁰⁵ and SAPT2006^{106,107} quantum chemistry software packages.

Single point energy calculations were obtained for the structures at various levels of

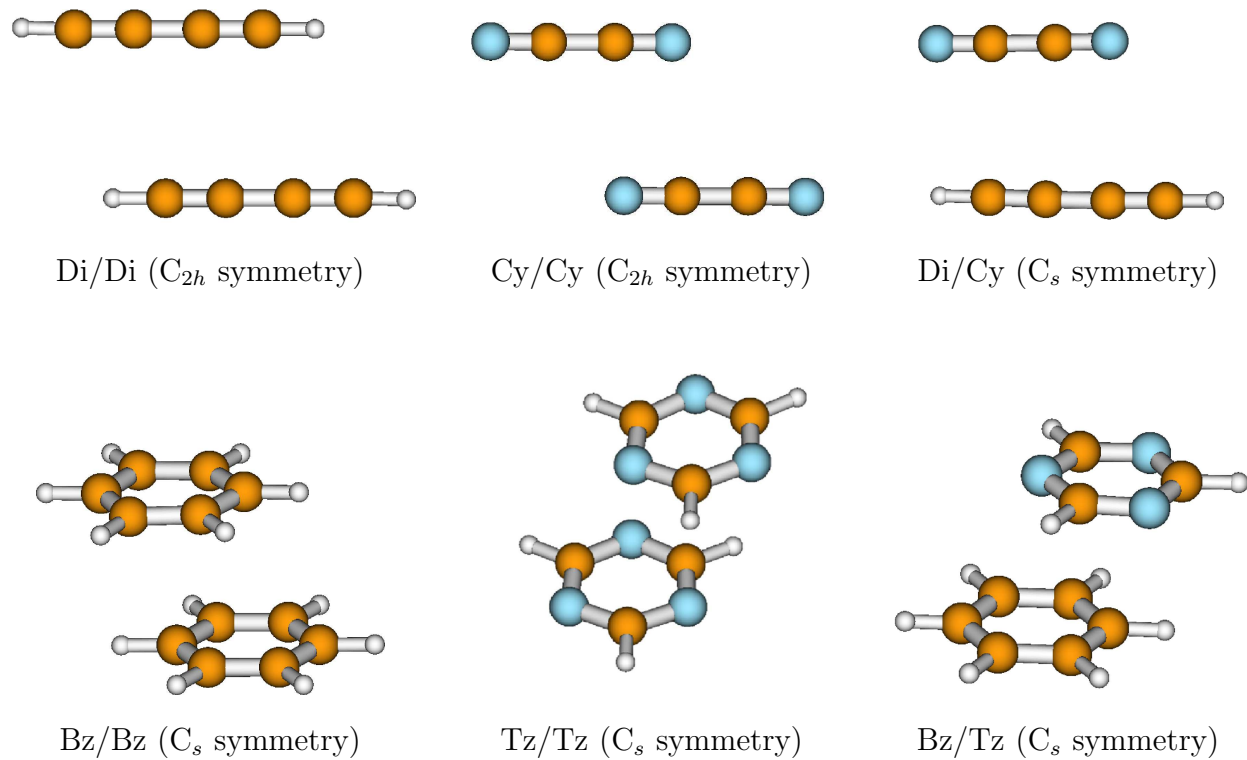


Figure 2.1: Parallel-slipped structures of the six dimers used to study the effects of heterogeneity on $\pi \cdots \pi$ stacking interactions.

theory. MP2 energies were computed with a series of Dunning’s correlation consistent basis sets where diffuse functions have been added to C and N but not to H (i.e., cc-pVXZ for H and aug-cc-pVXZ for C and N, where $X = D, T, Q, 5$). Hereafter, this basis set is denoted haXZ. All MP2 calculations (including geometry optimizations) were performed with the frozen core approximation. The MP2 CBS limit is obtained by using the haQZ and ha5Z energies in the two-parameter extrapolation scheme suggested by Helgaker *et al.*¹⁰⁸

The effect of higher-order excitations was investigated with CCSD(T)/haDZ single point energy calculations, again employing the frozen core approximation. Basis set superposition error (BSSE),¹³ is known to cause discrepancies in interaction energies computed with small basis sets and, therefore, is addressed here by performing Boys and Bernardi counterpoise (CP) corrections.^{109,110} In particular, the effect of CP corrections on the difference between MP2 and CCSD(T) interaction energies ($\delta E_{\text{MP2}}^{\text{CCSD(T)}}$) is examined. The CCSD(T) CBS limit of E_{int} is obtained by combining the MP2 CBS interaction energy with this correction for higher-order correlation effects.^{16,59,83,111,112}

$$E_{\text{int}}^{\text{CCSD(T)/CBS}} = E_{\text{int}}^{\text{MP2/CBS}} + \delta E_{\text{MP2}}^{\text{CCSD(T)}} \quad (2.1)$$

The spin-component-scaled (SCS) method developed by Grimme allows for separate scaling of the parallel and antiparallel contributions to the MP2 energy.³⁷ Later, Hill and Platts introduced the SCSN-MP2 method by reoptimizing the scaling parameters for nucleobase interactions.³⁸ The SCSN parameters were designed for aug-cc-pVTZ basis set. Both methods provide improved results for $\pi \cdots \pi$ interactions.³⁸ SCS-MP2 and SCSN-MP2 interaction

energies have also been evaluated here. Note that the efficient parallel MPQC program that was used to perform the largest MP2 calculations does not currently provide these contributions to the correlation energy. The parallel and antiparallel components were obtained with a different software package that could not perform the larger MP2 calculations on the same computational resources.

To gain some insight into why E_{int} changes as heteroatoms are introduced, a series of second-order symmetry-adapted perturbation theory (SAPT2)¹¹³ computations have been performed on the dimers with the haDZ and haTZ basis set. (SAPT2 calculations with the haTZ basis set are feasible only for the smaller dimers: Di/Di, Cy/Cy, Di/Cy.) The electrostatic, exchange, induction and dispersion contributions to the interaction energy are reported. As in recent work on substituted benzene dimers, the exchange-induction and exchange-dispersion terms are included as part of the induction and dispersion energies, respectively.

2.4 RESULTS

Table 2.1 displays the MP2 interaction energies computed with the haXZ basis sets. When considering this data along with comparable values for Bz/Bz ($E_{\text{int}}^{\text{MP2/CBS}} = -4.79 \text{ kcal mol}^{-1}$ ⁸⁸) one sees that the mixed dimers (Di/Cy and Bz/Tz) are substantially more stable than their homogeneous counterparts (Di/Di, Cy/Cy, Bz/Bz and Tz/Tz).

Table 2.2 shows the differences between the MP2/haDZ and CCSD(T)/haDZ interaction energies ($\delta E_{\text{MP2}}^{\text{CCSD(T)}}$) both with and without CP corrections. As expected, MP2 substantially overbinds relative to CCSD(T) (by as much as $2.35 \text{ kcal mol}^{-1}$). It was somewhat surprising

Table 2.1: MP2 interaction energies in kcal mol⁻¹ without CP corrections.

Basis Set	Di/Di	Cy/Cy	Di/Cy	Tz/Tz	Bz/Tz
haDZ	-3.12	-2.89	-4.02	-6.79	-8.71
haTZ	-2.82	-2.84	-3.81	-5.51	-7.10
haQZ	-2.65	-2.67	-3.65	-4.92	-6.41
ha5Z	-2.58	-2.59	-3.58	-4.71	-6.17

Table 2.2: Corrections to E_{int} from higher-order correlation effects obtained from CCSD(T) and MP2 calculations ($\delta E_{\text{MP2}}^{\text{CCSD(T)}}$) with the haDZ basis set. All values are in kcal mol⁻¹.

	Di/Di	Cy/Cy	Di/Cy	Tz/Tz	Bz/Tz
CP Uncorrected	+1.10	+0.83	+1.04	+1.43	+2.18
CP corrected	+1.15	+0.87	+1.10	+1.58	+2.35

to see that the $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ corrections are so insensitive to BSSE. To our knowledge, this is the first time this interesting and very useful result has been reported in the literature.

The MP2 CBS limit interaction energies are shown in Table 2.3 along with the non-CP corrected $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ values. These two pieces of data are combined via Equation 1 to give the CCSD(T) CBS limits shown in the last column. The MP2/6-311++G(3df,p) E_{int} of -5.28 kcal mol⁻¹ for the parallel-slipped configuration of the Bz/Tz dimer reported by Massera *et al.* is very similar to our calculated MP2 CBS limit of -5.93 kcal mol⁻¹.⁹² However, the importance of higher-order correlation effects is reflected in the $+2.18$ kcal mol⁻¹ change (or 58% of $E_{\text{int}}^{\text{CCSD(T)/CBS}}$) from the $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ correction.

At both the MP2 and CCSD(T) CBS limits, the mixed dimers (Di/Cy and Bz/Tz) are far more strongly bound than the other dimers. For example, the magnitude of $E_{\text{int}}^{\text{CCSD(T)/CBS}}$ for Di/Cy is roughly 1.7 times larger than for Di/Di and 1.4 times larger than for Cy/Cy.

Table 2.3: MP2 and CCSD(T) CBS interaction energies in kcal mol⁻¹. The $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ corrections are the CP uncorrected values from Table 2.2

Structure	MP2 CBS	$\delta E_{\text{MP2}}^{\text{CCSD(T)}}$	CCSD(T) CBS Limit
Di/Di	-2.52	+1.10	-1.42
Cy/Cy	-2.52	+0.83	-1.69
Di/Cy	-3.51	+1.04	-2.47
Bz/Bz	-4.95 ^a	+2.18 ^b	-2.78 ^c
Tz/Tz	-4.46	+1.43	-3.03
Bz/Tz	-5.93	+2.18	-3.75

^a From CP corrected MP2R12/A values in Table 2 of Ref.⁷⁶

^b From CP corrected $\delta\text{CCSD(T)}/\text{aug-cc-pVDZ}$ values in Table 2 of Ref.⁷⁶

^c Similar values of -5.03, +2.25 and -2.78 kcal mol⁻¹ were reported in Ref.⁸⁸

The same trend is evident for the aromatic ring systems studied. At the CCSD(T) CBS limit, Bz/Tz has an estimated $\pi \cdots \pi$ -stacking interaction energy of -3.75 kcal mol⁻¹, which is larger than the corresponding values for Bz/Bz (-2.78 kcal mol^{-176,88}) and Tz/Tz (-3.03 kcal mol⁻¹).

The SCS-MP2 and SCSN-MP2 interaction energies can be found in Table 2.4. Both methods give very similar results that are much closer to the CCSD(T) CBS limit than the MP2 data in Table 2.1. The spin component scaled interaction energies obtained with the haQZ basis set for Di/Di, Cy/Cy and Di/Cy lie within 0.3 kcal mol⁻¹ of the CCSD(T) CBS limits. When the haTZ basis set is used, this deviation increases (to more than 1 kcal mol⁻¹ for Bz/Tz), and the SCSN-MP2 method slightly outperforms the SCS-MP2 method. Despite overbinding by as much as 20–30% for Bz/Tz with the haTZ basis set, the SCSN-MP2 and SCS-MP2 methods yield a qualitatively correct description of E_{int} and reproduce the CCSD(T)/CBS trends. The improved performance with the haQZ basis set is encourag-

Table 2.4: SCS-MP2 and SCSN-MP2 interaction energies in kcal mol⁻¹.

		haDZ	haTZ	haQZ
Di/Di	SCS-MP2	-2.16	-1.78	-1.57
	SCSN-MP2	-1.72	-1.62	-1.58
Cy/Cy	SCS-MP2	-2.27	-2.19	-1.99
	SCSN-MP2	-1.97	-1.98	-1.93
Di/Cy	SCS-MP2	-3.12	-2.84	-2.64
	SCSN-MP2	-2.81	-2.77	-2.73
Tz/Tz	SCS-MP2	-5.02	-3.68	-
	SCSN-MP2	-4.11	-3.44	-
Bz/Tz	SCS-MP2	-6.49	-4.82	-
	SCSN-MP2	-5.38	-4.40	-

ing given that these methods were designed for use with the aug-cc-pVTZ basis set. This observation is consistent with the results of Antony and Grimme.¹¹⁴

The SAPT2 results are shown in Table 2.5. It is interesting to note that although Di/Di and Cy/Cy have very similar interaction energies, the components of E_{int} are very different. The dispersion contribution for Cy/Cy is approximately 1.8 kcal mol⁻¹ less attractive than for Di/Di with both the haDZ and haTZ basis sets. This large change is offset by an equally large change in the exchange repulsion which is approximately 1.8 kcal mol⁻¹ less repulsive for Cy/Cy than for Di/Di. The SAPT2 calculations also reveal that the electrostatic component is largely responsible for the increased interaction energy in the Di/Di \rightarrow Cy/Cy \rightarrow Di/Cy series. Di/Cy has an electrostatic contribution that exceeds 2.5 kcal mol⁻¹, whereas Cy/Cy

Table 2.5: Contributions to the interaction energies (E_{int}) from SAPT2 calculations with the haDZ basis set. Values given in parentheses are E_{int} obtained with haTZ basis set. All values are in kcal mol⁻¹.

	Di/Di	Cy/Cy	Di/Cy	Bz/Bz ^a	Tz/Tz	Bz/Tz
Electrostatic	-1.80 (-1.63)	-1.95 (-1.85)	-2.64 (-2.52)	-1.85	-2.95	-5.46
Exchange	+4.94 (+4.72)	+3.07 (+2.93)	+4.67 (+4.46)	+2.93	+8.72	+12.77
Induction	-0.63 (-0.61)	-0.38 (-0.37)	-0.72 (-0.70)	-0.37	-0.86	-1.39
Dispersion	-4.67 (-5.10)	-2.96 (-3.23)	-4.39 (-4.81)	-3.23	-8.58	-10.97
E_{int}	-2.16 (-2.62)	-2.22 (-2.52)	-3.07 (-3.58)	-2.52	-3.67	-5.05

^a CCSD(T)/aug-cc-pVQZ* parallel-displaced equilibrium geometry taken from Ref.⁷⁹

and Di/Di have values less than 2.0 kcal mol⁻¹.

In light of the results obtained for the smaller, linear $\pi \cdots \pi$ stacking prototypes, it is tempting to correlate E_{int} with the quadrupole moment of the monomers. However, the SAPT2 data for the larger, cyclic prototypes (Bz/Bz, Tz/Tz, Bz/Tz) reiterate that simple electrostatic arguments, such as Hunter-Sanders rules, are not always sufficient to rationalize trends in $\pi \cdots \pi$ stacking interactions. The absolute values for all four components of E_{int} increase steadily in the Bz/Bz \rightarrow Tz/Tz \rightarrow Bz/Tz progression. In each case, both exchange and dispersion provide the largest contributions to E_{int} (in excess of 10 kcal mol⁻¹ in the case of Bz/Tz). Although there is substantial cancellation between the large repulsive exchange

component and the large attractive dispersion component, the absolute value of the former is 1.80 kcal mol⁻¹ larger than the latter in the Bz/Tz dimer. Only after considering the smaller attractive components (electrostatic and induction), does one get a clear picture of energetics for the cyclic dimers. To our knowledge, this is the first time SAPT2 results have been reported for the parallel-slipped Bz/Bz dimer.

2.5 CONCLUSIONS

Estimates of the $\pi \cdots \pi$ stacking interaction energies for a variety of parallel-slipped prototypes at the CCSD(T) CBS limit have been obtained. As expected, MP2 overbinds relative to CCSD(T) energies. The SCS-MP2 and SCSN-MP2 methods yield better results and reproduce trends in CCSD(T)/CBS interaction energies despite overbinding by as much as 20–30% with the haTZ basis set. The introduction of heteroatoms into $\pi \cdots \pi$ stacking systems can cause significant changes to E_{int} . The mixed dimers (Di/Cy and Bz/Tz) have appreciably larger interaction energies compared to their homogeneous counterparts. For the smaller $\pi \cdots \pi$ -type prototypes, the interaction energies increase in the following manner Di/Di < Cy/Cy << Di/Cy. Similarly, for the larger aromatic systems, the interaction energy from weakest to the strongest is Bz/Bz < Tz/Tz << Bz/Tz. SAPT2 calculations demonstrate that this trend in the interaction energies correlates rather well with the electrostatic component of E_{int} . However, closer inspection of the SAPT2 data for the cyclic dimers reveals that the contributions from exchange, induction and dispersion must also be considered to understand the observed trends. Another significant outcome of this work is

that the $\delta E_{\text{MP2}}^{\text{CCSD(T)}}$ correction is very insensitive to BSSE. As a result, CP corrections are not required to determine the higher-order correlation effects on E_{int} .

2.6 ACKNOWLEDGEMENTS

The authors acknowledge the Mississippi Center for Supercomputing Research for access to their computational resources. The authors also thank Dr. Brian W. Hopkins and Adel ElSohly for their valuable suggestions and helpful discussions. This work was financially supported in part by the National Science Foundation (EPS-0132618, CHE-0517067).

CHAPTER 3

CCSD(T) COMPLETE BASIS SET LIMIT RELATIVE ENERGIES FOR LOW-LYING WATER HEXAMER STRUCTURES

3.1 ABSTRACT

MP2 and CCSD(T) complete basis set (CBS) limit relative electronic energies (ΔE_e) have been determined for 8 low-lying structures of the water hexamer by combining explicitly correlated MP2-R12 computations with higher-order correlation corrections from CCSD(T) calculations. Higher-order correlation effects are quite substantial and increases ΔE_e by at least $+0.19$ kcal mol⁻¹ and as much as $+0.59$ kcal mol⁻¹. The effects from zero-point vibrational energy (ZPVE) have been assessed from unscaled harmonic vibrational frequencies computed at the MP2 level with a correlation consistent triple- ζ basis set (cc-pVTZ for H and aug-cc-pVTZ for O). ZPVE effects are even more significant than higher-order correlation effects and are uniformly negative, decreasing the relative energies by -0.16 to -1.61 kcal mol⁻¹. Although it has been widely accepted that the cage becomes the lowest-energy structure after ZPVE effects are included [*Nature*, **1996**, 381, 501–503], the prism is consistently the most stable structure in this work, lying 0.06 kcal mol⁻¹ below the nearly isoenergetic cage isomer at the electronic MP2 CBS limit, 0.25 kcal mol⁻¹ below at the electronic CCSD(T) CBS limit, and 0.09 kcal mol⁻¹ below at the harmonic ZPVE corrected

CCSD(T) CBS limit. Moreover, application of any uniform scaling factor less than unity to correct for anharmonicity further stabilizes the prism and increases the relative energies of the other structures.

3.2 INTRODUCTION

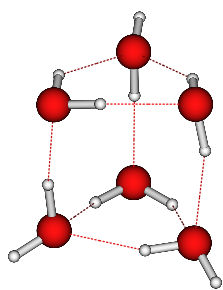
The water hexamer is an important and widely studied water cluster because it represents the crossover point from 2-dimensional to 3-dimensional hydrogen bonding networks.^{115–122} Even with only 6 water molecules, there is a staggering number of possible hydrogen bonding patterns for $(\text{H}_2\text{O})_6$.¹²³ Fortunately, only few fundamental motifs give rise to the most stable structures, and they are labeled with descriptive monikers such as “bag”, “book”, “cage”, “cyclic” and “prism”. Some examples of these structures are shown in Figure 3.1. Note that different isomers can be obtained through subtle changes in the relative orientations of the H atoms not involved in hydrogen bonding. This distinction is particularly important for the book [Figures 3.1(e) and 3.1(f)] and cyclic-boat structures [Figures 3.1(g) and 3.1(h)] where the same name is sometimes used to describe different $(\text{H}_2\text{O})_6$ structures. In this work, book-1 and cyclic-boat-1 are used to denote the conformation of these isomers with the lower electronic energy while the number 2 is appended to the higher-energy structure.

Electronic structure computations indicate that most of the low-lying $(\text{H}_2\text{O})_6$ isomers depicted in Figure 3.1 have very similar electronic energies,^{115,120} and several forms of the water hexamer have been observed experimentally under various conditions.^{116,121,124–130} In some cases, however, definitive assignment of the observed spectra to a particular structure

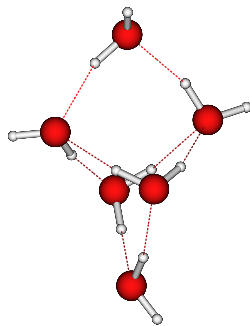
was not possible, potentially due to the presences of multiple isomers.

In recent years, rather sophisticated and demanding electronic structure computations have been performed on these water hexamer isomers to help resolve their relative electronic energies. For example, in 2002, MP2 complete basis set (CBS) limit relative electronic energies were reported for the book, cage, prism and cyclic-chair isomers of $(\text{H}_2\text{O})_6$.¹³¹ Electronically, the prism and cage structures [Figures 3.1(a) and 3.1(b)] were found to be isoenergetic with the cage only 0.07 kcal mol⁻¹ above the prism. The book and cyclic structures were only slightly higher in energy at the MP2 CBS limit, 0.25 and 1.00 kcal mol⁻¹, respectively, above the prism isomer. More recently, two groups have examined higher-order correlation effects in this system by computing CCSD(T) relative electronic energies for water hexamer isomers^{132,133} with correlation consistent triple- ζ basis sets augmented with diffuse functions. The CCSD(T) results are in qualitative agreement with the MP2 CBS data, but there are slight quantitative differences in the relative electronic energies on the order of a few tenths of a kcal mol⁻¹, which is not unexpected given the differences in optimized structures and basis sets. (See the Computational Details section for more detail about the structures examined in References 131, 132 and 133.)

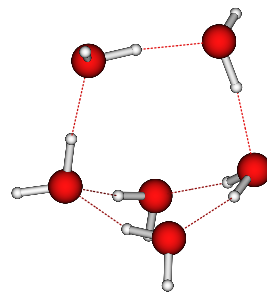
The effects of zero-point vibrational energy (ZPVE)¹³³⁻¹³⁶ and temperature (thermal energy)¹³⁷ on the relative energies of the water hexamer isomers have also been examined. The ZPVE represents a large fraction of the total binding energy and significantly changes the relative energies of the isomers. Temperature can also have a significant effect on the



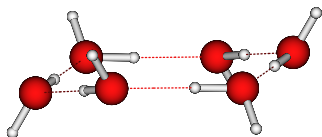
(a) prism



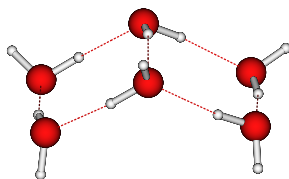
(b) cage



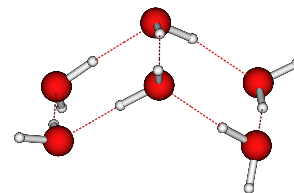
(c) bag



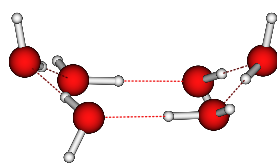
(d) cyclic-chair



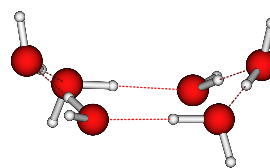
(e) book-1



(f) book-2



(g) cyclic-boat-1



(h) cyclic-boat-2

Figure 3.1: Structures of 8 low-lying $(\text{H}_2\text{O})_6$ isomers.

energetics of the $(\text{H}_2\text{O})_6$ system. For example, in a study of small water clusters that included the prism, cage and cyclic-chair water hexamers, a variety of popular model chemistries (e.g., G2, G3, CBS-APNO) revealed that, while the cage and prism forms of the hexamer are the lowest energy structures at very low temperatures, the cyclic-chair structure structure becomes more favored at higher temperatures.

This work builds on recent high-accuracy electronic structure studies.¹³¹⁻¹³³ Explicitly correlated MP2-R12 energies are combined with higher-order correlation corrections from CCSD(T) calculations to estimate the CBS limit CCSD(T) relative electronic energies (ΔE_e) for the eight water hexamer structures shown in Figure 3.1. Harmonic vibrational frequencies are computed with the MP2 method and a correlation consistent triple- ζ basis set (with diffuse functions on O atoms) and used to examine ZPVE effects on the relative energies (ΔE_0).

3.3 COMPUTATIONAL DETAILS

The authors of References 131, 132 and 133 graciously provided the Cartesian coordinates for their $(\text{H}_2\text{O})_6$ structures which enabled us to correlate them with those shown in Figure 3.1. The book and ring hexamers of Reference 131 correspond to book-1 and cyclic-chair in this work. The book and boat structures of Reference 133 are identical to the book-1 and cyclic-boat-1 isomers in Figure 3.1 while the book and boat structures of Reference 132 correspond to book-2 and cyclic-boat-2 here. The prism, cage and cyclic-chair structures were consistent throughout the studies. This work examines these seven unique structures

as well as the bag isomer from Reference 133.

All structures in this study have been fully optimized at the MP2 level with a correlation consistent triple- ζ basis set (cc-pVTZ basis set for H and the aug-cc-pVTZ for O). This basis set will, hereafter, be denoted haTZ. Cartesian coordinates for the MP2/haTZ optimized structures of the prism, cage, bag, cyclic-chair, book-1 and boat-1 were obtained from the supporting information for Reference 133. The other structures were optimized in this work. It is worth noting that both cyclic-boat structures deviate only slightly from C_2 symmetry. However, if the boat structures are re-optimized at the MP2/haTZ level in C_2 symmetry, the electronic energies increase by approximately $24 \mu E_h$ ($0.015 \text{ kcal mol}^{-1}$). Residual Cartesian gradients of the optimized structures reported here are less than $1 \times 10^{-4} E_h \text{ bohr}^{-1}$.

Previous studies have shown that the correlation energy converges to the CBS limit slowly when using correlation consistent basis sets.²⁷ However, dramatic progress in the field of explicitly correlated R12 methods now allows one to “bypass the slow convergence of the conventional methods, by augmenting the traditional orbital expansions with small number of terms that depend explicitly on the inter-electronic distance r_{12} .”²⁷ In this work, the MP2 CBS limit relative electronic energies ($\Delta E_e^{\text{MP2/CBS}}$) of the 8 structures are determined with explicitly correlated MP2-R12 computations²² employing the massive K2-- basis set^{138,139} (222 basis functions per monomer, compared to 74 for the haTZ basis set). This procedure provides MP2 CBS limit interaction energies comparable to those obtained with extrapolation schemes for correlation consistent basis sets.^{7,8,76,138,140–142} A correction for higher-order correlation effects was calculated from the difference between the MP2

and CCSD(T) relative energies with the haTZ basis set ($\delta_{\text{MP2}}^{\text{CCSD(T)}}$). Reliable estimates of CCSD(T) CBS limit relative energies ($\Delta E_e^{\text{CCSD(T)/CBS}}$) are routinely obtained by combining terms.^{7, 8, 59, 76, 83, 143–147}

$$\Delta E_e^{\text{CCSD(T)/CBS}} = \Delta E_e^{\text{MP2/CBS}} + \delta_{\text{MP2}}^{\text{CCSD(T)}} \quad (3.1)$$

In all MP2, MP2-R12 and CCSD(T) computations, the 1s-like core orbitals of O were excluded from the correlation procedure (i.e., the frozen core approximation). The geometry optimizations and MP2-R12 calculations were performed with the **MPQC** software package,^{26, 105} and the latter employed the A' resolution of the identity approximation.²⁴ Harmonic vibrational frequencies were obtained with the analytical MP2 Hessians available in **Gaussian 03**.¹⁴⁸ Finally, the CCSD(T) computations were performed with the **MOLPRO**¹⁴⁹ and **PSI3**¹⁵⁰ programs. Electronic energies were converged to at least $1 \times 10^{-7} E_h$ in all single point energy computations. Counterpoise (CP) corrections^{15, 151} for basis set superposition error (BSSE)^{14, 152} were not applied because (i) the MP2-R12/K2— energies are essentially at the CBS limit where BSSE is zero by definition, and (ii) the $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ correction for higher-order correlation effects in weakly bound non-covalent clusters is rather insensitive to BSSE.^{7, 8, 153}

3.4 RESULTS

3.4.1 RELATIVE ELECTRONIC ENERGIES

The MP2 CBS limit relative electronic energies ($\Delta E_e^{\text{MP2/CBS}}$) of the 8 water hexamer structure are given in the second column of Table 3.1. The $\Delta E_e^{\text{MP2/CBS}}$ values for the cage, book-1, cyclic-chair and 6 isolated monomers (i.e., the electronic dissociation energy, D_e) are virtually identical to the MP2 CBS limits of 0.1, 0.3, 1.1 and 45.9 kcal mol⁻¹, respectively, obtained in an earlier study by Xantheas, Burnham and Harrison who applied a customized extrapolation procedure to both CP corrected and un-corrected MP2/aug-cc-pVXZ energies ($X=D,T,Q,5$).¹³¹ The two sets of results not only suggest that the MP2 CBS results are converged to 0.1 kcal mol⁻¹, but they also confirm that CP corrections need not be applied to the MP2-R12/K2— relative energies.

Corrections for higher-order correlation effects from MP2 and CCSD(T) computations ($\delta_{\text{MP2}}^{\text{CCSD(T)}}$) with the haTZ basis set are presented in the third column of Table 3.1. Note that these corrections significantly increase the energies of the (H₂O)₆ isomers relative to the prism. This stabilization of the prism isomer by the CCSD(T) method with respect to MP2 relative energies is consistent with other studies of the water hexamer at the CCSD(T) level.^{132,133} All of the $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ values are positive, increasing ΔE_e in an absolute sense by +0.19 kcal mol⁻¹ to +0.59 kcal mol⁻¹ and in a relative sense by 25% to 316% but having almost no effect on D_e (+0.06 kcal mol⁻¹ or <0.02%). Although other studies of other weakly bound complexes have observed that the $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ term is quite insensitive to BSSE,^{7,8,153} the CP corrected D_e of the prism isomer was computed to demonstrate this trend holds

Table 3.1: Higher-order correlation effects, $\delta_{\text{MP2}}^{\text{CCSD(T)}}$, and relative electronic energies, ΔE_e , at the MP2 and CCSD(T) CBS limits for the eight $(\text{H}_2\text{O})_6$ structures.^a

Structure	$\Delta E_e^{\text{MP2/CBS}}$	$\delta_{\text{MP2}}^{\text{CCSD(T)}}$	$\Delta E_e^{\text{CCSD(T)/CBS}}$	$\Delta E_e^{\text{CCSD(T)/haTZ}^b}$	$\Delta E_e^{\text{CCSD(T)/haTZ}^c}$
prism	0.00	+0.00	0.00	0.00	0.00
cage	0.06	+0.19	0.25	0.21	0.28
bag	1.23	+0.39	1.62	1.57	...
cyclic-chair	1.21	+0.59	1.80	1.83	1.81
book-1	0.33	+0.39	0.72	0.71	...
book-2	0.64	+0.39	1.02	...	1.06
cyclic-boat-1	2.20	+0.59	2.79	2.84	...
cyclic-boat-2	2.28	+0.57	2.85	...	2.99
6 monomers	45.86 ^d	+0.06	45.92 ^d	46.71 ^d	46.6 ^d

^a All values in kcal mol⁻¹

^b Reference 132

^c Reference 133

^d D_e of the prism

true for the water hexamer. When a CP correction is applied, the higher-order correlation correction changes by 0.01 kcal mol⁻¹ from +0.06 to +0.07 kcal mol⁻¹ with the haTZ basis set.

The CCSD(T) CBS limit relative electronic energies were obtained by applying Equation 3.1 to the data in the 2nd and 3rd columns of Table 3.1, and these results are given in the 4th column. One of the most interesting features of these data is that the near degeneracy of the prism and cage structures at the MP2 CBS limit is lifted at the CCSD(T) CBS limit. Although they are virtually isoenergetic at the former limit, the prism is 0.25 kcal mol⁻¹ more stable than the cage at the CCSD(T) CBS limit.

The $\Delta E_e^{\text{CCSD(T)/CBS}}$ data are in nearly perfect agreement with the CCSD(T)/haTZ rel-

ative electronic binding energies reported in References 132 and 133 (shown in the last two columns of Table 3.1). Only for the cyclic-boat-2 structure do results differ by more than 0.05 kcal mol⁻¹, and in that case, the deviation still does not exceed 0.14 kcal mol⁻¹. Even the CCSD(T) dissociation energies of the prism reported in the table differ by less than 0.8 kcal mol⁻¹, which corresponds to a relative difference of less than 2%.

3.4.2 ZPVE INCLUSIVE RELATIVE ENERGIES

As mentioned in the Introduction, the zero-point vibrational energy (ZPVE) can significantly affect the relative energies of these isomers. The first column of data in Table 3.2 lists the effect of ZPVE (to 2 decimal places for consistency) on the relative energies of the isomers obtained from unscaled MP2/haTZ harmonic vibrational frequencies. These MP2/haTZ δ_{ZPVE} terms are added to the $\Delta E_e^{\text{CCSD(T)}/\text{CBS}}$ values from Table 3.1 to obtain ZPVE corrected relative energies at the CCSD(T) CBS limit (ΔE_0), which are listed in the last column of Table 3.2.

The δ_{ZPVE} corrections are all negative, decreasing the energies of the isomers relative to the prism structure by as little as -0.16 kcal mol⁻¹ for the cage and by as much as -1.61 kcal mol⁻¹ for the cyclic-boat-2 structure. The ZPVE has a much larger absolute effect on the dissociation energy, -13.71 kcal mol⁻¹. Despite these significant negative corrections, the prism remains the lowest energy isomer at the ZPVE corrected CCSD(T) CBS limit. In a sense, the δ_{ZPVE} shifts essentially reverse the effects of the $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ corrections. For example, $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ increases the relative energy of the cage by $+0.19$ kcal mol⁻¹, while δ_{ZPVE} pushes it back down by -0.16 kcal mol⁻¹. As a result, the cage is, once again, virtually isoenergetic

with the prism at the harmonic ZPVE corrected CCSD(T) CBS limit ($\Delta E_0 = +0.09$ kcal mol⁻¹). The ZPVE corrections to the relative energies of the other isomers are even more pronounced, which effectively compresses the energetic spectrum of (H₂O)₆ structures. The electronic energies of the 8 isomers are separated by 2.85 kcal mol⁻¹ at the CCSD(T) CBS limit but only by 1.28 kcal mol⁻¹ after MP2/haTZ harmonic ZPVE effects are included.

Table 3.2: Harmonic ZPVE corrections, δ_{ZPVE} , and ZPVE corrected CCSD(T) CBS limit relative energies, ΔE_0 , for the eight (H₂O)₆ structures.^a

Structure	δ_{ZPVE}	$\Delta E_0^{\text{CCSD(T)/CBS}}$	δ_{ZPVE}^b	δ_{ZPVE}^c	δ_{ZPVE}^d
prism	+0.00	0.00	+0.00	+0.00	+0.00
cage	-0.16	0.09	-0.35	-0.20	+1.10
bag	-0.78	0.84	+0.49
cyclic-chair	-1.29	0.51	-1.59	-1.18	+1.25
book-1	-0.51	0.21	+0.76
book-2	-0.54	0.48	...	-0.58	...
cyclic-boat-1	-1.51	1.28	-1.89	...	-0.23
cyclic-boat-2	-1.61	1.24	...	-1.49	...
6 monomers	-13.71	32.21 ^e	...	-13.85	...

^a All values in kcal mol⁻¹

^b HF/6-311G(*d*, *p*) values from Reference 134

^c MP2/aug-cc-pVDZ values from Reference 135

^d MP2/haTZ values from Reference 133

^e D_0 of the prism

While the large amplitude vibrational motions in weakly bound, non-covalent clusters tend to be highly anharmonic, this anharmonicity will not likely lead to qualitative changes in relative energetics of the (H₂O)₆ structures examined here. Appropriate empirical scaling factors are a popular and straightforward means to estimate the anharmonic ZPVE.^{154, 155} In

this particular case, all of the harmonic δ_{ZPVE} terms in Table 3.2 are negative. Consequently, any scaling factor less than unity will stabilize the prism and increase ΔE_0 of the other isomers. Only with frequency scaling factors greater than unity could another isomer end up with a ZPVE inclusive energy lower than that of the prism. In fact, a scaling factor > 1.39 is required to produce an isomer with an energy that is lower than that of the prism. (See figure in Appendix 7.2) Typical ZPVE scaling factors for MP2 harmonic vibrational frequencies are slightly less than unity (≈ 0.95). Consequently, corrections for anharmonicity are not likely to change the overall conclusions drawn from the ΔE_0 values reported in Table 3.2. While diffusion quantum Monte Carlo (DQMC) calculations with the VRT(ASP-W)III potential also predict that ZPVE stabilizes the cage isomer,¹⁵⁶ the effect is an order of magnitude larger ($-1.6 \text{ kcal mol}^{-1}$) than the harmonic value reported here. However, the DQMC ZPVE corrections were obtained utilizing a 2-body (with many-body polarization components), rigid monomer potential fit to experimental microwave and far-IR transitions for $(\text{D}_2\text{O})_2$, and direct comparison to our harmonic ZPVE data for fully flexible monomers is not entirely rigorous.

The MP2/haTZ ZPVE corrections reported in the 2nd column of Table 3.2 are very similar to those in the 4th and 5th columns from HF/6-311G(d, p) and MP2/aug-cc-pVDZ computations, respectively.^{134, 135} All three sets of δ_{ZPVE} data are uniformly negative, which indicates the same overall effects from ZPVE and leads to consistent conclusions (i.e., $\Delta E_0 < \Delta E_e$). These results are in stark contrast to a recent study of the $(\text{H}_2\text{O})_6$ system where the corresponding ΔE_0 values were almost always larger ΔE_e when computed with a va-

riety of density functional theory (DFT) techniques methods as well as the MP2 method with the haTZ basis set.¹³³ In attempt to resolve this discrepancy, we have also computed ZPVE corrections to the relative energies of the (H₂O)₆ isomers using five of the same DFT method/basis set combinations. These DFT δ_{ZPVE} results are reported in Appendix 7.2 and are consistent with the MP2 data from this work as well as the δ_{ZPVE} values from References 134 and 135. Although scaling factors were used Reference 133 to determine the ZPVE corrected relative binding energies, they cannot account for discrepancies in the sign of δ_{ZPVE} in situations where the same method and basis set have been used to compute the harmonic vibrational frequencies (*vida supra*). While we can readily reproduce the electronic energies reported in Reference 133, we have, as yet, not been able to reproduce their ZPVE corrected data. Therefore, raw electronic and ZPVE inclusive energies are provided in the Appendix 7.2 to support the data reported here.

3.5 CONCLUSIONS

The MP2 and CCSD(T) CBS limit relative energies for 8 low-lying structures of the water hexamer have been presented. Although the prism is the lowest-energy structure at both limits, the energies of the other structures relative to the prism (ΔE_e) increase significantly when higher-order correlation effects are included. The $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ correction increases ΔE_e by at least +0.19 kcal mol⁻¹ for the cage isomer and by as much as +0.59 kcal mol⁻¹ for the cyclic-chair and cyclic-boat-1 structures. Only when computing D_e of the prism do higher-order correlation effects have a negligible effect. The CCSD(T) electronic dissociation energy

of the prism differs from the MP2 value by only $+0.06$ kcal mol⁻¹.

Corrections for ZPVE (δ_{ZPVE}) from MP2/haTZ harmonic vibrational frequencies have the opposite sign of those for higher-order correlation effects ($\delta_{\text{MP2}}^{\text{CCSD(T)}}$) and tend to be somewhat larger. The δ_{ZPVE} terms decrease ΔE_e by at least -0.16 kcal mol⁻¹ for the cage structure and as much as -1.61 kcal mol⁻¹ for the cyclic-boat-2 isomer. Thus, the ZPVE effectively compresses the energetic separation between the 8 isomers. At the CCSD(T) CBS limit, the largest $\Delta E_e = +2.85$ kcal mol⁻¹ while the maximum $\Delta E_0 = +1.28$ kcal mol⁻¹.

Despite the significant corrections from higher-order correlation effects and ZPVE, the relative energetics of these (H₂O)₆ isomers are qualitatively similar at the electronic and ZPVE corrected CCSD(T) CBS limits. The prism is consistently the lowest energy structure, and the cage is nearly isoenergetic with the prism ($\Delta E_e = +0.25$ kcal mol⁻¹ and $\Delta E_0 = +0.09$ kcal mol⁻¹). The book isomers are slightly higher in energy. The bag and cyclic-chair structures are a bit further up the energetic spectrum while the cyclic-boats are consistently the highest-energy structures examined in this work.

Since the 1996 *Nature* paper by Liu, *et al.*, it has been widely accepted that the cage becomes the most stable structure after ZPVE effects are included.¹¹⁷ In contrast, this work indicates that ZPVE corrections do not change the energetic ordering of the minima as long as sufficiently sophisticated electronic structure techniques are employed to capture higher-order correlation effects. In light of the data presented here, it is certainly reasonable to expect that the prism and cage structures (and even the book-1 isomer) would be observed in very low temperature experiments. However, one must hesitate from concluding that the

prism, for example, is the most “stable” structure given the fleeting nature of these $(\text{H}_2\text{O})_6$ species.¹⁵⁷

3.6 ACKNOWLEDGMENTS

The authors acknowledge the Mississippi Center for Supercomputing Research for CPU time and the National Science Foundation for support (CHE-0517067 and EPS-0556308). We also thank the authors of References 131, 132 and 133 for providing the Cartesian coordinates of their $(\text{H}_2\text{O})_6$ structures.

3.7 NOTE ADDED IN PROOF

During the review process, a closely related work¹⁵⁸ was published that reports benchmark electronic energies from diffusion Monte Carlo (DMC) computations. The DMC relative energies of $+0.84 \text{ kcal mol}^{-1}$, $+1.43 \text{ kcal mol}^{-1}$ and $+3.88 \text{ kcal mol}^{-1}$ for the cage, book-1 and cyclic-chair structures, respectively, are approximately 2 times larger than the corresponding the $\Delta E_e^{\text{CCSD(T)}/\text{CBS}}$ values reported in Table 3.1. Yet, both sets of relative energies are consistent to within the statistical errors of the DMC computations. Combining the DMC electronic energies with our harmonic ZPVE corrections from Table 3.2 leads to the prism being significantly more stable than the cage ($0.68 \text{ kcal mol}^{-1}$ versus our $\Delta E_0^{\text{CCSD(T)}/\text{CBS}}$ value of $0.09 \text{ kcal mol}^{-1}$).

CHAPTER 4

DEVELOPMENT OF A 3-BODY:MANY-BODY INTEGRATED FRAGMENTATION METHOD FOR WEAKLY BOUND CLUSTERS AND APPLICATION TO WATER CLUSTERS $(\text{H}_2\text{O})_n$

$n=3-9$

4.1 ABSTRACT

A 3-body:Many-body integrated QM:QM fragmentation method for non-covalent clusters is introduced within the ONIOM formalism. The technique captures all 1-, 2- and 3-body interactions with a high-level electronic structure method while a less demanding low-level method is employed to recover 4-body and higher-order interactions. When applied to more than 40 low-lying $(\text{H}_2\text{O})_n$ isomers ranging in size from $n= 3$ to 10, the CCSD(T):MP2 3-body:many-body fragmentation scheme deviates from the full CCSD(T) interaction energy by no more than $0.07 \text{ kcal mol}^{-1}$ (or $0.007 \text{ kcal mol}^{-1}$ per water). The CCSD(T):MP2 procedure is also very efficient because the CCSD(T) computations only need to be performed on subsets of the cluster containing 1, 2 or 3 fragments, which in the current context means the largest CCSD(T) calculations are for 3 water molecules regardless of the cluster size.

4.2 INTRODUCTION

Non-covalent, weakly interacting clusters have been and continue to be of significant interest to many researchers. Because of the vast influence weak non-covalent interactions have on many phenomena, it is highly desirable to find a way to accurately predict structures, energetics and chemical properties for such interactions. These interactions, which are generally weaker than a typical chemical bond, influence many important processes such as solvation, crystallization, absorption, bulk-phase properties. For example, hydrogen-bonding and π -stacking interactions are responsible for the structure of the double-helix structure of DNA and RNA.¹⁵⁹

Water is arguably the most important solvent because of its significant role in Nature, and water clusters are the standard model for understanding hydrogen bonding. An accurate computational description of hydrogen bonding often requires sophisticated electronic structures methods for which computational demands scale steeply with system size and prohibit their routine application to larger systems. In recent years, several groups have extended the size of water clusters that can now be analyzed with these demanding model chemistries. The N-body decomposition (NBD) method, developed by Jordan in 2005, decomposes the interaction energy into 1, 2, \dots , N-body components and truncates the expansion to reduce the computational costs for application to larger systems.⁴⁰ In addition to a many-body expansion, the fragment molecular orbital (FMO) method also embeds each fragment in an electrostatic field from the whole system.^{30,31,36,160,161} *Ab initio* integrated multi-center molecular orbital method developed by Sakai and Morita in 2005 can be used to calcu-

late geometrical parameters, total energies, relative energies and vibrational frequencies for large water clusters that deviate only slightly from results obtained with the full *ab initio* method.¹⁶² Several other methods have also been developed for studying weakly interacting (non-covalent) clusters, such as Truhlar’s electrostatically embedded many-body (EE-MB) expansion,^{41,163,164} the elongation method,¹⁶⁵ the systematic fragmentation method (SFM),¹⁶⁶ molecular fragmentation with conjugated caps (MFCC),^{167,168} molecular tailoring approach¹⁶⁹ and many more.^{35,47,48,170–172} Readers interested in a more detailed overview of methods for studying large molecular systems and a thorough explanation of SFM, FMO and EFP methods are directed to Ref. 46 and references therein.

In this paper, we first review the development of our fragmentation method for clusters^{49,50,173} within the ONIOM framework of Maseras and Morokuma.¹⁷⁴ Then we outline the extension of this approach to a very accurate and efficient 3-body:Many-body QM:QM implementation. Finally, the procedure is applied to a series of water clusters ranging in size from 3 – 10 water molecules. We examine more than 40 different isomers for the clusters some of which are separated by less than 0.2 kcal mol⁻¹ electronically.

4.3 THEORETICAL METHODS

For size-consistent methods, the interaction energy of a weakly bound cluster, denoted ΔE in this work, is the energy of the complex, $E[f_1 f_2 \dots f_n]$, relative to the energy of the isolated fragments, f_i^* , where n is the total number of monomers within the cluster. (An asterisk

denotes a fragment at its optimal geometry in isolation.)

$$\Delta E = E[f_1 f_2 \dots f_n] - \sum_{i=1}^n E[f_i^*] \quad (4.1)$$

Because of the unfavorable scaling of high-level *ab initio* methods, which are often necessary to reliably determine ΔE for weakly bound clusters, routine application to large clusters is unfeasible. However, Xantheas and coworkers demonstrated for large water clusters, (up to $n = 24$), CCSD(T)/aug-cc-pVTZ calculations are possible with highly scalable software and massively parallel architectures.¹⁷⁵ In an attempt to overcome this obstacle, approximations are sometimes invoked. For example, the 2-body approximation (ΔE^{2b}) for weakly bound clusters is a simple summation of all the interaction energies of each unique pair of fragments within a cluster that is corrected for redundant 1-body terms.

$$\begin{aligned} \Delta E^{2b} &= \sum_{i=1}^n \sum_{j>i}^n E[f_i f_j] \\ &\quad - (n-2) \sum_{i=1}^n E[f_i] \\ &\quad - \sum_{i=1}^n E[f_i^*] \end{aligned} \quad (4.2)$$

A 2-body approximation can be very accurate when cooperative effects are relatively small ΔE . In 2005, for example, Tauer and Sherrill demonstrated that a simple 2-body approximation recovers 98% total interaction energy when applied to the the benzene tetramer.¹⁷⁶ In this case, the 2% error arises from the neglect of 3- and 4-body effects. In general, this

deviation from pairwise additivity is known as the non-additivity or cooperativity, which will hereafter be denoted as $\delta E^{\geq 3b}$.

$$\Delta E = \Delta E^{2b} + \delta E^{\geq 3b} \quad (4.3)$$

An analogous 3-body approximation to ΔE can be obtained from the interaction energies of each unique triad of fragments within the cluster after correcting for redundant 1- and 2-body contributions.

$$\begin{aligned} \Delta E^{3b} = & \sum_{i=1}^n \sum_{j>i}^n \sum_{k>j}^n E[f_i f_j f_k] \\ & - (n-3) \sum_{i=1}^n \sum_{j>i}^n E[f_i f_j] \\ & + \frac{(n-3)(n-2)}{2} \sum_{i=1}^n E[f_i] \\ & - \sum_{i=1}^n E[f_i^*] \end{aligned} \quad (4.4)$$

The multicentered (MC) approach^{177,178} developed by our group for integrated computational methods allows one to recast this many-body approximation within the ONIOM formalism.^{49,50,173} In the initial 2-body:Many-body implementation of this integrated fragmentation for non-covalent clusters a high-level electronic structure method is used to compute up through the 2-body interactions while a low-level method recovers the higher-order interactions. In practice, a low-level calculation is applied to the entire system ($E_{Lo}[f_1 f_2 \dots f_n]$) while high-level calculations only need to be performed for all the unique pairs and the frag-

ments ($E_{\text{Hi}}[f_i f_j]$ and $E_{\text{Hi}}[f_i]$) within the cluster. The resulting 2-body:Many-body interaction energy is:

$$\begin{aligned}
\Delta E^{2\text{bHi:Lo}} &= E_{\text{Lo}}[f_1 f_2 \dots f_n] \\
&+ \sum_{i=1}^n \sum_{j>i}^n [E_{\text{Hi}}[f_i f_j] - E_{\text{Lo}}[f_i f_j]] \\
&- (n-2) \sum_{i=1}^n [E_{\text{Hi}}[f_i] - E_{\text{Lo}}[f_i]] \\
&- \sum_{i=1}^n E_{\text{Hi}}[f_i^*] \tag{4.5}
\end{aligned}$$

where an inclusion-exclusion principle has been applied to remove redundant contributions. More details can be found in Refs.,¹⁷³⁴⁹ and.⁵⁰ Note that unlike the 2-body approach (Eq. 4.2), the 2-body:many-body expression in Eq. 4.5 recovers the higher-order effects ($\delta E^{\geq 3b}$) at the low-level. Consequently, the error associated with a particular 2-body:Many-body QM:QM procedure is the difference between the high-level and low-level non-additivity. When a 2-body:many-body QM:QM fragmentation approach was applied to weakly bound clusters of He, Ne, HF, and water, results showed only a 1% error when comparing interaction energies to the high-level calculation.⁴⁹

The 2-body:Many-body fragmentation method offers an accurate and efficient approach for extending sophisticated electronic structure theory methods to larger clusters and introduces errors on order of a few tenths of a kcal mol⁻¹.⁴⁹ However, greater accuracy can be required in certain pathological cases where nearly isoenergetic isomers of a cluster may only be separated by 0.1 kcal mol⁻¹. Take, for example, the water hexamer where the prism and

cage isomers are only separated electronically by 0.06 kcal mol⁻¹ at the MP2 CBS limit.^{131,179}

To improve the accuracy of our QM:QM fragmentation method, we have extended the 2-body:many-body approach to a 3-body:Many-body procedure. Here the high-level method captures the 1-, 2- and 3-body interactions with the low-level method is used to describe the interactions of 4th order and higher.

$$\begin{aligned}
\Delta E^{3\text{bHi:Lo}} &= E_{\text{Lo}}[f_1 f_2 \dots f_n] \\
&+ \sum_{i=1}^n \sum_{j>i}^n \sum_{k>j}^n [E_{\text{Hi}}[f_i f_j f_k] - E_{\text{Lo}}[f_i f_j f_k]] \\
&- (n-3) \sum_{i=1}^n \sum_{j>i}^n [E_{\text{Hi}}[f_i f_j] - E_{\text{Lo}}[f_i f_j]] \\
&+ \frac{(n-2)(n-3)}{2} \sum_{i=1}^n [E_{\text{Hi}}[f_i] - E_{\text{Lo}}[f_i]] \\
&- \sum_{i=1}^n E_{\text{Hi}}[f_i^*] \tag{4.6}
\end{aligned}$$

Again an inclusion-exclusion principle is applied to correct for redundancies from overlapping subsets of fragments. The resulting procedure effectively reduces the high-level calculation on the entire cluster to a series of computations on the unique triads, pairs and fragments in the cluster i.e., $E_{\text{Hi}}[f_i f_j f_k]$, $E_{\text{Hi}}[f_i f_j]$ and $E_{\text{Hi}}[f_i]$.

4.4 COMPUTATIONAL METHODS

All structures have been optimized at the MP2 level of theory with a triple- ζ basis set, aug-cc-pVTZ for O and cc-pVTZ for H (henceforth denoted as haTZ). Cartesian coordinates for all $(\text{H}_2\text{O})_n$ structures were taken from Refs.,¹⁸⁰¹⁷⁹ and.¹³³ Cartesian coordinates for all

structures are provided in the Supporting Information.¹⁸¹ Residual Cartesian gradients of the optimized structures reported here are less than $1 \times 10^{-4} E_h \text{ bohr}^{-1}$.

Electronic energies of all clusters and components associated with Eqs. 4.5 and 4.6 (triads, pairs and fragments) have been computed with Hartree-Fock (HF) method, second-order Møller-Plesset perturbation theory (MP2) and the coupled cluster method CCSD(T). All single-point energy calculations also employed the haTZ basis set. The electronic energies have then been combined to compute QM:QM interaction energies according to both Eqs. 4.5 and 4.6 (CCSD(T):HF and CCSD(T):MP2). All coupled cluster energies were converged at least $1 \times 10^{-7} E_h$. The $1s$ -like core orbitals of the oxygen atoms were frozen for all MP2 and CCSD(T) calculations. All computations were performed with Gaussian03, Molpro, PQS and MPQC software packages.^{105,148,149}

Counterpoise (CP) corrections^{15,151} for basis set superposition error (BSSE)¹⁵² were not performed in this study. Previous studies have shown that CP corrections dramatically increase the basis set incomplete error (BSIE) associated with the haTZ basis set for hydrogen bonding systems.^{182,183} Furthermore, the QM:QM schemes that provide the most accurate interaction energies without CP corrections tend to also provide the most accurate CP corrected interaction energies.^{184,185}

4.5 RESULTS AND DISCUSSION

4.5.1 2-BODY:MANY-BODY APPROXIMATION

For comparison, MP2 and CCSD(T) interaction energies computed for all water clusters with the haTZ basis set are reported in Table 1. The last column of Table 1 reports MP2 deviations from the CCSD(T) ΔE values. In a relative sense the MP2 method is quite accurate and never differs from the ΔE at the CCSD(T) level by more than 1%. However, the absolute errors can exceed $0.5 \text{ kcal mol}^{-1}$ which can be quite significant when the isomers are separated by less than a few tenths of a kcal mol^{-1} .

Table 2 reports deviations relative to the CCSD(T) interaction energies from Table 1 for all QM:QM approaches used in this study, as well as the 2-body and 3-body approximations with the CCSD(T) method. As expected, the performance of the 2-body approximation (first column of data) is abysmal due to the significant cooperative effects in these systems. The 2-body errors (i.e., $\delta E^{\geq 3b}$ in Eq. 4.2) grows as large $23.33 \text{ kcal mol}^{-1}$ in an absolute sense (for the Global Min isomer of the decamer) or 30% in a relative sense (for isomer C_5 of the pentamer). These results are consistent with the work of Xantheas who demonstrated that the inclusion of 3-body terms are imperative for predicting accurate interaction energies for water clusters.¹⁸⁶

The $\delta E^{\geq 3b}$ errors for the 2-body:Many-body approaches are reported in the second and third columns of data in Table 2. With HF as a low-level calculation, the largest absolute error is $0.15 \text{ kcal mol}^{-1}$ (for isomer D_{2d} of the octamer) and 0.5% in a relative sense (for isomer C_{4h} of the tetramer). The largest absolute error for the 2-body:Many-body approach

with MP2 as the low-level is $0.23 \text{ kcal mol}^{-1}$ (for isomer S_4 of the octamer and D of the water nonamer) or in a relative sense $<1\%$ of the total interaction energy for the prism structure of the water hexamer.

It is interesting to note that within the 2-body:Many-body fragmentation scheme where CCSD(T) is the high-level method, the low-level MP2 errors are often larger than the low-level HF errors for the larger, 3-dimensional clusters. This result indicates that HF reproduces the CCSD(T) non-additivity ($\delta E^{\geq 3b}$) better than MP2 for some of these systems. Nevertheless, both the HF and MP2 errors associated with the 2-body:Many-body approach are significantly smaller than the 2-body approximation with CCSD(T) method.

Table 4.1: MP2 and CCSD(T) interaction energies (ΔE in kcal mol⁻¹) obtained with the haTZ basis set for various water clusters, as well as the deviation of ΔE MP2 from ΔE CCSD(T).

		ΔE CCSD(T)	ΔE MP2	$\Delta\Delta E$ MP2
$(\text{H}_2\text{O})_3$				
	C_1	-15.86	- 15.93	-0.07
	C_3	-15.04	- 15.16	-0.12
	C_{3h}	-14.39	- 14.57	-0.18
$(\text{H}_2\text{O})_4$				
	S_4	-27.75	- 28.00	-0.26
	C_i	-26.82	- 27.08	-0.26
	C_4	-25.52	- 25.83	-0.31
	C_{4h}	-24.61	- 24.97	-0.37
$(\text{H}_2\text{O})_5$				
	C_1	-36.38	- 36.79	-0.41
	C_5	-33.61	- 34.04	-0.43
	C_{5h}	-32.95	- 33.43	-0.48
$(\text{H}_2\text{O})_6$				
	Prism	-46.71	-46.65	0.06
	Cage	-46.50	-46.64	-0.13
	Book1	-46.00	-46.34	-0.33
	Book2	-45.70	-46.04	-0.34
	Bag	-45.14	-45.47	-0.33
	Boat1	-43.87	-44.40	-0.53
	Boat2	-43.81	-44.32	-0.51
	Cyclic	-44.88	-45.40	-0.53
$(\text{H}_2\text{O})_7$				
	A	-58.23	-58.40	-0.17
	B	-56.49	-56.81	-0.33
	C	-54.43	-54.92	-0.49
	D	-54.39	-54.82	-0.43
	E	-51.77	-52.37	-0.60
$(\text{H}_2\text{O})_8$				
	D_{2d}	-73.85	-74.08	-0.23
	S_4	-73.80	-74.05	-0.25
	C_i	-70.90	-71.09	-0.18
	C_s	-70.02	-70.21	-0.19
	C_2	-70.97	-71.14	-0.17
	B	-69.91	-70.06	-0.15

– continued from previous page				
	A	–69.83	–69.94	–0.12
	C	–69.76	–69.89	–0.13
	Noncubic1	–68.98	–69.23	–0.25
	Noncubic2	–68.60	–68.83	–0.24
(H ₂ O) ₉	A	–83.16	–83.56	–0.40
	C	–78.17	–78.39	–0.22
	D	–76.69	–76.69	0.00
(H ₂ O) ₁₀	Global Min	–94.64	–95.06	–0.42
	Prism1	–94.77	–95.16	–0.23
	Chair	–92.82	–93.05	–0.39
	Butterfly	–88.62	–89.00	–0.37

4.5.2 3-BODY:MANY-BODY APPROXIMATION

The performance of the CCSD(T) 3-body approximation (fourth column of data in Table 2) is dramatically better than the CCSD(T) 2-body approximation. Although, the errors for the former are typically an order of magnitude smaller than those for the latter, the 3-body CCSD(T) errors still grow as large as 2.36 kcal mol^{–1} and are generally too large to reliably discern between the most stable configurations of a cluster. The errors for the 3-body:Many-body approaches are given in the last two columns of Table 2. When HF is employed as the low-level calculation, the largest absolute error is 0.35 kcal mol^{–1} for isomer E of the water heptamer. In a relative sense, the deviations are less than 1% of the total interaction energy. When MP2 is used as the low-level calculation, the largest absolute error is 0.07 kcal mol^{–1} for the Chair isomer of the water decamer, and in a relative sense the error is always less than 0.1% of the total interaction energy for all clusters studied. Unlike the

2-body:Many-Body approach, MP2 consistently outperforms HF as a low-level method for the 3-body:Many-body fragmentation technique when CCSD(T) is the high-level method. Errors for the CCSD(T):MP2 method are typically on the order of a few hundredths of a kcal mol⁻¹ while those for the CCSD(T):HF approach are on the order of a few tenths of a kcal mol⁻¹.

Table 4.2: Errors (in kcal mol⁻¹) relative to ΔE CCSD(T) for various 2-body and 3-body methods obtained with the haTZ basis set.

		2-body CCSD(T):Low			3-body CCSD(T):Low		
		None ^a	HF ^b	MP2 ^c	None ^d	HF ^e	MP2 ^f
(H ₂ O) ₃							
	C ₁	-2.47	0.02	0.04
	C ₃	-2.33	0.06	0.03
	C _{3h}	-1.93	0.11	0.03
(H ₂ O) ₄							
	S ₄	-6.80	-0.03	0.02	-0.58	-0.09	-0.02
	C _i	-6.57	-0.03	0.02	-0.57	-0.09	-0.02
	C ₄	-6.32	0.01	0.01	-0.55	-0.08	-0.02
	C _{4h}	-5.64	0.12	0.01	-0.49	-0.07	-0.01
(H ₂ O) ₅							
	C ₁	-10.38	-0.07	-0.03	-1.33	-0.21	-0.02
	C ₅	-9.77	0.00	-0.03	-1.25	-0.19	-0.02
	C _{5h}	-9.20	0.11	-0.02	-1.15	-0.17	-0.02
(H ₂ O) ₆							
	Prism	-9.39	0.11	0.21	-0.56	-0.05	0.00
	Cage	-9.60	0.10	0.18	-0.50	-0.06	0.01
	Book1	-11.50	-0.07	0.03	-1.14	-0.18	-0.01
	Book2	-11.14	-0.11	0.04	-1.02	-0.16	-0.01
	Bag	-11.55	-0.03	0.07	-1.20	-0.19	-0.01
	Boat1	-13.03	-0.09	-0.05	-1.82	-0.29	-0.02
	Boat2	-13.00	-0.04	-0.04	-1.80	-0.27	-0.02
	Cyclic	-13.62	-0.07	-0.06	-2.00	-0.31	-0.02
(H ₂ O) ₇							
	A	-12.86	0.07	0.19	-0.90	-0.11	0.01
	B	-13.12	0.03	0.11	-1.05	-0.15	0.01
	C	-14.77	-0.13	-0.02	-1.86	-0.30	-0.02
	D	-13.12	-0.11	0.00	-0.97	-0.13	-0.01

– continued from previous page							
	E	-15.83	-0.08	-0.08	-2.36	-0.35	-0.02
(H ₂ O) ₈	D _{2d}	-16.01	0.15	0.21	-0.60	0.09	0.03
	S ₄	-16.48	0.04	0.23	-0.78	-0.02	0.02
	C _i	-15.63	-0.04	0.21	-0.82	-0.02	0.01
	C _s	-15.74	-0.05	0.20	-0.86	-0.02	0.01
	C ₂	-15.20	0.05	0.19	-0.67	0.07	0.02
	B	-14.89	0.01	0.19	-0.71	0.03	0.02
	A	-14.20	0.06	0.18	-0.56	0.09	0.02
	C	-14.59	-0.03	0.20	-0.69	0.00	0.02
	Noncubic1	-14.88	0.13	0.19	-0.88	-0.04	0.04
	Noncubic2	-14.68	0.13	0.20	-0.84	-0.03	0.04
(H ₂ O) ₉	A	-19.83	0.10	0.15	-1.50	-0.11	0.02
	C	-17.09	0.12	0.15	-0.90	-0.08	0.00
	D	-15.06	0.14	0.23	-0.42	0.12	0.03
(H ₂ O) ₁₀	Global Min	-23.33	-0.04	0.15	-1.99	-0.24	0.06
	Prism1	-22.64	0.07	0.14	-1.72	-0.11	0.03
	Chair	-19.81	0.11	0.17	-0.62	0.13	0.07
	Butterfly	-20.54	-0.05	0.22	-1.14	-0.19	0.05
	^a CCSD(T) $\delta E^{\geq 3b}$				^d CCSD(T) $\delta E^{\geq 4b}$		
	^b HF $\delta E^{\geq 3b}$ – CCSD(T) $\delta E^{\geq 3b}$				^e HF $\delta E^{\geq 4b}$ – CCSD(T) $\delta E^{\geq 4b}$		
	^c MP2 $\delta E^{\geq 3b}$ – CCSD(T) $\delta E^{\geq 3b}$				^f MP2 $\delta E^{\geq 4b}$ – CCSD(T) $\delta E^{\geq 4b}$		

4.6 CONCLUSIONS

The 2-body:Many-body integrated QM:QM fragmentation method for non-covalent clusters has been extended to a 3-body:Many-body technique that captures the 1-, 2- and 3-body interactions at the high-level while recovering all 4-body and higher-order interactions with the low-level method. This new 3-body:Many-body fragmentation method improves on the accuracy of the 2-body:Many-body approach. For the (H₂O)_n clusters examined in this work,

where $n= 3 - 10$, the CCSD(T):MP2 3-body:Many-body fragmentation method provides interaction energies that are nearly identical to the CCSD(T) method. Interaction energies computed with these two approaches never differ by more than a total of $0.07 \text{ kcal mol}^{-1}$ regardless of the value of n .

Comparison of the 2-body:Many-body and 3-body:Many-body approaches reveals that HF frequently reproduces the CCSD(T) non-additivity ($\delta E^{\geq 3b}$) better than MP2. The situation, however, reverses for higher-order (4-body and beyond) interactions. MP2 consistently reproduces the CCSD(T) $\delta E^{\geq 4b}$ values to within a few hundredths of a kcal mol^{-1} .

While the computational demands of the 3-body:Many-body CCSD(T):MP2 fragmentation method are more significant than the analogous 2-body:Many-body scheme, they are still usually orders of magnitude less than the full CCSD(T) computation. Because the high-level method is only used to compute the 1-, 2- and 3-body interactions, high-level computations only need to be performed on subsets of the cluster that contain 1, 2 or 3 fragments. Consequently, for the $(\text{H}_2\text{O})_n$ clusters examined here, the extremely accurate CCSD(T):MP2 3-body:Many-body fragmentation method does not require CCSD(T) computations on any system larger than $(\text{H}_2\text{O})_3$ regardless of the size of the cluster.

Analytic derivatives are being developed and implemented for these fragmentation methods to enable the calculation of excited states, optimized geometries, vibrational frequencies and NMR chemical shifts. We are also investigating the use of spatial and energetic thresholds to further improve the performance of these integrated QM:QM fragmentation methods for non-covalent clusters.

4.7 ACKNOWLEDGEMENTS

We would like to acknowledge the Mississippi Center for Super Computing Research for CPU time. We would like to thank Dr. Brian Hopkins for his helpful discussions. This work was financially supported by the National Science Foundation (CHE-0957317 and EPS-0903787). TJ was also supported by the National Science Foundation under award number CHE-0911541, and by the Mildred B. Cooper Chair at the University of Arkansas. Acquisition of the Star of Arkansas supercomputer was supported in part by the National Science Foundation under award number MRI-0722625.

CHAPTER 5

EFFICIENT AND ACCURATE METHODS FOR THE GEOMETRY OPTIMIZATION OF WATER CLUSTERS: APPLICATION OF ANALYTIC GRADIENTS FOR THE 2-BODY:MANY-BODY QM:QM FRAGMENTATION METHOD TO $(\text{H}_2\text{O})_n$, $n = 3 - 10$

5.1 ABSTRACT

The structures of more than 70 low-lying water clusters ranging in size from $(\text{H}_2\text{O})_3$ to $(\text{H}_2\text{O})_{10}$ have been fully optimized with several different quantum mechanical electronic structure methods, including second-order Møller-Plesset perturbation theory (MP2) in conjunction with correlation consistent triple- ζ basis sets (aug-cc-pVTZ for O and cc-pVTZ for H, abbreviated haTZ). Optimized structures obtained with less demanding computational procedures were compared to the MP2/haTZ ones using both MP2/haTZ single point energies and the root mean square (RMS) deviations of unweighted Cartesian coordinates. Based on these criteria, B3LYP/6-31+G($d, 2p$) substantially outperforms both HF/haTZ and MP2/6-31G*. B3LYP/6-31+G($d, 2p$) structures never deviate from the MP2/haTZ geometries by more than $0.44 \text{ kcal mol}^{-1}$ on the MP2/haTZ potential energy surface, whereas the errors associated with the HF/haTZ and MP2/6-31G* structures grow as large as $12.20 \text{ kcal mol}^{-1}$ and $2.98 \text{ kcal mol}^{-1}$, respectively. The most accurate results, however, were obtained with the 2-body:Many-body QM:QM fragmentation method for weakly bound clusters, in which

all 1- and 2-body interactions are calculated at the high-level while a low-level calculation is performed on the entire cluster to capture the cooperative effects (non-additivity). With the haTZ basis set, the MP2:HF 2-body:Many-body fragmentation method generates structures that deviate from the MP2/haTZ ones by 0.01 kcal mol⁻¹ on average and not by more than 0.03 kcal mol⁻¹.

5.2 INTRODUCTION

Hydrogen bonding is widely studied, particularly in water, because of its key role in biological phenomena as well as a plethora of important chemical and physical processes.^{122, 145, 187–191} The characterization of molecular clusters with sophisticated quantum mechanical (QM) electronic structure techniques is often highly desirable.^{7–9, 59, 76, 131, 141, 142, 179, 192–196} High-accuracy computational procedures are frequently necessary to reliably describe the properties (e.g., structures and energetics) of weakly-bound clusters. Such computations can also help unravel the chemical physics of the non-covalent interactions that hold the clusters together. Unfortunately, the computational demands of the most reliable QM methods scale steeply with the size of the cluster, thereby prohibiting their routine application to large systems.

A wide variety of computational techniques have been introduced that partition a cluster into fragments (not necessarily monomers) in an attempt to extend high-accuracy computational methods to previously inaccessible size regimes.^{30, 31, 35, 36, 40, 41, 46–48, 160–165, 167–172} The integrated QM:QM fragmentation methods being developed by our group fall into this cat-

egory, and they facilitate the computation of not only energies but also properties. In this paper, we review the 2-body:Many-body fragmentation method and its analytic gradients. The technique is then used to optimize the geometries of more than 70 $(\text{H}_2\text{O})_n$ clusters where $n = 3 - 10$. The errors associated with these 2-body:Many-body optimized structures are assessed and compared to those obtained with 3 other relatively inexpensive electronic structure methods.

5.3 THEORETICAL BACKGROUND

Through careful application of the inclusion-exclusion principle, integrated computational chemistry methods (QM:QM, QM:MM, ONIOM, etc.) have been extended from systems with a single chemically important subset (or reaction center) to systems with an arbitrary number of subsets that can overlap.^{177,178} With this “multicentered” approach to integrated computations, the traditional many-body energy decomposition for weakly bound clusters has been recast^{49,173} in the ONIOM formalism of Morokuma and co-workers.¹⁷⁴ The result is effectively a QM:QM fragmentation scheme for non-covalent clusters.

In the original 2-body:Many-body implementation,^{49,173} an accurate but computationally demanding high-level QM method is employed to compute the 1- and 2-body interactions within a cluster while a less demanding low-level QM method is used to recover the higher-order (≥ 3 -body) interactions, which are also commonly referred to as the cooperative effects or non-additive effects. Consequently, a high-level calculation on the entire cluster $[f_1 f_2 \dots f_n]$ can be avoided, and high-level computations only need to be performed on the

fragments $[f_i]$ or unique pairs of fragments $[f_i f_j]$ in the cluster. An expression for the total energy of the cluster can then be obtained by combining the high-level electronic energies with low-level computations on the entire cluster as well as the fragments and pairs.

$$\begin{aligned}
E^{2\text{bHi:Lo}} &= E_{\text{Lo}}[f_1 f_2 \dots f_n] \\
&+ \sum_{i=1}^n \sum_{j>i}^n (E_{\text{Hi}}[f_i f_j] - E_{\text{Lo}}[f_i f_j]) \\
&- (n-2) \sum_{i=1}^n (E_{\text{Hi}}[f_i] - E_{\text{Lo}}[f_i])
\end{aligned} \tag{5.1}$$

When an appropriate low-level method is used (i.e., one that accurately reproduced the high-level ≥ 3 -body effects), the method is quite accurate, and errors typically do not exceed 0.2 kcal mol⁻¹. It is also quite efficient because the demands of the high-level computations only increase quadratically with the size of the cluster, and the high-level computations are ideally suited for coarse-grained parallelization. An analogous 3-body:Many-body procedure has also been developed to examine the convergence of the series.¹⁹⁷

$$\begin{aligned}
\Delta E^{3\text{bHi:Lo}} &= E_{\text{Lo}}[f_1 f_2 \dots f_n] \\
&+ \sum_{i=1}^n \sum_{j>i}^n \sum_{k>j}^n (E_{\text{Hi}}[f_i f_j f_k] - E_{\text{Lo}}[f_i f_j f_k]) \\
&- (n-3) \sum_{i=1}^n \sum_{j>i}^n (E_{\text{Hi}}[f_i f_j] - E_{\text{Lo}}[f_i f_j]) \\
&+ \frac{(n-2)(n-3)}{2} \sum_{i=1}^n (E_{\text{Hi}}[f_i] - E_{\text{Lo}}[f_i])
\end{aligned} \tag{5.2}$$

For the 3-body:Many-body CCSD(T):MP2 approach, errors tend to decrease by an order

of magnitude relative to the 2-body:Many-body method, suggesting that the series quickly converges and the error can be systematically controlled.

These QM:QM fragmentation schemes have been developed within the ONIOM framework to facilitate the computation of properties, not just energies. An extremely important feature of the expression for cluster energies in Equations 5.1 and 5.2 is that they are linear with respect to the computed energies. Consequently for a linear operator like the gradient, one obtains analogous expressions for the gradient by taking linear combinations of the appropriate components from a series of high- and low-level gradient calculations. For example, the 2-body:Many-body gradient can be expressed in the following manner.

$$\begin{aligned}
\nabla E^{2\text{bHi:Lo}} &= \nabla E_{\text{Lo}}[f_1 f_2 \dots f_n] \\
&+ \sum_{i=1}^n \sum_{j>i}^n (\nabla E_{\text{Hi}}[f_i f_j] - \nabla E_{\text{Lo}}[f_i f_j]) \\
&- (n-2) \sum_{i=1}^n (\nabla E_{\text{Hi}}[f_i] - \nabla E_{\text{Lo}}[f_i])
\end{aligned} \tag{5.3}$$

Evaluation of these 2-body:Many-body Cartesian gradients is fairly straightforward as long as all gradients are rotated into the same reference frame. The high- and low-level gradients for the fragments $[f_i]$ and pairs $[f_i f_j]$ in Equation 5.3 only contribute to a few components of the composite Cartesian gradient. If atom a is contained in fragment j , then the only non-zero contributions to the component of the Cartesian gradient along the R

coordinate ($R = x, y, z$) of atom a can be obtained with the following expression.

$$\begin{aligned}
 \frac{\partial E^{2\text{bHi:Lo}}}{\partial R_a} &= \frac{\partial E_{\text{Lo}}[f_1 f_2 \dots f_n]}{\partial R_a} \\
 &+ \sum_{i \neq j}^n \left(\frac{\partial E_{\text{Hi}}[f_i f_j]}{\partial R_a} - \frac{\partial E_{\text{Lo}}[f_i f_j]}{\partial R_a} \right) \\
 &+ (n-2) \left(\frac{\partial E_{\text{Lo}}[f_j]}{\partial R_a} - \frac{\partial E_{\text{Hi}}[f_j]}{\partial R_a} \right)
 \end{aligned} \tag{5.4}$$

These analytic gradients were originally implemented in a stand-alone interface to the MPQC *ab initio* software package¹⁰⁵ and applied to the geometry optimization of 15 different hydrogen-bonded clusters of hydrogen fluoride, water and methanol.⁵⁰ In the current implementation, Cartesian gradients are computed with MPQC, rotated into a common reference frame and combined to form a composite 2-body:Many-body gradient that is then passed to the Gaussian03 optimizer via the “external” keyword.

5.4 COMPUTATIONAL METHODS

All water clusters were optimized with the HF and MP2 methods, the MP2:HF QM:QM fragmentation method and the B3LYP density functional. Residual Cartesian gradients of all optimized structures were smaller than $4.5 \times 10^{-4} E_h \text{ bohr}^{-1}$. The 6-31+G($d, 2p$) basis set was used with the B3LYP optimizations because it has been shown that this methodology provides quite accurate structures for $(\text{H}_2\text{O})_6$ isomers.¹³³ All B3LYP computations used a pruned grid, composed of 99 radial shells and 590 angular points per shell. Both HF and MP2 optimizations were performed with a triple- ζ correlation consistent basis set, aug-cc-

pVTZ for O and cc-pVTZ for H (henceforth denoted haTZ). MP2 optimizations were also performed with the 6-31G* basis set, a prescription that has been used to accurately predict the energetics of cluster formation for the same range of water clusters that are the focus of this study.¹⁹⁸ The QM:QM fragmentation optimizations employed MP2/haTZ as the high-level method and HF/haTZ for the low-level calculations.

For all computations, the change in the root mean square (RMS) density between SCF iterations was converged to at least 1×10^{-8} , yielding energies to converged approximately $1 \times 10^{-10} E_h$. The 1s like core orbitals of the oxygen atoms were frozen in all MP2 calculations. All atomic orbital basis sets employed in this work utilized spherical harmonic functions (5*d*, 7*f*) rather than their Cartesian counterparts (6*d*, 10*f*). MP2/haTZ single point energy calculations were performed on all optimized structures to compare the relative energies on the MP2/haTZ potential energy surface (PES). All calculations were performed with Gaussian03,¹⁴⁸ Gaussian09¹⁹⁹ and MPQC¹⁰⁵ software packages.

5.5 RESULTS AND DISCUSSION

Two independent means were used to compare the optimized structures obtained with the various computational methods. The first and more straight forward comparison utilized the minimal RMS deviation of the unweighted Cartesian coordinates optimized with the **superpose** program in TINKER.²⁰⁰ The second metric is based on energetics. MP2/haTZ single point energies were computed for all optimized structures. By definition, the MP2/haTZ optimized structure corresponds to the lowest point associated with a particular minimum

on the MP2/haTZ PES. All other optimized structures lie above that minimum. Optimization procedures that most accurately reproduce the MP2/haTZ optimized structure will lie closest to the bottom of the well and, therefore, also have the smallest deviation from the MP2/haTZ//MP2/haTZ cluster energy.

All trimer, tetramer and pentamer structures are commonly studied low-lying stationary points. Hexamer structures were taken from Reference 179. Heptamer, octamer, nonamer and decamer structures were taken from Reference 198 with a few additional structures taken from Reference 180. The BI2, BI3 and CH3 isomers of $(\text{H}_2\text{O})_7$ along with DP9 of the $(\text{H}_2\text{O})_{10}$ have been omitted because they could not be located on the MP2/haTZ PES. We note, however, that exhaustive searches were not performed because they collapse to other structures on the PES. Because the number of possible configurations grows very quickly with n , only structures within 5 kcal mol⁻¹ of the lowest lying isomer were examined in this study.

Table 5.1 contains the minimal RMS deviations of the unweighted Cartesian coordinates for various optimized structures compared to MP2/haTZ optimized structures. The first column of data shows the deviations associated with the HF/haTZ structures. As expected, HF/haTZ structures have large deviations from the MP2/haTZ structures. The second column of data in Table 5.1 reports the RMS deviations for the MP2/6-31G* optimized structures. Overall, MP2/6-31G* has improved accuracy compared to HF/haTZ methodology. Occasionally however, the MP2/6-31G* RMS values exceed those for the HF/haTZ structures. The values in the last two columns of Table 5.1 are appreciably smaller, indicating

that the B3LYP/6-31+G(*d*, 2*p*) and MP2/haTZ:HF/haTZ optimized structures deviate only slightly from the MP2/haTZ ones. The 2-body:Many-body approach consistently reproduces the MP2/haTZ structures more accurately than any other procedure.

Table 5.1: RMS deviations (in Å) for optimized structures relative to the MP2/haTZ optimized structures.

Method	HF	MP2	B3LYP	MP2:HF
Basis Set	haTZ	6-31G*	6-31+G(<i>d</i> , 2 <i>p</i>)	haTZ
<hr/> (H ₂ O) ₃ <hr/>				
C ₁	0.121	0.127	0.010	0.004
C ₃	0.160	0.186	0.012	0.004
C _{3h}	0.083	0.014	0.009	0.002
<hr/> (H ₂ O) ₄ <hr/>				
S ₄	0.131	0.059	0.017	0.006
C _i	0.146	0.093	0.010	0.007
C ₄	0.204	0.150	0.010	0.007
C _{4h}	0.095	0.019	0.012	0.003
<hr/> (H ₂ O) ₅ <hr/>				
C ₁	0.160	0.126	0.041	0.013
C ₅	0.243	0.136	0.014	0.011
C _{5h}	0.113	0.025	0.018	0.006
<hr/> (H ₂ O) ₆ <hr/>				
prism	0.147	0.081	0.031	0.011
cage	0.180	0.115	0.040	0.008
Book 1	0.203	0.076	0.020	0.008
Book 2	0.262	^a	0.081	0.017
Bag	0.192	^b	0.055	0.010
Boat 1	0.297	0.332	0.078	0.038
Boat 2	0.304	0.280	0.080	0.015
Cyclic	0.154	0.134	0.017	0.012
<hr/> (H ₂ O) ₇ <hr/>				
A	0.173	0.103	0.036	0.007
B	0.196	0.399	0.039	0.009
C	0.274	0.330	0.046	0.012
D	0.328	0.255	0.048	0.024
PR2	0.178	0.131	0.047	0.009
PR3	0.217	0.145	0.036	0.038
CA1	0.242	0.010	0.029	0.010
CA2	0.386	0.174	0.025	0.011
CH1	0.292	0.320	0.101	0.019
BI1	0.465	0.280	0.020	0.015
CH2	0.235	0.294	0.043	0.015
<hr/> (H ₂ O) ₈ <hr/>				
C ₁ a	0.166	0.073	0.029	0.010
C ₁ b	0.167	0.068	0.030	0.009

continued from previous page					
C_1	<i>c</i>	0.173	0.070	0.029	0.007
C_2		0.186	0.021	0.050	0.006
C_i		0.200	0.050	0.031	0.004
C_s		0.179	0.076	0.032	0.007
D_{2d}		0.164	0.068	0.024	0.009
Noncubic 1		0.272	0.308	0.184	0.013
S_4		0.165	0.069	0.026	0.008
$(H_2O)_9$					
$D_{2d}DDh$		0.184	0.091	0.025	0.009
S_4Dah 1		0.185	0.092	0.024	0.008
S_4Dah 2		0.192	0.107	0.025	0.010
S_4DDh 1		0.184	0.103	0.026	0.010
S_4DDh 2		0.187	0.101	0.025	0.008
$D_{2d}Dah$		0.183	0.102	0.022	0.007
S_4Danh 1		0.185	0.125	0.023	0.008
S_4Danh 2		0.187	0.089	0.024	0.010
$(H_2O)_{10}$					
PP1	<i>c</i>		0.073	0.029	0.013
PP2		0.185	0.076	0.026	0.007
PP3		0.186	0.087	0.027	0.021
PP4		0.191	0.084	0.028	0.007
PP5		0.185	0.086	0.026	0.009
OB1		0.200	0.096	0.026	0.011
OB2		0.202	0.094	0.028	0.008
OB3		0.201	0.089	0.028	0.010
DP1	<i>d</i>		0.102	0.028	0.008
OB4		0.202	0.083	0.034	0.015
OB5		0.200	0.203	0.023	0.012
DP2		0.190	0.109	0.033	0.008
OB6		0.204	0.106	0.037	0.017
OB7		0.205	0.070	0.033	0.016
OB8		0.205	0.073	0.032	0.014
DP3		0.204	0.203	0.060	0.010
DP4		0.210	0.081	0.102	0.205
DP5		0.204	0.089	0.050	0.014
DP6		0.198	0.199	0.022	0.009
OB9		0.220	0.221	0.031	0.012
DP7		0.225	0.143	0.032	0.011
DP8		0.208	0.120	0.052	0.013
OB10		0.199	0.318	0.030	0.006
OB11		0.284	0.279	0.043	0.018

continued from previous page				
DP10	0.210	0.103	0.065	0.022
DP11	0.212	0.118	0.026	0.014
C1	0.233	0.100	0.035	0.012
C2	0.199	0.101	0.033	0.024
C3	0.222	0.122	0.039	0.020

^a Collapsed to prism structure.

^b Not located on the MP2/6-31G* PES.

^c Not located on the HF/haTZ PES.

^d Collapsed to DP2 structure.

Table 5.2 summarizes the results of Table 5.1 with the average and maximum RMS deviations associated with each method for each value of n . The second column lists the number of isomers used to compute the average (unless otherwise noted) . For example, the largest RMS deviation between the HF/haTZ and MP2/haTZ structures is 0.465 Å (for isomer BI1 of the water heptamer). In general, the average and maximum RMS deviations of the HF/haTZ and MP2/6-31G* approaches are comparable, with the later exhibiting slightly better performance overall. The average values are roughly 1 order of magnitude smaller for the B3LYP/6-31+G($d, 2p$) optimized structures. The last two columns of Table 5.2 list the average and maximum RMS deviations associated with the 2-body:Many-body fragmentation method employing MP2/haTZ for the high-level calculation and HF/haTZ for the low-level calculations. Regardless of the size of the cluster, this QM:QM fragmentation procedure yields the smallest average errors relative to the MP2/haTZ optimized structures. In fact the RMS deviations never exceed 0.038 Å.

Table 5.2: Average and maximum RMS deviations (in Å) for various optimized structures relative to the MP2/haTZ optimized structures for various $(\text{H}_2\text{O})_n$ clusters with $n=3-10$.

n	#	HF/haTZ		MP2/6-31G*		B3LYP/6-31+G($d, 2p$)		MP2:HF/haTZ	
		Avg	Max	Avg	Max	Avg	Max	Avg	Max
3	3	0.121	0.160	0.109	0.186	0.011	0.012	0.003	0.004
4	4	0.144	0.204	0.080	0.150	0.012	0.017	0.006	0.007
5	3	0.137	0.243	0.096	0.136	0.024	0.041	0.010	0.013
6	8	0.218	0.304	0.170 ^a	0.332 ^a	0.050	0.081	0.015	0.038
7	11	0.272	0.465	0.230	0.399	0.043	0.101	0.015	0.038
8	9	0.186	0.272	0.089	0.308	0.048	0.184	0.008	0.013
9	8	0.186	0.192	0.101	0.125	0.024	0.026	0.009	0.010
10	29	0.207 ^b	0.284	0.125	0.318	0.037	0.102	0.013	0.025

^aExcludes the bag and book 2 isomers.

^bExcludes the PP1 and DP1 isomers.

Table 5.3 is similar to Table 5.1, but it reports energetic, rather than structural, deviations from the MP2/haTZ optimized structures (i.e., from the MP2/haTZ//MP2/haTZ energies). For example, the first column of data reports the MP2/haTZ//HF/haTZ errors associated with the total cluster energy compared to the MP2/haTZ//MP2/haTZ energies. The MP2/haTZ//HF/haTZ errors are always the largest, which is entirely consistent with the RMS deviations. In contrast, the MP2/haTZ//MP2/6-31G* errors are much smaller despite having RMS deviations comparable to the HF/haTZ optimized structures. The errors associated with the B3LYP/6-31+G($d, 2p$) structures are listed in the penultimate column, and they are significantly smaller than the errors associated with the HF/haTZ and MP2/6-31G* optimized structures. The last column of data shows the energetic errors associated with the 2-body:Many-body scheme. Structures optimized with the MP2:HF QM:QM frag-

mentation method and the haTZ basis set are typically one or two hundredths of a kcal mol⁻¹ above the MP2/haTZ optimized structures.

Table 5.3: Errors associated with MP2/haTZ energies (in kcal mol⁻¹) performed on various optimized structures relative to the MP2/haTZ//MP2/haTZ values.

Method	HF	MP2	B3LYP	MP2:HF
Basis Set	haTZ	6-31G*	6-31+G(<i>d</i> , 2 <i>p</i>)	haTZ
<hr/> (H₂O)₃ <hr/>				
C ₁	2.57	0.74	0.10	0.00
C ₃	2.59	0.97	0.11	0.00
C _{3h}	2.27	0.25	0.10	0.00
<hr/> (H₂O)₄ <hr/>				
S ₄	3.82	0.71	0.14	0.01
C _{<i>i</i>}	3.83	0.78	0.15	0.01
C ₄	3.83	1.05	0.16	0.00
C _{4h}	3.25	0.37	0.16	0.00
<hr/> (H₂O)₅ <hr/>				
C ₁	4.81	0.87	0.19	0.01
C ₅	4.69	1.07	0.21	0.01
C _{5h}	4.16	0.51	0.22	0.00
<hr/> (H₂O)₆ <hr/>				
Prism	6.47	0.79	0.25	0.01
Cage	6.50	0.98	0.26	0.01
Book 1	6.18	1.15	0.23	0.01
Book 2	6.25	^a	0.25	0.01
Bag	6.32	^b	0.24	0.01
Boat 1	5.72	1.34	0.25	0.01
Boat 2	5.64	1.22	0.25	0.01
Cyclic	5.67	1.12	0.23	0.01
<hr/> (H₂O)₇ <hr/>				
A	7.77	1.14	0.30	0.01
B	7.64	2.36	0.29	0.01
C	7.21	0.28	0.28	0.01
D	6.89	1.58	0.25	0.01
PR2	7.82	1.02	0.31	0.01
PR3	7.21	1.07	0.30	0.01
CA1	7.07	1.09	0.28	0.01
CA2	6.53	1.26	0.28	0.01
CH1	7.20	1.21	0.28	0.01
BI1	6.82	1.44	0.25	0.01
CH2	7.20	2.20	0.27	0.01
<hr/> (H₂O)₈ <hr/>				
C ₁ a	9.29	1.05	0.34	0.01

continued from previous page				
C ₁ b	9.36	1.03	0.35	0.02
C ₁ c	9.32	1.03	0.35	0.01
C ₂	9.65	0.78	0.39	0.01
C _i	9.73	0.86	0.35	0.01
C _s	9.53	1.01	0.35	0.02
D _{2d}	9.47	1.17	0.35	0.02
Noncubic 1	9.13	1.83	0.41	0.02
S ₄	9.50	1.19	0.36	0.02
<hr/>				
(H ₂ O) ₉				
D _{2d} DDh	10.57	1.31	0.38	0.02
S ₄ Dah 1	10.69	1.30	0.38	0.02
S ₄ Dah 2	10.66	1.35	0.38	0.02
S ₄ DDh 1	10.61	1.35	0.39	0.02
S ₄ DDh 2	10.60	1.36	0.39	0.02
D _{2d} Dah	10.65	1.34	0.37	0.02
S ₄ Danh 1	10.68	1.39	0.38	0.02
S ₄ Danh 2	10.66	1.35	0.38	0.02
<hr/>				
(H ₂ O) ₁₀				
PP1	^c	1.30	0.38	0.00
PP2	11.99	1.35	0.43	0.02
PP3	12.03	1.37	0.41	0.01
PP4	12.20	1.36	0.40	0.02
PP5	12.03	1.41	0.43	0.02
OB1	11.91	1.45	0.43	0.02
OB2	11.94	1.47	0.44	0.01
OB3	11.94	1.48	0.44	0.02
DP1	^d	1.42	0.41	0.02
OB4	11.93	1.42	0.42	0.02
OB5	11.94	1.98	0.43	0.02
DP2	11.81	1.50	0.43	0.02
OB6	11.95	1.51	0.43	0.02
OB7	11.95	1.36	0.43	0.02
OB8	11.94	1.38	0.42	0.02
DP3	11.85	1.48	0.37	0.02
DP4	11.72	1.44	0.29	0.01
DP5	11.69	1.53	0.37	0.02
OB9	12.03	1.90	0.43	0.02
DP6	11.87	1.71	0.41	0.02
DP7	11.33	1.65	0.41	0.02
DP8	11.74	1.82	0.07	0.02
OB10	11.92	2.98	0.42	0.01

continued from previous page				
OB11	11.55	1.92	0.33	0.02
DP10	11.75	1.63	0.41	0.01
DP11	11.83	1.59	0.41	0.03
C1	11.84	1.48	0.40	0.02
C2	11.49	1.55	0.41	0.02
C3	11.63	1.42	0.41	0.02

^a Collapsed to prism structure.

^b Not located on the MP2/6-31G* PES.

^c Not located on the HF/haTZ PES.

^d Collapsed to DP2 structure.

Again, average and maximum errors are tabulated to help summarize all of the data in Table 5.3. For example, the data in Table 5.4 shows that the errors associated with the HF/haTZ structures optimized structures increase with the size of the cluster and grow as large as 12.20 kcal mol⁻¹ (for the PP4 isomer of the water decamer). The energetic errors associated with the MP2/6-31G* optimized structures also tend to increase with the value of n , but do not exceed 2.98 kcal mol⁻¹ (for isomer OB10 of the water decamer). The combination of the B3LYP density functional with the 6-31+G(d , 2 p) basis set appears to be a good way to quickly and reliably identify low-lying structures of (H₂O) _{n} clusters. The largest MP2/haTZ//B3LYP/6-31+G(d , 2 p) error is only 0.44 kcal mol⁻¹ (for both OB2 and OB3 structures of the water decamer). The 2-body:Many-body integrated fragmentation technique for non-covalent clusters provides even more accurate results. The errors associated with the structures optimized with the MP2:HF method and the haTZ basis set never exceed 0.03 kcal mol⁻¹ (for DP11 isomer of the water decamer). The average error for the MP2:HF fragmentation method is 0.01 kcal mol⁻¹ for all of 75 water clusters examined. These 2-body:Many-body results are particularly encouraging for certain pathological cases where

water clusters are virtually isoenergetic and separated by less than $0.10 \text{ kcal mol}^{-1}$. For example, the MP2 complete basis set limit interaction energies for the prism and cage isomers of the water hexamer are separated electronically by only $0.06 \text{ kcal mol}^{-1}$.^{131,179}

Table 5.4: Average and maximum errors for MP2/haTZ energies (in kcal mol⁻¹) performed on various structures relative to MP2/haTZ//MP2/haTZ values for (H₂O)_n clusters with n=3-10.

n	#	HF/haTZ		MP2/6-31G*		B3LYP/6-31+G(d, 2p)		MP2:HF/haTZ	
		Avg	Max	Avg	Max	Avg	Max	Avg	Max
3	3	2.48	2.59	0.65	0.97	0.10	0.11	0.00	0.00
4	4	3.68	3.83	0.73	1.05	0.15	0.16	0.00	0.01
5	3	4.49	4.81	0.81	1.07	0.21	0.21	0.01	0.01
6	8	6.09	6.50	1.10 ^a	1.34 ^a	0.24	0.26	0.01	0.01
7	11	7.23	7.82	1.47	2.36	0.28	0.31	0.01	0.02
8	9	9.44	9.73	1.11	1.83	0.36	0.41	0.01	0.02
9	8	10.64	10.69	1.34	1.39	0.38	0.39	0.02	0.02
10	29	11.84 ^b	12.20	1.58	2.98	0.40	0.44	0.02	0.03

^aExcludes the bag and book 2 isomers.

^bExcludes the PP1 and DP1 isomers.

5.6 CONCLUSIONS

Analytic gradient techniques for the 2-body:Many-body fragmentation method for weakly bound clusters were used to optimize the geometries of more than 70 water clusters ranging in size from (H₂O)₃ to (H₂O)₁₀. In this application, MP2/haTZ was used as the high-level method to compute the 1- and 2- body interactions while HF/haTZ was employed as the low-level method to recover the higher-order (≥ 3 -body) interactions. This procedure proved to be quite efficient because the largest MP2 computations associated with the MP2:HF calculations involve a pair water of molecules (i.e., a dimer), regardless of the size of the cluster. Consequently, the HF/haTZ computation on the entire cluster was always the rate determining step in these 2-body:Many-body fragmentation calculations. Structures optimized with this QM:QM fragmentation procedure were compared to those obtained from conven-

tional MP2/haTZ optimizations using two different metrics, the minimum RMS deviation of unweighted Cartesian coordinates and the MP2/haTZ energy. The 2-body:Many-body optimized structures were virtually identical to those from the MP2/haTZ optimizations. On average, the structures optimized with these two methods were within 0.01 kcal mol⁻¹ of each other on the MP2/haTZ PES, and they never differed by more than 0.03 kcal mol⁻¹. For comparison, HF/haTZ and MP2/6-31G* optimized structures deviated by as much as 12.20 and 2.98 kcal mol⁻¹, respectively, from the MP2/haTZ structures. This work also demonstrated that the B3LYP/6-31+G(*d*, 2*p*) structures did not differ from the MP2/haTZ ones by more than 0.44 kcal mol⁻¹ on the MP2/haTZ PES.

5.7 ACKNOWLEDGEMENTS

We acknowledge the Mississippi Center for Super Computing Research for CPU time and the National Science Foundation for funding (CHE-0957317 and EPS-0903787).

CHAPTER 6

CONCLUSIONS

Detailed conclusions for all of this work are presented in the previous chapters. However, some key results are worth repeating here. The research presented here has:

1. investigated the effect of the inclusion of heteroatoms within parallel-slipped $\pi \cdots \pi$ dimers and showed that the addition of nitrogen atoms to aromatic systems dramatically increases the interaction energy;
2. showed that mixed dimers have appreciably larger interaction energies when compared to their homogenous counterparts;
3. generated CCSD(T) complete basis set limit interaction energies for low-lying structures of the water hexamer and concluded that the inclusion of zero point vibrational energy does not change the relative stabilities of the cage and prism isomers as previously thought;
4. introduced and calibrated a 3-body:Many-body fragmentation approach for determining CCSD(T) quality interaction energies for weakly bound clusters;
5. demonstrated that with proper selections of a high- and low-level method, the 3-body:Many-body fragmentation method can reliably and efficiently reproduce CCSD(T)

- benchmark interaction energies within a few kcal mol⁻¹; and,
6. applied the 2-body:Many-body fragmentation method to computing Cartesian gradients for low-lying structures of water clusters $n=3-10$ which produced structures virtually identical to the target MP2/haTZ structures.

BIBLIOGRAPHY

- [1] J. S. McDougal, M. S. Kennedy, J. M. Sligh, S. P. Cort, A. Mawle, and J. K. A. Nicholson, *Science*, 1986, **231**, 382–385.
- [2] C. Borchers and K. B. Tomer, *Biochemistry*, 2000, **38**, 11734–11740.
- [3] R. Eisenschitz and F. London, *Z. Physik*, 1930, **60**, 491–527.
- [4] F. London, *Z. Phys. Chem. (B)*, 1930, **11**, 222.
- [5] F. London, *Z. Physik*, 1930, **63**, 245–279.
- [6] F. London, *Trans. Faraday Soc.*, 1937, **33**, 8–26.
- [7] G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer, and M. Quack, *J. Chem. Phys.*, 2002, **116**, 690–701.
- [8] J. A. Anderson, K. Cramer, L. Fedoroff, and G. S. Tschumper, *J. Chem. Phys.*, 2004, **121**(22), 11023–11029.
- [9] B. W. Hopkins and G. S. Tschumper, *J. Phys. Chem. A*, 2004, **108**, 2941–2948.
- [10] T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.
- [11] R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

- [12] D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1993, **98**(2), 1358–1371.
- [13] B. Liu and A. McLean, *J. Chem. Phys.*, 1973, **59**, 4557.
- [14] N. R. Kestner, *J. Chem. Phys.*, 1968, **48**(1), 252–257.
- [15] S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
- [16] G. S. Tschumper in *Rev. Comput. Chem.*, ed. K. B. Lipkowitz and D. B. Boyd, Vol. 26; VCH, Inc., New York, 2009; p. invited review.
- [17] D. Feller, *J. Chem. Phys.*, 1992, **96**(8), 6104–6114.
- [18] D. Feller, *J. Chem. Phys.*, 1993, **98**(9), 7059–7071.
- [19] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.*, 1996, **106**(23), 9639–9646.
- [20] E. F. Valeev, *Chem. Phys. Lett.*, 2004, **395**, 190–195.
- [21] W. Klopper, *J. Chem. Phys.*, 1995, **102**, 6168.
- [22] W. Kutzelnigg and W. Klopper, *J. Chem. Phys.*, 1991, **94**(3), 1985–2001.
- [23] W. Klopper, *Chem. Phys. Lett.*, 1991, **186**(6), 583–585.
- [24] W. Klopper, W. Kutzelnigg, H. Müller, J. Noga, and S. Vogtner, *Top. Curr. Chem.*, 1999, **203**, 21–41.
- [25] W. Klopper and C. C. M. Samson, *J. Chem. Phys.*, 2002, **116**(15), 6397–6410.

- [26] E. F. Valeev and C. L. Janssen, *J. Chem. Phys.*, 2004, **121**(3), 1214–1227.
- [27] W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev, *Int. Rev. Phys. Chem.*, 2006, **25**(3), 427–468.
- [28] E. A. Hylleraas, *Z. Phys.*, 1929, **54**, 347.
- [29] K. Kitaura, T. Sawai, T. Asada, T. Nakano, and M. Uebayasi, *Chem. Phys. Lett.*, 1999, **312**(2–4), 319–324.
- [30] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, *Chemical Physics Letters*, 1999, **313**(3–4), 701 – 706.
- [31] D. G. Fedorov and K. Kitaura, *J. Phys. Chem. A*, 2007, **111**(30), 6904–6914.
- [32] M. Kamiya, S. Hirata, and M. Valiev, *J. Chem. Phys.*, 2008, **128**(7), 074103.
- [33] W. Yang, *Phys. Rev. Letts*, 1991, **66**(11), 1438–1441.
- [34] T. -S. Lee, J. P. Lewis, and W. Yang, *Comp. Mat. Sci*, 1998, **12**(3), 259–277.
- [35] D. Min and W. Yang, *J. Chem. Phys.*, 2008, **128**(9), 094106.
- [36] S. Hirata, M. Valiev, M. Dupuis, S. S. Xantheas, S. Sugiki, and H. Sekino, *Mol. Phys.*, 2005, **103**(15–16), 2255–2265.
- [37] S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095.
- [38] J. G. Hill and J. A. Platts, *J. Chem. Theory Comput.*, 2007, **3**, 80–85.

- [39] Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, *J. Chem. Phys.*, 2004, **121**(20), 9793–9802.
- [40] R. A. Christie and K. D. Jordan in *Intermolecular Forces and Clusters II*, ed. D. J. Wales, Vol. 116 of *Structure and Bonding*; Springer, Germany, 2005; pp. 27–41.
- [41] E. Dahlke and D. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**(1), 46–53.
- [42] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, and D. Cohen, *J. Chem. Phys.*, 1996, **105**(5), 1968–1986.
- [43] P. N. Day, R. Pachter, M. S. Gordon, and G. N. Merrill, *J. Chem. Phys.*, 2000, **112**(5), 2063.
- [44] M. S. Gordon, H. L. Lyudmilla Slipchenko and, and J. H. Jensen in *Annual Reports in Computational Chemistry*, ed. D. C. Spellmeyer and R. Wheeler, Vol. 3; Elsevier, Amsterdam, 2007; pp. 177–193.
- [45] I. Adamovic and M. S. Gordon, *J. Phys. Chem. A*, 2006, **110**(34), 10267–10273.
- [46] M. S. Gordon, J. M. Mullin, S. R. Pruitt, L. B. Roskop, L. V. Slipchenko, and J. A. Boatz, *J. Phys. Chem. B*, 2009, **113**(29), 9646–9663.
- [47] S. Li, J. Ma, and Y. Jiang, *J. Comp. Chem.*, 2002, **23**, 237–244.
- [48] W. Li, P. Piecuch, J. R. Gour, and S. Li, *The Journal of Chemical Physics*, 2009, **131**(11), 114109.

- [49] G. S. Tschumper, *Chem. Phys. Lett.*, 2006, **427**(1–3), 185–191.
- [50] A. M. ElSohly, C. L. Shaw, M. E. Guice, B. D. Smith, and G. S. Tschumper, *Mol. Phys.*, 2007, **105**(19–22), 2777–2782.
- [51] G. J. O. Beran, *J. Chem. Phys.*, 2009, **130**, 164115–164124.
- [52] A. Sebetci and G. J. O. Beran, *J. Chem. Theory Comput.*, 2010, **6**, 155–167.
- [53] J. Šponer, P. Jurečka, I. Marchan, F. J. Luque, M. Orozco, and P. Hobza, *Chem. Eur. J.*, 2006, **12**, 2854–2865.
- [54] J. Rejnek and P. Hobza, *J. Phys. Chem. B*, 2007, **111**, 641–645.
- [55] R. R. Toczyłowski and S. M. Cybulski, *J. Phys. Chem. A*, 2003, **107**, 418–426.
- [56] P. Hobza and J. Šponer, *J. Am. Chem. Soc.*, 2002, **124**, 11802–11808.
- [57] G. B. McGaughey, M. Gagne, and A. K. Rappe, *J. Bio. Chem.*, 1998, **273**, 15458–15463.
- [58] A. C. Cheng and A. D. Frankel, *J. Am. Chem. Soc.*, 2004, **126**, 434–435.
- [59] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**(17), 1985–1993.
- [60] Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 6624–6627.
- [61] G. Alagona, C. Ghio, and S. Monti, *Int. J. Quant. Chem.*, 1999, **73**, 175–186.

- [62] L. Mao, Y. Wang, Y. Liu, and X. Hu, *J. Mol. Bio.*, 2004, **336**, 787–807.
- [63] P. Chene, *Nat. Rev. Drug Discovery*, 2002, **1**, 665–673.
- [64] M. Saeki, H. Akagi, and M. Fujii, *J. Chem. Theory Comput.*, 2006, **2**, 1176–1183.
- [65] D. M. Rogers, J. D. Hirst, E. P. F. Lee, and T. G. Wright, *Chem. Phys. Lett.*, 2006, **427**, 410–413.
- [66] Y. Wang and X. Hu, *J. Chem. Phys.*, 2002, **117**, 1–4.
- [67] S. Gundersun, S. Samdal, T. G. Strand, and H. V. Volden, *J. Mol. Struct.*, 2007, **832**, 164–171.
- [68] W. Scherzer, O. Kratzschmar, H. L. Selzle, and E. W. Schlag, *Z. Naturforsch., A: Phys. Sci.*, 1992, **47**(12), 1248–1252.
- [69] E. Arunan and H. S. Gutowsky, *J. Chem. Phys.*, 1993, **98**(5), 4294–4296.
- [70] P. Hobza, H. L. Selzle, and E. W. Schlag, *J. Am. Chem. Soc.*, 1994, **116**(8), 3500–3506.
- [71] H. J. Neusser and H. Krause, *Chem. Rev.*, 1994, **94**(7), 1829–1843.
- [72] P. Hobza, H. Selzle, and E. Schlag, *J. Phys. Chem.*, 1996, **100**(48), 18790–18794.
- [73] P. Hobza, V. Spirko, H. L. Selzle, and E. W. Schlag, *J. Phys. Chem. A*, 1998, **102**(15), 2501–2504.
- [74] C. A. Hunter, *Angew. Chem. Int. Ed.*, 1993, **32**(11), 1584–1586.

- [75] R. L. Jaffe and G. D. Smith, *J. Chem. Phys.*, 1996, **105**(7), 2780–2788.
- [76] M. Sinnokrot, E. Valeev, and C. Sherrill, *J. Am. Chem. Soc.*, 2002, **124**(36), 10887–10893.
- [77] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, and K. Tanabe, *J. Am. Chem. Soc.*, 2002, **124**(1), 104–112.
- [78] M. Mons, I. Dimicoli, and F. Piuzzi, *Int. Rev. Phys. Chem.*, 2002, **21**(1), 101–135.
- [79] M. Sinnokrot and C. Sherrill, *J. Phys. Chem. A*, 2004, **108**(46), 10200–10207.
- [80] O. A. Zhikol, O. V. Shishkin, K. A. Lyssenko, and J. Leszczynski, *J. Chem. Phys.*, 2005, **122**(14), 144104.
- [81] Y. C. Park and J. S. Lee, *J. Phys. Chem. A*, 2006, **110**(15), 5091–5095.
- [82] J. G. Hill, J. A. Platts, and H. J. Werner, *Phys. Chem. Chem. Phys.*, 2006, **8**(35), 4072–4078.
- [83] M. Sinnokrot and C. Sherrill, *J. Phys. Chem. A*, 2006, **110**(37), 10656–10668.
- [84] B. W. Hopkins, A. M. ElSohly, and G. S. Tschumper, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1550–1558.
- [85] M. Sinnokrot and C. Sherrill, *J. Phys. Chem. A*, 2003, **107**(41), 8377–8379.
- [86] M. O. Sinnokrot and C. D. Sherrill, *J. Am. Chem. Soc.*, 2004, **126**, 7690–7697.

- [87] A. L. Ringer, M. O. Sinnokrot, R. P. Lively, and C. D. Sherrill, *Chem. Eur. J.*, 2006, **12**, 3821–3828.
- [88] E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, and K. S. Kim, *J. Phys. Chem. A*, 2007, **111**, 3446–3457.
- [89] P. Hobza and J. Šponer, *Chem. Rev.*, 1999, **99**(11), 3247–3276.
- [90] S. Tsuzuki, K. Honda, and R. Azumi, *J. Am. Chem. Soc.*, 2002, **124**, 12200–12209.
- [91] F. Rodriguez-Ropero, J. Casanovas, and C. Aleman, *J. Comput. Chem.*, 2007, **29**, 69–78.
- [92] F. Ugozzoli and C. Massera, *Cryst. Eng. Comm.*, 2005, **7**, 121–128.
- [93] P. Hobza, *Annu. Rep. Prog. Chem. Sect. C*, 1996, **93**, 257.
- [94] J. Šponer and P. Hobza, *Chem. Phys. Lett.*, 1997, **267**, 263.
- [95] M. Zaccheddu, C. Filippi, and F. Buda, *J. Phys. Chem. A*, 2008, **112**, 1627–1632.
- [96] S. Huzinaga, *J. Chem. Phys.*, 1965, **42**(4), 1293–1302.
- [97] T. H. Dunning, *J. Chem. Phys.*, 1970, **53**, 2823.
- [98] T. J. Lee and H. F. Schaefer, *J. Chem. Phys.*, 1985, **83**(4), 1784–1794.
- [99] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A.

- Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision C.02, 2003.
- [100] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, M. L. Leininger, R. A. King, S. T. Brown, C. L. Janssen, E. T. Seidl, Y. Yamaguchi, W. D. Allen, Y. Xie, G. Vacek, T. P. Hamilton, C. B. Kellogg, R. B. Remington, and H. F. Schaefer III, Psi 3.0 PSITECH, Inc. , Watkinsville, GA 30677, U.S.A., 1999.
- [101] C. L. Janssen, E. T. Seidl, and M. E. Colvin, in *Parallel Computing in Computational Chemistry*, ed. T. G. Mattson, American Chemical Society, Washington, DC, 1995; p. 47; ACS Symposium Series 592.
- [102] I. M. B. Nielsen and E. T. Seidl, *J. Comput. Chem.*, 1995, **16**, 1301.
- [103] I. M. B. Nielsen, *Chem. Phys. Lett.*, 1996, **255**, 210.

- [104] I. M. B. Nielsen and C. L. Janssen, *Comput. Phys. Comm.*, 2000, **128**, 238.
- [105] C. L. Janssen, I. B. Nielsen, M. L. Leininger, E. F. Valeev, and E. T. Seidl, The massively parallel quantum chemistry program (mpqc) version 2.3.1 Sandia National Laboratories, Livermore, CA, USA, <http://www.mpqc.org>, 2004.
- [106] B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.*, 1994, **94**(7), 1887–1930.
- [107] R. Bukowski, W. Cencek, P. Jankowski, B. Jeziorski, M. Jeziorska, S. A. Kucharski, V. F. Lotrich, A. J. Misquitta, R. Moszynski, K. Patkowski, S. Rybak, K. Szalewicz, H. L. Williams, R. J. Wheatley, P. E. S. Wormer, and P. S. Zuchowski, Sapt2006: an ab initio program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies. <http://www.physics.udel.edu/~szalewic/SAPT>, 2006.
- [108] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639.
- [109] S. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
- [110] H. B. Jansen and P. Ros, *Chem. Phys. Lett.*, 1969, **3**(3), 140–143.
- [111] S. Tsuzuki and T. Uchimaru, *Curr. Org. Chem.*, 2006, **10**(7), 745–762.
- [112] P. Hobza, R. Zahradnik, and K. Müller-Dethlefs, *Collect. Czech. Chem. Commun.*, 2006, **71**(4), 443–531.
- [113] H. L. Williams, K. Szalewicz, B. Jeziorski, R. Moszynski, and S. Rybak, *J. Chem. Phys.*, 1993, **98**, 1279–1292.

- [114] J. Antony and S. Grimme, *J. Phys. Chem. A.*, 2007, **111**, 4862–4868.
- [115] B. J. Mhin, J. S. Kim, and S. Lee, *J. Chem. Phys.*, 1994, **100**(6), 4484–4486.
- [116] K. Kim, K. D. Jordan, and T. S. Zwier, *J. Am. Chem. Soc.*, 1994, **116**(25), 11568–11569.
- [117] K. Liu, M. G. Brown, and C. Carter, *Nature*, 1996, **381**(6582), 501–503.
- [118] D. A. Estrin, L. Paglieri, G. Corogiu, and E. Clementi, *J. Phys. Chem.*, 1996, **100**, 8701.
- [119] H. M. Lee, S. B. Suh, J. Y. Lee, P. Tarakeshwar, and K. S. Kim, *The Journal of Chemical Physics*, 2000, **112**(22), 9759–9772.
- [120] J. Kim and K. S. Kim, *J. Chem. Phys.*, 1998, **109**(14), 5886–5895.
- [121] K. Nauta and R. E. Miller, *Science*, 2000, **287**(5451), 293–295.
- [122] R. Ludwig, *Angew. Chem. Int. Ed. Engl.*, 2001, **40**(10), 1808–1827.
- [123] M. D. Tissandier, S. J. Singer, and J. V. Coe, *J. Phys. Chem. A*, 2000, **104**(4), 752–757.
- [124] K. Lui, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory, and D. C. Clary, *Nature*, 1996, **381**, 501–503.
- [125] J. B. Paul, C. P. Collier, R. J. Saykally, J. J. Scherer, and A. O’Keefe, *J. Phys. Chem. A*, 1997, **101**(29), 5211–5214.
- [126] K. Liu, M. G. Brown, and R. J. Saykally, *J. Phys. Chem. A*, 1997, **101**(48), 8995–9010.

- [127] M. E. Fajardo and S. Tam, *J. Chem. Phys.*, 2001, **115**(15), 6807–6810.
- [128] C. Steinbach, P. Andersson, M. Melzer, J. K. Kazimirski, U. Buck, and V. Buch, *Phys. Chem. Chem. Phys.*, 2004, **6**, 3320–3324.
- [129] S. Hirabayashi and K. M. T. Yamada, *J. Mol. Struct. (Theochem)*, 2006, **795**, 78–83.
- [130] S. Hirabayashi and K. M. T. Yamada, *Chem. Phys. Lett.*, 2006, **435**, 74–78.
- [131] S. S. Xantheas, C. J. Burnham, and R. J. Harrison, *J. Chem. Phys.*, 2002, **116**(4), 1493–1499.
- [132] R. M. Olson, J. L. Bentz, R. A. Kendall, M. W. Schmidt, and M. S. Gordon, *J. Chem. Theory Comput.*, 2007, **3**(4), 1312–1328.
- [133] E. E. Dahlke, R. M. Olson, H. R. Leverentz, and D. G. Truhlar, *J. Phys. Chem. A*, 2008, **112**(17), 3976–3984.
- [134] C. Kozmutza, E. Kryachko, and E. Tfirst, *J. Mol. Struct. (Theochem)*, 2000, **501-502**, 435–444.
- [135] M. Losada and S. Leutwyler, *J. Chem. Phys.*, 2002, **117**(5), 2003–2016.
- [136] K. Diri, E. M. Myshakin, and K. D. Jordan, *J. Phys. Chem. A*, 2005, **109**, 4005–4009.
- [137] M. E. Dunn, E. K. Pokon, and G. C. Shields, *J. Am. Chem. Soc.*, 2004, **126**, 2647–2653.
- [138] W. Klopper, *J. Chem. Phys.*, 1995, **102**(15), 6168–6179.
- [139] E. F. Valeev, *Chem. Phys. Lett.*, 2004, **395**(4–6), 190–195.

- [140] W. Klopper, H. P. Luthi, T. Brupbacher, and A. Bauder, *J. Chem. Phys.*, 1994, **101**(11), 9747–9754.
- [141] W. Klopper and M. Schuetz, *Ber. Bunsenges. Phys. Chem.*, 1995, **99**, 469–473.
- [142] W. Klopper, M. Quack, and M. A. Suhm, *Mol. Phys.*, 1998, **94**(1), 105–116.
- [143] Y. Park and J. Lee, *J. Phys. Chem. A*, 2006, **110**(15), 5091–5095.
- [144] S. Tsuzuki, K. Honda, T. Uchimaru, and M. Mikami, *J. Chem. Phys.*, 2006, **124**(11), 114304.
- [145] J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2007, **9**(39), 5281–5388.
- [146] R. A. DiStasio, G. von Helden, R. P. Steele, and M. Head-Gordon, *Chem. Phys. Lett.*, 2007, **437**(4-6), 277–283.
- [147] C. D. Sherrill in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and T. R. Cundari, Vol. 26; Wiley-VCH, Inc., Hoboken, NJ, 2009; pp. 1–38.
- [148] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. R. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y.

- Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision E.01, 2003.
- [149] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, and T. Thorsteinsson, Molpro, version 2006.1, a package of ab initio programs, 2006.
- [150] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, *J. Comput. Chem*, 2007, **28**, 1610–1616.
- [151] H. B. Jansen and P. Ros, *Chem. Phys. Lett.*, 1969, **3**, 140–143.
- [152] B. Liu and A. D. McLean, *J. Chem. Phys.*, 1973, **59**, 4557–4558.
- [153] D. M. Bates, J. A. Anderson, P. Oloyede, and G. S. Tschumper, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2775–2779.

- [154] R. S. Grev, C. L. Janssen, and H. F. Schaefer, *J. Chem. Phys.*, 1991, **95**(7), 5128–5132.
- [155] A. P. Scott and L. Radom, *J. Phys. Chem.*, 1996, **100**(41), 16502–16513.
- [156] N. Goldman and R. J. Saykally, *J. Chem. Phys.*, 2004, **120**(10), 4777–4789.
- [157] R. Hoffmann, P. von Ragué Schleyer, and H. F. Schaefer, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**(38), 7164–7167.
- [158] B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, *J. Chem. Phys.*, 2008, **129**(19), 194111.
- [159] J. Rezac, P. Hobza, and S. A. Harris, *Biophys. J.*, 2010, **98**, 101.
- [160] T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi, and K. Kitaura, *Chemical Physics Letters*, 2000, **318**(6), 614 – 618.
- [161] D. G. Fedorov, K. Ishimura, T. Ishida, K. Kitaura, P. Pulay, and S. Nagase, *J. Comp. Chem.*, 2007, **28**(9), 1476–1484.
- [162] S. Sakai and S. Morita, *J. Phys. Chem. A*, 2005, **109**(37), 8424–8429.
- [163] E. Dahlke and D. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**(4), 1342–1348.
- [164] E. Dahlke, H. Leverentz, and D. Truhlar, *J. Chem. Theory Comput.*, 2008, **4**(1), 33–41.
- [165] A. Imamura, Y. Aoki, and K. Maekawa, *The Journal of Chemical Physics*, 1991, **95**(7), 5419–5431.
- [166] V. Deev and M. A. Collins, *The Journal of Chemical Physics*, 2005, **122**(15), 154102.

- [167] D. W. Zhang and J. Z. H. Zhang, *The Journal of Chemical Physics*, 2003, **119**(7), 3599–3605.
- [168] N. Jiang, J. Ma, and Y. Jiang, *J. Chem. Phys.*, 2006, **124**(11), 114112.
- [169] S. R. Gadre, R. N. Shirsat, and A. C. Limaye, *J. Chem. Phys.*, 1994, **98**, 9103–9169.
- [170] S. Sb and P. Pulay, *Chemical Physics Letters*, 1985, **113**(1), 13 – 18.
- [171] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, *The Journal of Chemical Physics*, 1996, **105**(5), 1968–1986.
- [172] G. Rauhut, P. Pulay, and H. J. Werner, *J. Comp. Chem.*, 1998, **19**(11), 1241–1254.
- [173] B. W. Hopkins and G. S. Tschumper, *Chem. Phys. Lett.*, 2005, **407**(4–6), 362–367.
- [174] F. Maseras and K. Morokuma, *J. Comput. Chem.*, 1995, **16**, 1170–1179.
- [175] E. Apra, A. P. Rendell, R. J. Harrison, V. Tipparaju, W. A. deJong, and S. S. Xanthreas, In proceedings of the conference on high performance computing, networking, storage and analysis, 2009.
- [176] T. P. Tauer and C. D. Sherrill, *J. Phys. Chem. A*, 2005, **109**(46), 10475–10478.
- [177] B. W. Hopkins and G. S. Tschumper, *J. Comput. Chem.*, 2003, **24**(13), 1563–1568.
- [178] B. W. Hopkins and G. S. Tschumper, *Mol. Phys.*, 2005, **103**(2–3), 309–315.
- [179] D. M. Bates and G. S. Tschumper, *J. Phys. Chem. A*, 2009, **113**(15), 3427–3708.

- [180] P. Qian, W. Song, L. Lu, and Z. Yang, *Int. J. Quantum Chem.*, 2009, **110**, 1923–1937.
- [181] D. M. Bates, J. R. Smith, P. Pulay, and G. S. Tschumper, Supporting information, 2010.
- [182] A. M. ElSohly and G. S. Tschumper, *Int. J. Quantum Chem.*, 2009, **109**(1), 91–96.
- [183] G. S. Tschumper in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and T. R. Cundari, Vol. 26; Wiley-VCH, Inc., Hoboken, NJ, 2009; pp. 39–90.
- [184] G. S. Tschumper and K. Morokuma, *J. Mol. Struct. (Theochem)*, 2002, **592**(1–3), 137–147.
- [185] B. W. Hopkins and G. S. Tschumper, *Int. J. Quantum Chem.*, 2004, **96**(4), 294–302.
- [186] S. S. Xantheas, *J. Chem. Phys.*, 2000, **258**, 225–231.
- [187] A. Castleman and K. Bowen, *J. Phys. Chem.*, 1996, **100**(31), 12911–12944.
- [188] Z. Bačić and R. E. Miller, *J. Phys. Chem.*, 1996, **100**(31), 12945–12959.
- [189] G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, England, 1997.
- [190] P. Schuster and P. Wolschann, *Monat. Chem.*, 1999, **130**(8), 947–960.
- [191] T. Steiner, *Angew. Chem. Int. Ed. Engl.*, 2002, **41**(1), 48–76.
- [192] P. Hobza, *Annu. Rep. Prog. Chem., Sect. C*, 2004, **100**, 3–27.

- [193] Y. Zhao and D. G. Truhlar, *J. Chem. Theo. Comput.*, 2005, **1**, 415–432.
- [194] A. Boese, J. Martin, and W. Klopper, *J. Phys. Chem. A*, 2007, **111**(43), 11122–11133.
- [195] K. L. Copeland, J. A. Anderson, A. R. Farley, J. R. Cox, and G. S. Tschumper, *J. Phys. Chem. B*, 2008, **112**(45), 14291–14295.
- [196] A. M. ElSohly, B. W. Hopkins, K. L. Copeland, and G. S. Tschumper, *Mol. Phys.*, 2009, p. accepted.
- [197] D. M. Bates, T. Janowski, J. R. Smith, and G. S. Tschumper, *J. Chem. Phys.*, 2011, p. submitted .
- [198] R. M. Shields, B. Temelso, K. A. Archer, T. E. Morrell, and G. C. Shields, *J. Phys. Chem. A*, 2010, **114**(43), 11725–11737.
- [199] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G.

Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels,
O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09
Revision A.2.

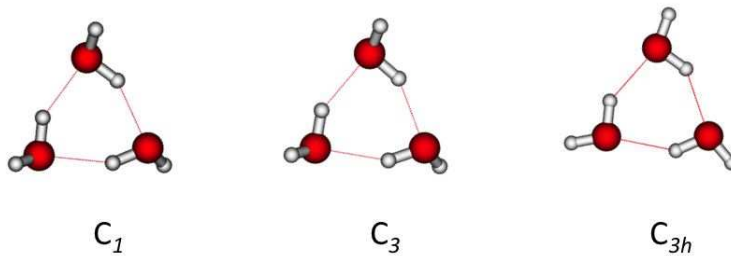
[200] P. JW, Tinker - software tools for molecular design, version 5.1.09.

CHAPTER 7

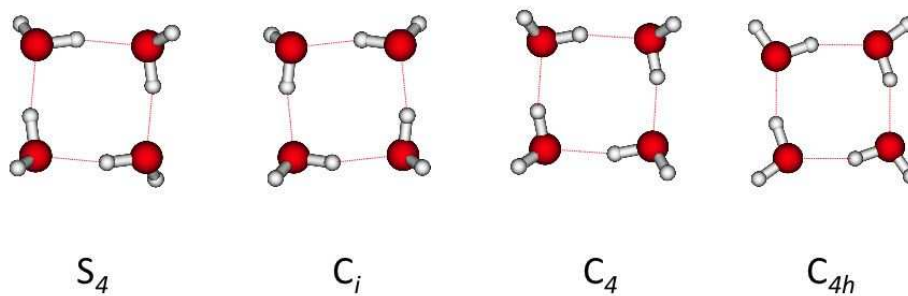
APPENDIX

7.1 STRUCTURE OF VARIOUS WATER CLUSTERS $(\text{H}_2\text{O})_n$ $n=3-10$

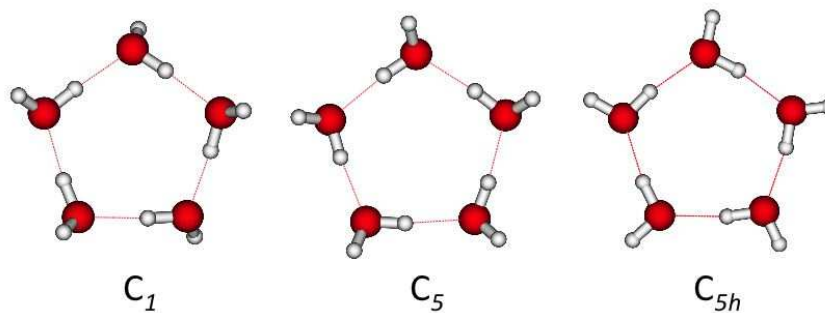
$(\text{H}_2\text{O})_n$ Clusters, $n=3-10$



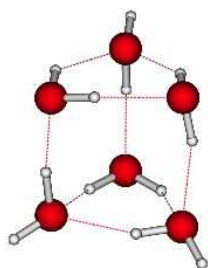
(a) $(\text{H}_2\text{O})_3$ isomers



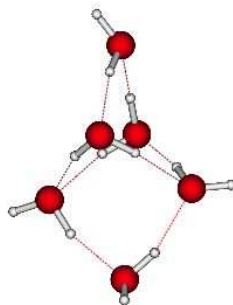
(b) $(\text{H}_2\text{O})_4$ isomers



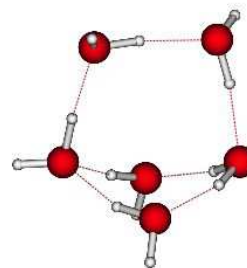
(c) $(\text{H}_2\text{O})_5$ isomers



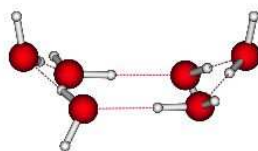
Prism



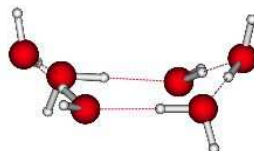
Cage



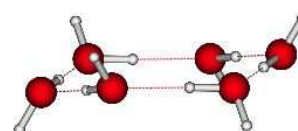
Bag



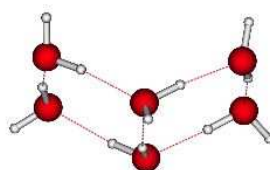
Boat 1



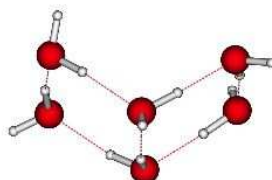
Boat 2



Cyclic

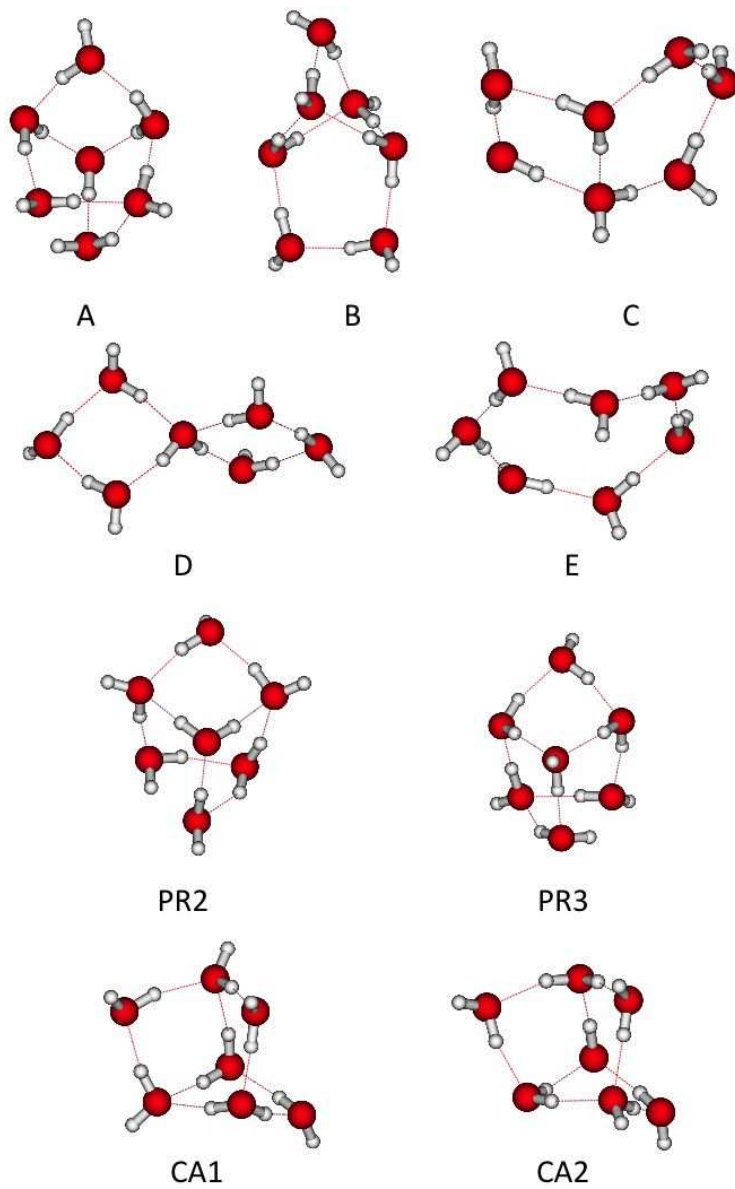


Book 1

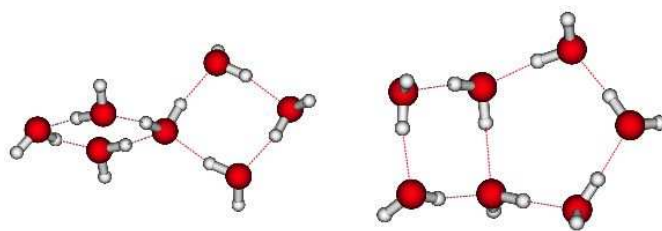


Book 2

(d) $(\text{H}_2\text{O})_6$ isomers

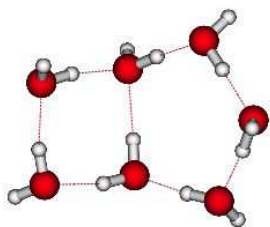


(e) $(\text{H}_2\text{O})_7$ isomers

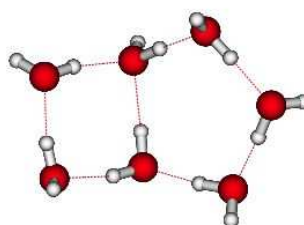


BI1

CH1

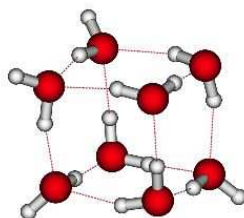


CH2

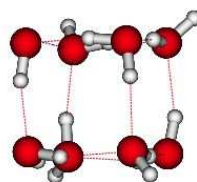


CH3

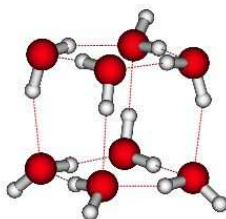
(f) (H₂O)₇ isomers continued



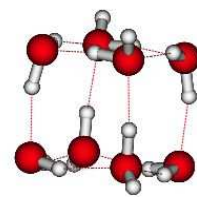
A



B

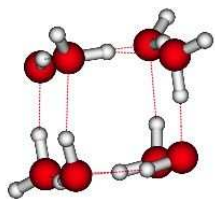


C

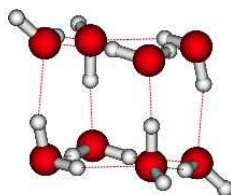


C₂

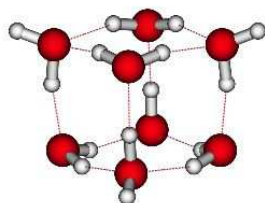
(g) (H₂O)₈ isomers



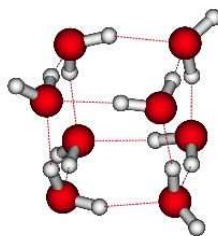
C_s



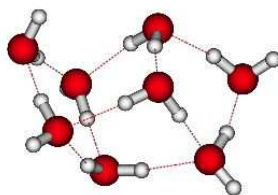
C_i



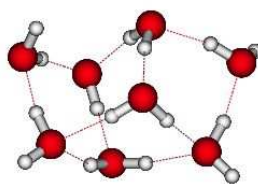
D_{2d}



S_4

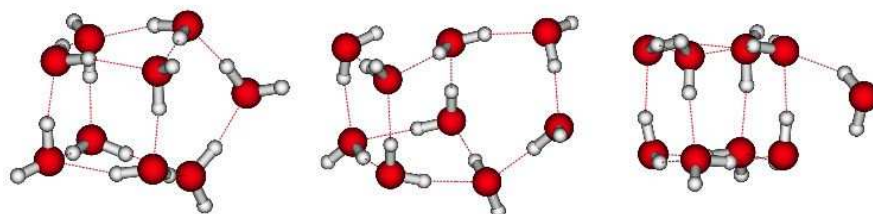


Noncubic 1



Noncubic 2

(h) $(H_2O)_8$ isomers continued

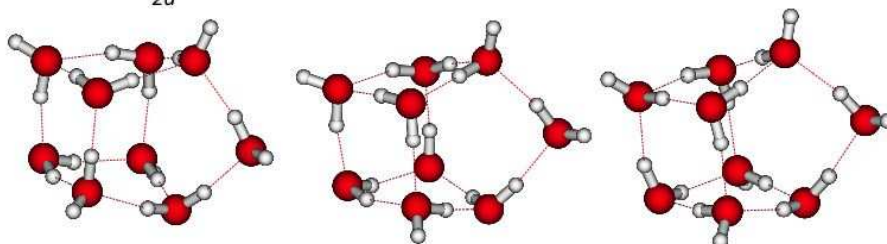


A

C

D

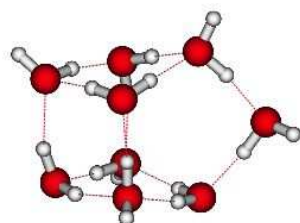
Also referred
to as $D_{2d}DDh$



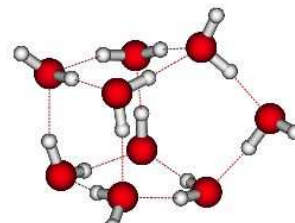
S_4Dah 1

S_4Dah 2

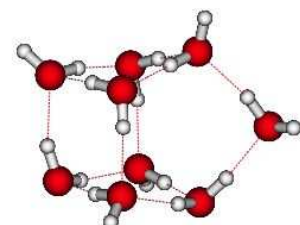
$D_{2d}Dah$



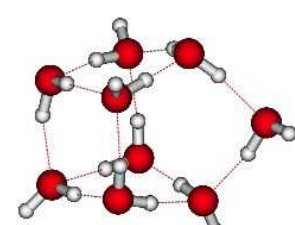
S_4DDh 1



S_4DDh 2

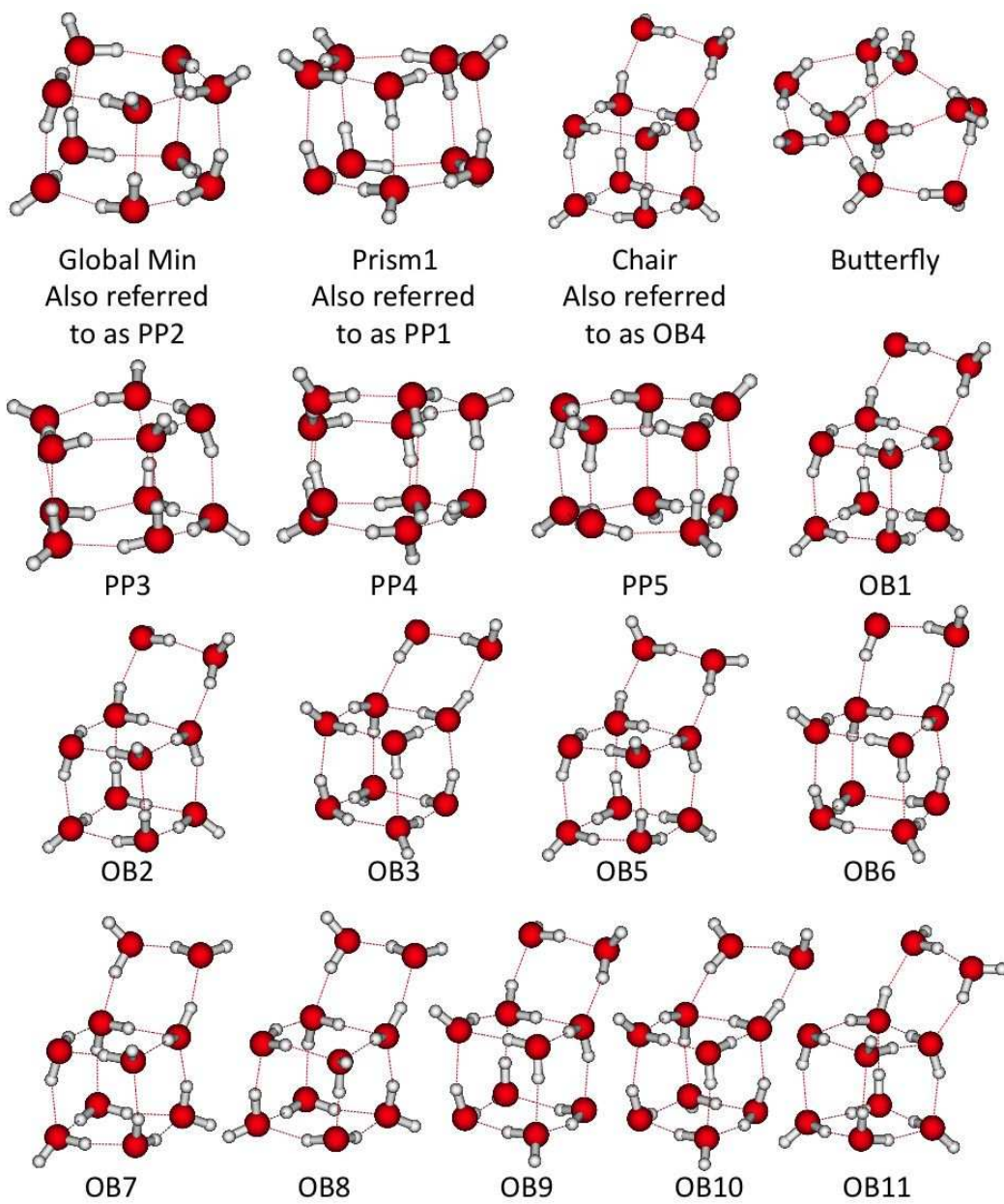


S_4Danh 1

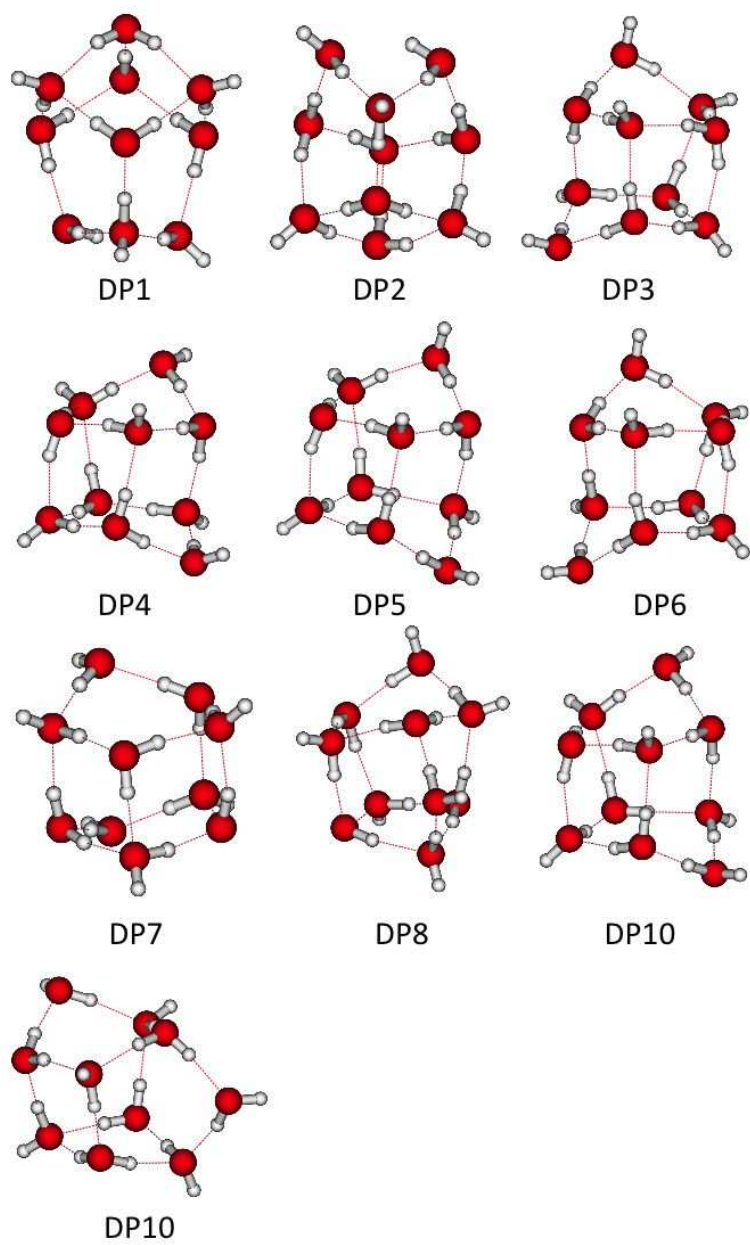


S_4Danh 2

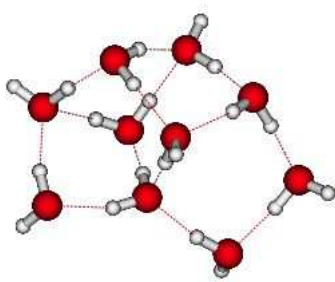
(i) $(H_2O)_9$ isomers



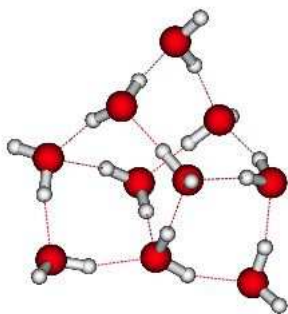
(j) (H₂O)₁₀ isomers



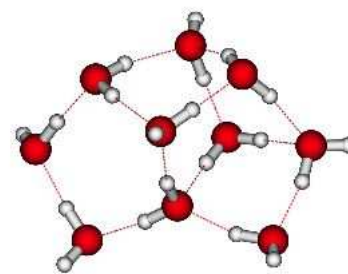
(k) (H₂O)₁₀ isomers continued



C1



C2



C3

(l) (H₂O)₁₀ isomers continued

7.2 SUPPORTING INFORMATION FOR CCSD(T) COMPLETE BASIS SET LIMIT RELATIVE ENERGIES FOR LOW-LYING WATER HEXAMER STRUCTURES

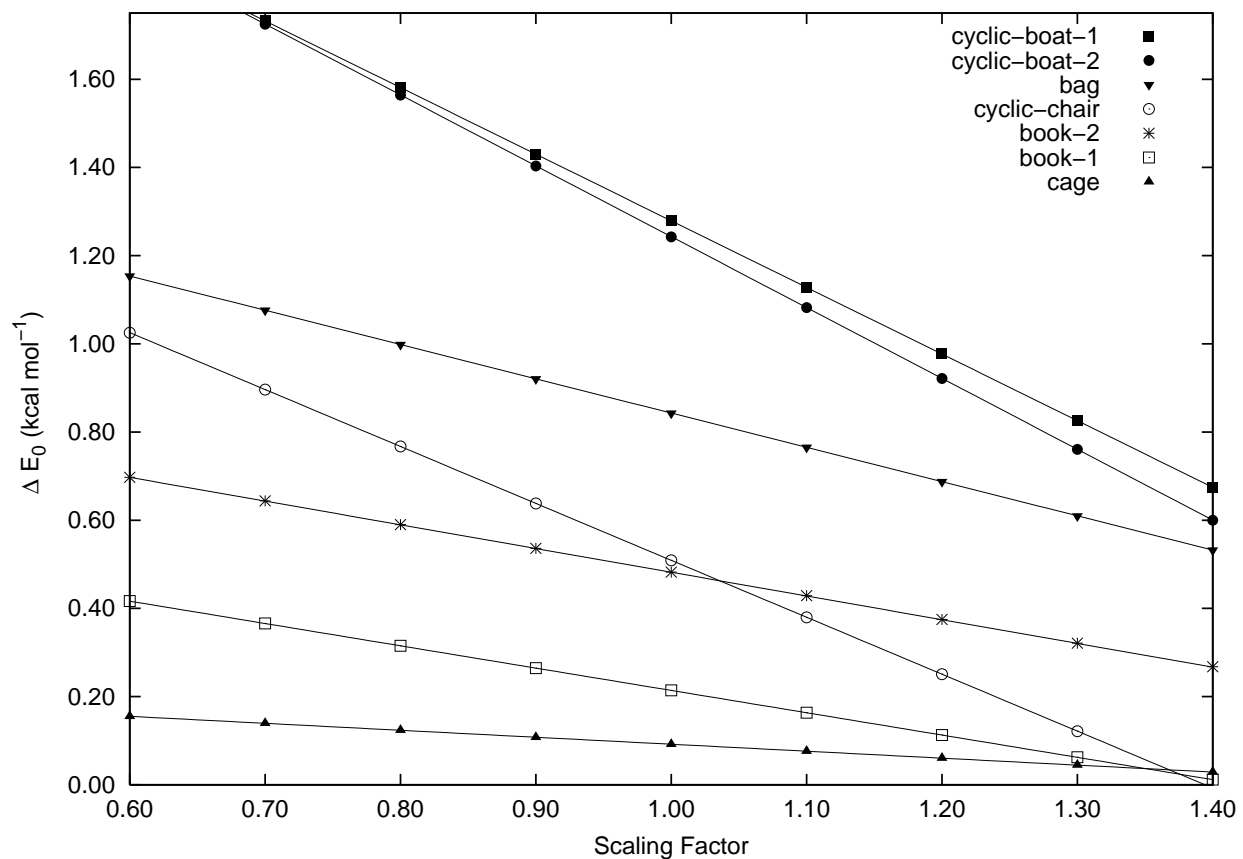


Figure S0: Effect of scaling factors for MP2/haTZ harmonic vibrational frequencies on the relative energies of the water hexamer isomers. All relative energies increase with respect to the prism for scaling factors less than 1. Only with a scaling factor larger than 1.39 does another structure (cyclic-chair) become more stable than the prism. The relative stability cyclic-chair and book-2 structures inverts at a value of 1.04.

Table S1: Harmonic ZPVE corrections (δ_{ZPVE} in kcal mol⁻¹) computed with various methods and basis sets.

Method	MP2	BLYP	B3LYP	BLYP	B3LYP	PBE
Basis	haTZ	6-31+G(<i>d, p</i>)	6-31+G(<i>d, 2p</i>)	MG3S	MG3S	MG3S
prism	+0.00	+0.00	+0.00	+0.00	+0.00	+0.00
cage	-0.16	-0.13	-0.14	-0.02	-0.16	-0.08
bag	-0.78	-0.67	-0.62	-0.60	-0.69	-0.60
cyclic-ring	-1.29	-1.11	-1.04	-0.92	-1.14	-0.93
book-1	-0.51	-0.40	-0.38	-0.32	-0.45	-0.34
book-2	-0.54	-0.37	-0.30	-0.33	-0.47	-0.32
cyclic-boat-1	-1.51	-1.27	-1.29	-1.10	-1.28	-1.10
cyclic-boat-2	-1.61	-1.36	-1.41	-1.18	-1.46	-1.11
6 monomers	-13.71	-14.21	-14.07	-13.35	-13.82	-13.42

Table S2: Total electronic energies and sum of electronic and zero-point vibrational energies (E_e and E_{+ZPVE} , respectively, in E_h) from the MP2/haTZ computations as well as the corresponding relative energies (ΔE_e and ΔE_{+ZPVE} in kcal mol⁻¹).

Structure	$E_e^{\text{MP2/haTZ}}$	$\Delta E_e^{\text{MP2/haTZ}}$	$E_{+ZPVE}^{\text{MP2/haTZ}}$	$\Delta E_{+ZPVE}^{\text{MP2/haTZ}}$	δ_{ZPVE}
prism	-458.042669	+0.00	-457.892242	+0.00	+0.00
cage	-458.042650	-0.15	-457.892475	+0.01	-0.16
bag	-458.040790	+0.40	-457.891601	+1.18	-0.78
cyclic-ring	-458.040683	-0.05	-457.892314	+1.25	-1.29
book-1	-458.042171	-0.19	-457.892550	+0.31	-0.51
book-2	-458.041693	+0.07	-457.892123	+0.61	-0.54
cyclic-boat-1	-458.039088	+0.74	-457.891068	+2.25	-1.51
6yclic-boat-2	-458.038958	+0.72	-457.891092	+2.33	-1.61
6 monomers	-457.968327	+32.94	-457.839750	+46.65	-13.71

Curriculum Vitae

Desiree M. Bates

Department of Chemistry and Biochemistry

University of Mississippi, University, MS 38677-1848 USA

Phone: 662 801 5526 Fax: 662 915 7300

Electronic Mail: dmbonsac@olemiss.edu

<http://quantum.chem.olemiss.edu/people/desiree.html>

I. Educational Background

Ph. D. Theoretical Chemistry	University of Mississippi	2011
(Advisor: Gregory S. Tschumper)	University, MS	(projected March)
B. S. Chemistry and B. S. Biochemistry	Winona State University	2006
(Advisors: Thomas Nalli and Jeanne Franz)	Winona, MN	

II. Research Experience

- Performed SAPT calculations to examine the composition of the interaction energy in select $\pi \cdots \pi$ systems.
- Determined CCSD(T) complete basis set limit relative energies for low-lying structures of the water hexamer.
- Investigating model chemistries for determining accurate electron affinities for molecules with highly delocalized singly occupied molecular orbitals.
- Exploring various QM:QM techniques for developing benchmark energies and struc-

tures of weakly bound clusters. This is part of a on going attempt to expand the chemical systems that can be characterized with highly accurate electronic structure methods.

III. Professional Experience

Computational Chemistry Cluster

System Administrator	University of Mississippi	2008-2011
Physical Chemistry Tutor	University of Mississippi	2008-2011
Research Assistant	University of Mississippi	2006-2011
Teaching Assistant	Winona State University	2004-2006

IV. Skills

Software Expertise:

ACES	CFOUR	Gaussian
Molpro	MPQC	PSI
PQS	Tinker	NWChem
ATMOL	Mathematica	MathCad

V. Professional Affiliations

American Physical Society (APS)

American Chemical Society (ACS)

VI. Honors and Awards

Disseration Fellowship Award	University of Mississippi	2010
Outstanding Poster Presentation	Conference on Current Trends in Computational Chemistry Jackson, MS	2008
Graduate Student Travel Award	University of Mississippi	2008
Undergraduate Research Grant	Winona State University	2005
Undergraduate Research Travel Grant	Winona State University	2005

VII. Professional Workshops

Center for Experimental

and Theoretical

Physical Chemistry

University of Mississippi

2009 and 2010

Telluride School on

Theoretical Chemistry

Telluride Science Research Center

2009

VIII. Publications

Electronic links to publications can be found at <http://quantum.chem.olemiss.edu>.

1. D. M. Bates, J. R. Smith and G. S. Tschumper, *J. Chem. Theory Comput. Submitted*; March 14, 2011 “Efficient and Accurate Methods for the Geometry Optimization of Water Clusters: Application of Analytic Gradients for the 2-body:Many-Body QM:QM Fragmentation Method to $(\text{H}_2\text{O})_n$, $n = 3 - 10$ ”
2. D. M. Bates, J. R. Smith, T. Janowski and G. S. Tschumper, *J. Chem. Phys. Submitted*; March 3, 2011 “Development of a 3-body:Many-body Integrated Fragmentation Method for Weakly Bound Clusters and Application to Water Clusters $(\text{H}_2\text{O})_n$ $n=3-10$ ”
3. D. M. Bates and G. S. Tschumper, *J. Phys. Chem. A* **113**, 3555–3559 (2009), “CCSD(T) Complete Basis Set Limit Relative Energies for Low -Lying Water Hexamer Structures.” (Cover article)

4. D. M. Bates J. A. Anderson, P. Oloyede and G. S. Tschumper, *Phys. Chem. Chem. Phys.* **10**, 2775–2779 (2008), “Probing the Effects of Heterogeneity on Delocalized $\pi \cdots \pi$ Interaction Energies.”

IX. Professional Presentations

1. “Life of a Graduate Student” Joint meeting for MCSR Computational Minicamp 2010 and Center for Experimental and Theoretical Physical Chemistry at Olemiss 2010 Summer Research Institute, July 2010
2. “CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures” 2009 NCNPR Cochran Center Symposium University of Mississippi, University, MS October 2009 [poster]
3. “CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures” Center for Experimental and Theoretical Physical Chemistry at Olemiss, MS June 2009
4. “CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures” Mississippi EPSCoR Conference, Starkville, MS April 2009[poster]
5. “Introduction to Quantum Chemistry” Winona State University, MN March 2009
6. “CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures” Sanibel Symposium St. Simons Island, GA February 2009 [poster]

7. "Probing the Effects of Heterogeneity on Delocalized $\pi \cdots \pi$ Interaction Energies" Seventeenth Annual Conference on Current Trends in Computational Chemistry Jackson, MS October 2008 [poster]
8. "Generating Benchmark Interaction energies for Weakly Bound Non-Covalent Clusters with the 2-Body:Many-Body Multicentered QM:QM Method" Eighth Triennial Congress of The World Association of Theoretical and Computational Chemists 2008 Sydney, Australia September 2008 [poster]
9. "Probing the Effects of Heterogeneity on Delocalized $\pi \cdots \pi$ Interaction Energies" Fall 2008 National Meeting of the American Chemical Society Philadelphia, PA August 2008 [poster]
10. "Introduction to Quantum Chemistry", "Introduction to Z-matrices" and "Introduction to Gaussian Software" Summer Lecture Series, University of Mississippi, MS June 2008
11. "Deterioration of popular DFT model chemistries for electron Affinities" Sigma Xi Student Research Symposium University of Mississippi, University, MS April 2008 [poster]
12. "Deterioration of popular DFT model chemistries for electron Affinities" Spring 2008 National Meeting of the American Chemical Society New Orleans, LA March 2008 [poster]
13. "Deterioration of popular DFT model chemistries for electron Affinities" Mississippi

Center for Super Computing Research Symposium University of Mississippi, University,
MS September 2007 [poster]

14. "Phosphonium Polymers" Spring 2005 National Meeting of the American Chemical
Society San Diego, CA March 2005 [poster]

X. References

Gregory S. Tschumper, Ph. D.

Associate Professor of Physical Chemistry

Department of Chemistry and Biochemistry

University of Mississippi

University, MS 38677

+1 662 915 7301

tschumpr@olemiss.edu

Nathan I. Hammer, Ph. D.

Assistant Professor of Physical Chemistry

Department of Chemistry and Biochemistry

University of Mississippi

University, MS 38677

+1 662 915 7301

nhammer@olemiss.edu