Three-Body Noncovalent Interactions: Insights from Energy Decomposition and their influence on Halogen Bonding

by

Sharon Alice Ochieng

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 9, 2025

Keywords: quantum chemistry, intermolecular interactions, symmetry-adapted perturbation theory, nonadditive interactions

Approved by

Konrad Patkowski, Chair, Professor of Chemistry and Biochemistry
Vincent Ortiz, Professor Emeritus of Chemistry and Biochemistry
Evangelos Miliordos, Associate Professor of Chemistry and Biochemistry
Steven Mansoorabadi, Associate Professor of Chemistry and Biochemistry

Abstract

Accurate models are crucial to predict noncovalent interactions, which govern complex structure, energetics, material self-assembly and reaction dynamics in complex environments. While there exist ample datasets of accurate interaction energies for bimolecular complexes, benchmark data for nonadditive three-body interactions are quite scarce, limiting the development and validation of methods that capture many-body contributions. This work introduces two datasets to study two and three−body noncovalent interactions in heteromolecular trimers. The 3BHET benchmark dataset comprises 20 equilibrium noncovalent interaction energies for a small but diverse selection of 10 heteromolecular trimers that combine different interactions including $\pi-\pi$, anion$-\pi$, cation$-\pi$ and various motifs of hydrogen and halogen bonding. The 3BXB datastet in contrast, features a diverse selection of 107 heteromolecular trimers in 214 structures formed by augmenting halogen−bonded dimers with either methane or water. It presents complexes that model $\sigma-hole$ interactions within each trimer, focusing on the magnitude and influence of three-body effects on the halogen bonds. Both datasets are constructed using the same level of theory and subjected to the same analysis process. The benchmark interaction energies were computed using the composite CCSD(T)/CBS method and a detailed energy decomposition of the two- and three-body interaction energies performed using various flavors and variants symmetry-adapted perturbation theory (SAPT). Additionally, we develop a six-dimensional potential energy surface (6D PES) for the OH-H_2 system using RCCSD(T) and CCSDT corrections for the symmetric geometires and MRCI-F12 for asymmetric configurations. This PES allows for the quantum scattering calculations for the computation of elastic and inelastic cross-sections.

Acknowledgments

I would like to thank all the people who contributed to this incredible journey. First and foremost, I would like to thank Dr. Konrad Patkowski for being a phenomenal advisor. He has provided me with opportunities and challenges that have made me a better scientist, and helped to broaden my horizons. His knowledge, talent, and passion have always been very inspiring and motivating. I would like to acknowledge the efforts and guidance of my thesis committee: Dr. Vincent Ortiz, Dr. Evangelos Miliordos and Dr. Steven Mansoorabadi. Each has provided their own distinct insight of the field of chemistry.

To my dad, Cosmas Ochanyo, mom, Judy Ochieng, my sincere thanks for bringing me up to always be smart and pursue my dreams. They always emphasized the importance of education and supported my academic career. I would like to thank my brother, Adrian Ochanyo, for unconditional love, good vibes and constant support. My friends who became family, Ayuba Akinpelu and Eva Mwakazi. Last but not least, I would like to thank Segun Obisesan who has been my constant companion for his unbelievable support, kindness, and an open heart. He always reminds me that there is life beyond science and puts a big smile on my face whenever I need it. I have felt so alive with our road trips and amazing adventures.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Computational chemistry plays a fundamental role in describing, understanding and even predicting important chemical phenomena using a hierarchy of theoretical methods. It has been particularly important in understanding weak interactions ; hydrogen, halogen bond, $\pi - \pi$ stacking, OH-π, CH-π and so forth, which govern molecular behavior. From determining spectroscopic and scattering data for bimolecular complexes, properties of clusters, liquids, and solids, ranking the relative stability of polymorphs of organic molecular crystals to probing differential host$-$guest binding in pharmaceutical design, it is important that these interactions are modeled precisely. Accurately predicting these interactions requires robust quantum mechanical methods, particularly for dispersion, a quantum-mechanical force arising from electron correlation. This necessity lays emphasis on the importance of developing methods and theories dedicated to accurately describe these subtle forces.

Two-body interactions have been shown to be the largest contributors to the total interaction energies in clusters, solids, and liquids with up to 81% for cage and prism water hexamers,[3] leading traditional modeling to assume pairwise additivity. However, this approximation is not sufficient to tell the whole story as these pairwise models fail to account for the three-body and higher-order cooperative effects.[4] In systems with more than two bodies, especially clusters and condensed-phase systems, the interactions between three or more bodies introduce many-body nonadditive effects that cannot be recovered by the summation of pairwise contributions (equation 1.1). Of particular interest are the three-body nonadditive effects as they have been shown to contribute significantly (up to 30% in water clusters) to the total interaction energy and their contributions significantly increase with an increase in system size.[3,5]

$$
\begin{aligned}
E^3_{int} &= E_{int} - E^2_{int} \\
E_{int} &= E_{ABC} - E_A - E_B - E_C \\
E^2_{int} &= (E_{AB} - E_A - E_B) + (E_{BC} - E_B - E_C) + (E_{AC} - E_A - E_C)
\end{aligned}
\tag{1.1}
$$

The subtle (often <1kcal/mol) nature of three-body nonadditive interactions presents a unique challenge in accurately modeling them. In comparison to investigating energetic quantities in covalent interactions where one would be satisfied with chemical accuracy, usually 1kcal/mol, nonadditive interactions demand the accuracy and robustness of CCSD(T)/CBS and SAPT methods.[6] However, the computational cost of these methods are prohibitive due to the steep scaling ($\mathscr{O}(N^7)$ for CCSD(T)[7] and $\mathscr{O}(N^5) - \mathscr{O}(N^7)$ for SAPT methods[8]). DFT-based dispersion models(DFT-D) have been shown to excel for neutral systems but the approximation faces limitations for ionic systems[9] and overall overestimate three-body interactions.[10] Therefore well-designed benchmarks become necessary to validate existing approximate methods, parametrize new methods in addition to revealing trends across chemically meaningful systems.[11]

The field has benefited from a wide variety of benchmarks created specifically to represent a class of interacting systems and create theoretical frameworks to achieve accuracy. Their importance cannot be underscored as this lies in their role in evaluating the various frameworks/models performance, strengths and weaknesses. Several reference specific datasets have been introduced that tackle interactions in chemically relevant systems like S22,[12] S66,[13] X40,[14] S12L,[15] NENCI,[16] or more recently the NCI atlas by Řezáč et al. with several datasets trying to map dimer interactions over the entire chemical space one interaction at a time.[17–21] Each dataset features a specific theme. For example X40, S22, S66 represent small to medium dimer systems with less than 40 atoms computed using a composite CCSD(T)/CBS scheme. The NENCI dataset by DiStasio and coworkers represents a large dataset (∼8000) which features a set of non-equilibrium geometries starting from the S66 and S101[22] datasets. The construction of these high quality benchmarks serves to address the growing need for extensive high-quality quantum mechanical data in the

field. The broad and carefully curated datasets that enable accurate validation of models and also ensure methodological performance is transferable are especially useful.

Despite the robust growth in the size and diversity of benchmark datasets, they remain disproportionately focused on two-body complexes. They lack comprehensive data on three-body effects, especially at the interplay of short-range dispersion and induction with long-range screening effects in trimers.[23] Given the importance of the nonadditive three-body effects in the description of molecular clusters and condensed phase systems, 3B-69[24] and S22(3)[25] provide high-level trimer benchmarks for homogeneous systems. In keeping up with emerging trends in machine learning (ML) for chemistry, Low[26] and coworkers constructed a dataset from 3B-69, S22(3) and a 509-trimer collection extracted from an enzyme-inhibitor complex to aid predicting the three-body interaction energy contribution in trimer systems of up to 77 atoms with the resulting accuracy being below CCSD(T) level. Even with this growth, these homogeneous datasets limit insights into chemically symmetric environments. More recently, the 3BHET dataset presented in detail in Chapter 3 provides high quality benchmark data but is not sufficient to effectively and systematically probe directional polarization in mixed electrostatics and dispersive systems and addressing cooperative effects among different groups. To investigate these phenomena systematically, halogen bonding offers an ideal platform to study nonadditivity across different interaction scales.

Halogen bonds (XB) are a kind of noncovalent interaction where the the halogen atom acts as an electrophile interacting with a Lewis base. They, as illustrated in Fig.1.1, consist of R$-$ X $\cdots$Y (where X = Cl, Br, or I; Y = Lewis base; R = substituent) characterized by the existence of a localized region of depleted electron density at the tip of the covalently bound halogen atom. This area of positive electrostatic potential that forms from the anisotropic distribution of electron

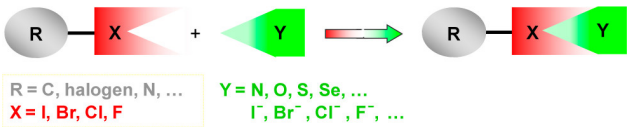


Figure 1.1: Schematic representation of the halogen bond.[1]

density is called the $\sigma-$ hole.[27,28] The $\sigma-$ hole anisotropy is responsible for the dual character

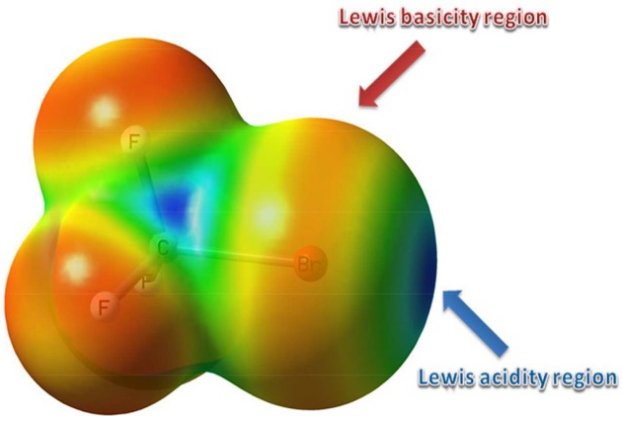


Figure 1.2: Electrostatic potential in CF_3Br molecule. Electron density isosurface of 0.006 e/bohr^3 value.[2]

(Fig.1.2) of the XB and thus the debates in the theoretical world regarding their nature as to whether they are primarily electrostatic or involve significant charge transfer and many-body effects.[29] As such, XB complexes make for an insightful framework to study many-body effects in part due to their strong directionality similar to hydrogen bonds (HB) and offering an excellent blend of electrostatics and dispersion which is enhanced by the polarizability of the atom that increases with atom size.[14] Furthermore, these characteristics allow for us to study the cooperative and nonadditive effects when a third molecule perturbs the interaction,[2] in that it can modulate the XB strength, geometry, or electrostatic profile depending on its placement and character.

Energy Decomposition Analysis (EDA) offers an important tool to elucidate and characterize the underlying physics of the nonadditive interaction energy component. Methods such as Symmetry-Adapted Perturbation Theory (SAPT)[30] , discussed in detail in Chapters 2,3, and 4, allow for the the decomposition of the interaction energy into physically meaningful components. It follows the antisymmetrization of the wavefunction with respect to interchanges of electrons between monomers[31] thus partitioning the interaction energy into electrostatics, exchange, induction and dispersion. This type of energy decomposition becomes particularly useful when studying the effects of cooperativity and identifying the dominant contributions in halogen bonding. In Chapters 3 and 4, we see in detail how SAPT analysis can reveal whether the addition of a third body to form a trimer enhances dispersion cooperativity or disrupts electrostatic alignment, helping to

rationalize observed trends.[2,32] These insights gained are not only limited to isolated trimers but are very important in force field design and development for use in various fields of chemistry such as crystal engineering, supramolecular chemistry,[23] and membrane−ligand interactions[28] where small nonadditive changes can shift equilibrium structures or binding preferences.

While the principal focus of this work is on understanding nonadditive three-body effects of noncovalent interactions, we simultaneously aim to develop an accurate six-dimensional potential energy surface (PES) for the OH-H_2 system. The OH radical is especially of interest due to its abundance in the atmosphere. In addition to controlling the removal of many trace gases in the atmosphere,[33] its chemistry is is vital in the network of reactions relevant to water and other oxygen-bearing molecules.[34] For our case, the OH-H2 system provides a prototype for tetratomic reactions that enables quantum scattering calculations which yield both elastic and inelastic cross-sections.[35,36] In addition, the rate coefficients derived from the collisions are of interest for astrophysical applications.[37] As an open-shell system, the OH radical presents a unique set of challenges, especially when interacting with closed shell species such as H_2. In particular, the OH-H_2 complex , exhibits two closely lying states, A' and A" in C_s symmetry. Therefore, to accurately describe the complex, particularly with the focus on vibrationally inelastic collisions, we employ explicitly correlated coupled cluster method [RCCSD(T)-F12b].[38] However, in the presence of near degeneracy, additional multireference methods are important in the description of correct energetic ordering and coupling between the states. For this reason, we employ MRCISD+Q with cluster corrected energies of Langhoff and Davidson.[39] This aspect is presented in detail in Chapter 5.

Chapter 2

Electronic Structure Theory

2.1. The Electronic Schrödinger Equation

The challenge of describing the behavior of multiple electrons presents a fundamental electronic problem. This complexity arises from the interaction of electrons with each other and with the nucleus, making it difficult to solve the Schrödinger equation exactly.[40] Thus at the heart of quantum chemistry lies the pursuit of approximate solutions to the non-relativistic time-independent Schrödinger equation which takes the form:

$$\hat{H}\Psi(r,R) = E\Psi(r,R) \tag{2.1}$$

where $\hat{H}$ is the total Hamiltonian of a system with N electrons and M nuclei. The wavefunction Ψ is dependent on $3N$ spatial coordinates r_i, and N spin coordinates σ_i collectively represented as x_i, and $3M$ spatial coordinates of the nuclei R_I.

$$\Psi \equiv \Psi(x_1, x_2, ...x_N, R_1, R_2...R_M) \tag{2.2}$$

The $\hat{H}$ operator presents the total interaction energy as sum of kinetic energy of electrons and the nuclei, electron-electron repulsion, electron-nuclear attraction and nucleus-nucleus repulsion. Presented in atomic units;

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>1}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.3}$$

The first and second terms of the equation describe the operator for the kinetic energy of the electrons and nuclei respectively. The Laplacian operators ∇_i^2 and ∇_A^2 involve differentiation with

respect to the *i*th electron and the *A*th nucleus. The remaining terms represent the Coulomb attraction between electrons and nuclei as well as the repulsion between electrons and between nuclei. Thus we can condense 2.3 as:

$$\hat{H} = \hat{T}_e + \hat{T_N} + \hat{V_{eN}} + \hat{V_{ee}} + \hat{V_{NN}} \tag{2.4}$$

The importance of equation 2.3 cannot be overstated, as it presents a rigorous framework for the description of molecules by accounting for all the motions of the electrons and the nuclei. The wavefuction Ψ possesses all the information of the system while the $\hat{H}$ governs the behavior of the system. However, due to the correlated nature of electrons and nuclei, solving the Schrödinger equation analytically for systems with more than one electron becomes mathematically unfeasible; it is only solvable for the hydrogen atom and other trivial systems. Chemical systems are neither simple nor trivial. As such, the Schrödinger equation is intractable for realistic systems and approximations are essential. In order to simplify the problem, the Born-Oppenheimer approximation is applied.

The Born-Oppenheimer approximation, one of the most fundamental approximations in quantum mechanics, takes advantage of the difference in mass between the electrons and the nuclei to decouple the nuclear and electronic motions.[41] This allows for the separation of the wavefunction and Hamiltonian to electronic and nuclear parts. The electronic part explicitly depends on the electronic coordinates *r* and parametrically on the nuclear coordinates *R*. This reduces the complexity of the Schrödinger equation which yields the electronic Hamiltonian that describes the motion of N electrons in the field of M point charges:

$$\hat{H_{elec}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}} \tag{2.5}$$

where the kinetic energy of the nucleus is omitted and the nuclear repulsion contribution is considered to be a constant. Any constant added to an operator only adds to the operator eigenvalues and has no effect on the operator eigenfunctions.

Nonetheless, even with this simplification, the resulting Schrödinger equation still proves to be computationally demanding especially due to electron correlation. To address this, *ab initio* quantum mechanical methods have been designed and developed to find approximate numerical solutions to the Schrödinger equation. They do so by solving the equation from first principles without relying on empirical parameters. These methods such as Hartree-Fock and post Hartree-Fock theories (MP2, CCSD) covered in the following sections allow for the symmetric approximations to the wavefunction. This allows us to predict molecular properties with increased accuracy.

2.2. Hartree-Fock Method

The Hartree-Fock approximation is central to quantum chemistry as it provides the simplest method to treat a many-particle system. It replaces the dynamic many-particle problem by an effective one-electron problem where the electron is assumed to be moving in an effective static potential field.[42,43] The Hartree-Fock method provides a formalism to find the best multi-particle state that can be expressed as a Slater determinant of single particle states. The elements of the Slater determinant are single electron spinorbitals with the spin and orbital part. Therefore, for a system with N electrons, the Slater[44,45] determinant takes the form:

$$\Psi_0(x_1,x_2,...,x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_N(x_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(x_N) & \chi_2(x_N) & \cdots & \chi_N(x_N) \end{vmatrix} \tag{2.6}$$

A nice feature of the Slater determinant is that it satisfies the Pauli's exclusion principle for fermions, i.e $(\chi_1 = \chi_2)$, then $\Psi_0(x_1,x_2) = 0$; as well as enforces antisymmetry when any two electrons are interchanged, $\Psi_0(x_1,x_2) = -\Psi_0(x_2,x_1)$. Given the simplified electronic Hamiltonian:

$$\widehat{H_{elec}} = \sum_i h(i) + \sum_{i<j} v(i,j) + V_{NN} \tag{2.7}$$

where the one electron operator $h(i)$:

$$h_(i) = \nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \tag{2.8}$$

and two electron operator $v(i,j)$

$$v(i,j) = \frac{1}{r_{i,j}} \tag{2.9}$$

with the Slater determinant (2.6), we can obtain the expectation value of the energy:

$$E_{el} = \langle \Psi | \hat{H} | \Psi \rangle \tag{2.10}$$

assuming the wavefunction is normalized.

Therefore, upon employing the variational theorem, we obtain better approximates to the wavefunction that minimize the energy by varying their parameters within the given functional space.[40]

$$E_{HF} = \sum_i \langle i|h|i\rangle + \frac{1}{2}\sum_{ij} \langle ii|jj\rangle - \langle ij|ji\rangle \tag{2.11}$$

where the one electron integral:

$$\langle i|h|j\rangle = \int dx_1 \chi_i^*(x_1) h(r_1) \chi_j(x_1) \tag{2.12}$$

and the two electron integral:

$$\langle ij|kl\rangle = \int dx_1 dx_2 \chi_i^*(x_1)\chi_j(x_1)\frac{1}{r_{12}}\chi_k^*(x_2)\chi_l(x_2) \tag{2.13}$$

It is possible to see that minimizing the E_{HF} in equation 2.11 under the constraint that the spin orbitals are orthonormal i.e $\langle \chi_i | \chi_j \rangle = \delta_{ij}$ is equivalent to solving the eigenvalue equation. Hence, we are able to obtain the set of spin orbitals which are the solutions to the Hartree-Fock equations:

$$F_i \chi_i = \varepsilon_i \chi_i \tag{2.14}$$

where F is the Fock operator defined as:

$$F_i = h_i + \sum_{j \neq i} J_j(i) - K_j(i), \tag{2.15}$$

The operators J and K are Coulomb and exchange operators.

$$J_j \chi_i(x_i) = \chi_i(x_i)[\int dx_j \chi_j^*(x_j) \frac{1}{r_{ij}} \chi_j(x_j)] \tag{2.16}$$

$$K_j \chi_i(x_i) = \chi_j(x_i)[\int dx_j \chi_j^*(x_j) \frac{1}{r_{ij}} \chi_i(x_j)] \tag{2.17}$$

For practical implementation of the Hartree -Fock theory to computation, it is prudent to make it a linear algebra problem. We introduce the basis sets in order to transform the HF to the Roothaan[46] equations denoting

$$\chi_i(r) = \sum_{\mu}^{k} C_{\mu i} \Phi_\mu \tag{2.18}$$

$$FC = SC\varepsilon \tag{2.19}$$

While Hartree-Fock has known shortfalls in modeling electron correlation, it provides a foundation through the Roothaan equations, and a set of virtual orbitals. These are useful for correlated wavefunction methods.

2.3. Correlated Methods

2.3.1. Second-Order Møller-Plesset (MP2) Perturbation Theory

While the Hartree-Fock theory provides a fundamental foundation, it suffers from a major limitation: it neglects electron correlation, especially dynamical correlation that is a result of electron repulsion. Perturbation theory was developed to overcome this issue. The Second-order Møller-Plesset perturbation theory (MP2) belongs to the family of the Many-Body Perturbation Theories (MBPT). It happens to be the simplest and one of the most useful levels of theory beyond the Hartree-Fock approximation. Starting from a Hartree-Fock wavefunction as a reference, it takes into account multiple excitations from the unperturbed occupied orbitals to the unoccupied

ones.[40,43] Hence, the exact Hamiltonian can now be written as:

$$H = H_0 + \lambda V \tag{2.20}$$

where V is the small perturbation, λ is a dimensionless parameter and H_0 is the Hartree-Fock Hamiltonian. The exact wavefunction and energy can be expanded as the following infinite series respectively:

$$\Psi = \Psi_0 + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \lambda^3\Psi^{(3)} + \cdots \tag{2.21}$$

$$E = E_0 + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \cdots \tag{2.22}$$

Upon substituting equations 2.21 and 2.22 into the Schrödinger equation and collecting terms that have the same powers of λ :

$$\begin{gathered} H_0\Psi_0 = E_0\Psi_0 \\ H_0\Psi^{(1)} + V\Psi_0 = E_0\Psi^{(1)} + E^{(1)}\Psi_0 \\ H_0\Psi^{(2)} + V\Psi^{(1)} = E_0\Psi^{(2)} + E^{(1)}\Psi^{(1)} + E^{(2)}\Psi_0 \\ \cdots \\ H_0\Psi^{(n)} + V\Psi^{(n-1)} = \textstyle\sum_{k=0}^{k} E^{(k)}\Psi^{(n-k)} \end{gathered} \tag{2.23}$$

and multiplying each of the above equations by Ψ_0 and integrating over all space yields the following expression for the nth-order (MPn) energy:

$$\begin{gathered} E^{(0)} = \langle\Psi_0|H_0|\Psi_0\rangle \\ E^{(1)} = \langle\Psi_0|V|\Psi_0\rangle \\ E^{(2)} = \langle\Psi_0|V|\Psi^{(1)}\rangle \\ \cdots \\ E^{(n)} = \langle\Psi_0|V|\Psi^{(n-1)}\rangle \end{gathered} \tag{2.24}$$

Consequently, the first-order perturbed wavefunction is obtained by taking a linear combination of the zeroth-order excited states that forms a complete set:

$$\Psi^{(1)} = \sum_n c_n^{(1)} \Psi_n^{(0)} \tag{2.25}$$

where the $c_n^{(1)}$ are coefficients to be determined. In Møller Plesset perturbation theory (MP2), where the zeroth-order Hamiltonian is a sum of Fock operators:

$$H_0 = \sum_i F_{(i)} \tag{2.26}$$

and the perturbation V as:

$$V = \sum_{i<j}^{N} \frac{1}{r_{12}} - \sum_j^N (J_{(i)} - K_{(i)}) \tag{2.27}$$

Therefore, the HF energy

$$E^{HF} = \langle \Psi_0 | H_0 + V | \Psi_0 \rangle \tag{2.28}$$

corresponds to the sum of zeroth- and first-order energies

$$E^{HF} = E_0 + E^{(1)} \tag{2.29}$$

while higher order corrections account for the correlation energy.

$$E_{corr} = E_0^{(2)} + E_0^{(3)} + E_0^{(4)} + \ldots \tag{2.30}$$

For MP2, the first-order wavefunction and second-order energy are denoted by:

$$\begin{aligned} \Psi^{(1)} &= \tfrac{1}{4} \textstyle\sum t_{ab}^{ij} |\Psi_{ab}^{ij}\rangle \\ E_0^{(2)} &= -\tfrac{1}{4} \textstyle\sum_{ab}^{virt} \sum_{ij}^{occ} \frac{|\langle ab||ij\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \end{aligned} \tag{2.31}$$

where i and j are occupied orbitals, a and b are unoccupied virtual orbitals, t_{ab}^{ij} is the double excitation amplitude from the HF wavefunction and the ε quantities represent the corresponding HF orbital energies. The MP2 equations above include double excitations which account for a majority of the correlation energy that is missed by Hartree Fock theory.

2.3.2. Coupled Cluster Theory

The coupled-cluster(CC) theory is the leading quantum-chemical wavefunction approach for the description of electronic structure in small and medium-sized systems, especially where the effects of electron correlation are important. The hierarchy of approximations in this family of methods gives an effective description of the electron correlation while retaining the size consistency even in a truncated form. This combination of robustness and accuracy has made it an attractive method for noncovalent interactions as it is able to recover $98\% - 99\%$ of the interaction energy at the CBS limit in the absence of static correlation.[47,48] It uses the exponential expansion to a reference wavefunction Φ_0:

$$\Psi_{CC} = e^{T}\Phi_0 \tag{2.32}$$

where the cluster operator T that determines the type of excitations taken into account is outlined as:

$$T = T_1 + T_2 + T_3 + \cdots + T_N \tag{2.33}$$

and due to the prohibitive ($\mathcal{O}(N^{2n+2})$ where N is the number of basis functions and n is the excitation level) scaling, usually truncated for practical purposes. The excited Slater determinants are generated from the reference:

$$T_1\Phi_0 = \sum_{i}^{occ}\sum_{a}^{virt} t_i^a \Phi_i^a \tag{2.34}$$

$$T_2\Phi_0 = \sum_{i<j}^{occ}\sum_{a<b}^{virt} t_{ij}^{ab} \Phi_{ab}^{ij} \tag{2.35}$$

where the t_i^a and t_{ij}^{ab} represent single and double-excitation amplitudes, respectively. One can substitute the exponential function by a Taylor series:

$$\begin{aligned} e^T &= 1 + T + \tfrac{1}{2}T^2 + \tfrac{1}{6}T^3 + \tfrac{1}{24}T^4 + \cdots + \tfrac{1}{k!}T^k + \cdots \\ &= 1 + T_1 + T_2 + \tfrac{1}{2}T_1^2 + T_3 + T_2T_1 + \tfrac{1}{6}T_1^3 + T_4 + T_3T_1 + \tfrac{1}{2}T_2^2 + \tfrac{1}{2}T_1^2T_2 + \tfrac{1}{24}T_1^4 + \cdots \end{aligned} \tag{2.36}$$

Based on the truncation level of the exponential operator e^T , we are able to obtain a variety of approximations. Coupled-cluster singles and doubles (CCSD) results from $T_1 + T_2$, coupled-cluster singles, doubles and triples (CCSDT) from $T_1 + T_2 + T_3$, and CCSDTQ from $T_1 + T_2 + T_3 + T_4$. The successive inclusion of higher excitations yields the full configuration interaction solution to the electronic Schödinger equation. However, while CCSD captures correlation effects, it is still not precise enough for applications such as weak interactions to the submillihartree accuracy. Therefore, it is prudent to include higher order excitations to guarantee accuracy. This comes at a high price of computational scaling of $\mathcal{O}(N^8)$ for CCSDT for example. Alternatively, the triple excitations can be introduced perturbatively as in CCSD(T)[7] that offers better computational scaling $\mathcal{O}(N^7)$ while matching the accuracy of CCSDT. The noniterative treatment of triples makes this approximation have a better cost to accuracy ratio, making it the "gold" standard method for noncovalent interaction benchmarks.[7,49]

CCSD(T), like any other approximation, suffers from a set of drawbacks. In addition to exhibiting slow basis set convergence, it cannot be applied rigorously for multi-reference systems like the case of open shell fragments and/or spaces where near degeneracies are observed. In the case of systems such the OH-H_2, with multiple electronic states, we would require appropriate methods to reference the multi-reference character of the system.

2.3.3. Multireference Configuration Interaction (MRCI)

Configuration interaction (CI) works on the general idea to calculate the many-electron wave function of an electronic state as a linear combination of pre-defined configuration state functions (CSF). Explicity, the wavefunction is written as:

$$|\Psi_0\rangle = c_o\,|\Phi_o\rangle + \sum_{ar} c_a^r\,|\Phi_a^r\rangle + \sum_{\substack{a<b\\r<s}} c_{ab}^{rs}\,|\Phi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c\\r<s<t}} c_{abc}^{rst}\,|\Phi_{abc}^{rst}\rangle + \cdots \tag{2.37}$$

where $|\Psi_0\rangle$ represents the many electron wavefuction (full CI), $|\Phi_o\rangle$ represents the HF determinant which is the reference wavefuction, $|\Phi_a^r\rangle$ a singly excited determinant where an electron from occupied spin orbital a is excited to the virtual orbital r up until the Nth excitation, c_o are expansion coefficients optimized variationally. CI calculations aim at getting the few lowest eigenvalues and the corresponding eigenvectors of the CI Hamiltonian matrices.[50] For full CI (FCI) calculations, the linear expansion continues up to and including the CSF in which all n electrons are promoted from occupied to virtual molecular spin orbitals. FCI results are exact in a given basis and therefore account for the complete electron correlation energy. These calculations are computationally too demanding so a truncation scheme is applied in order to keep the number of CSFs at a manageable size. However, these truncated methods lack size-extensivity; the non-interacting systems will lead not to the same energy as the sum of calculations on the isolated systems.[51]

In addition to FCI and truncated CI methods being costly and not size-extensive, they are single-reference and break down in systems with strong static correlation. Multi-Reference CI (MRCI) addresses this by using a multi-configurational reference space. MRCI calculations are performed by first constructing a space of reference configurations and then allowing a fixed number of excitations out of this reference space. The Complete Active Space Self-Consistent Field(CASSCF) offers the most common way to define this reference space. The CASSCF method allows all possible occupations within the active space and variationally optimizes both the orbitals and the CI coefficients.[52] The excitations, from the CASSCF method thus the foundation for MRCI, cover the dynamical correlation while the reference covers the non-dynamic electron correlation.[53] While both static and dynamical correlation are adequately captured through the excitations and multi-configuratonal reference, it is still not size-extensive. The size extensivity problem is partially remedied *a posteriori* by adding an energetic correction term at the end of an MRCI computation.[39]

$$E_{MRCI+Q} = E_{MRCI} + \Delta E_{DC} \tag{2.38}$$

$$\Delta E_{DC} = (1 - c_0^2)\Delta E \tag{2.39}$$

$$c_0^2 = \sum_{k=1}^{N_{ref}} (c_k^{ref})^2 = \sum_{k=1}^{N_{ref}} \langle \Psi | \Phi_k^{ref} \rangle^2 \tag{2.40}$$

2.3.4. Symmetry-Adapted Perturbation Theory

Symmetry-adapted perturbation theory (SAPT) provides a solid framework for calculating weak interactions between molecules. Being a perturbative method, it allows for the direct computation of the interaction energy as a perturbation to the Hamiltonian of the individual monomers. A key advantage of SAPT over supermolecular approaches is that in addition to not requiring error cancellation when computing interaction energies, it provides a decomposition of the interaction into physically meaningful components of electrostatics, exchange, induction, and dispersion.[54] This allows us to draw meaningful insights into the nature of these interactions and their physical origins.

The SAPT expansion starts from the assumption that the monomers are isolated (unperturbed) and their electronic Schrödinger equations can be solved exactly.[31]

$$\begin{gathered} H = H_A + H_B + V_{AB} = H_0 + V_{AB} \\ H_A \Psi_A^{(0)} = E_A^{(0)} \Psi_A^{(0)} \quad H_B \Psi_B^{(0)} = E_B^{(0)} \Psi_B^{(0)} \quad \Psi_{AB}^{(0)} = \Psi_A^{(0)} \Psi_B^{(0)} \quad E_{AB}^{(0)} = E_A^{(0)} + E_B^{(0)} \end{gathered} \tag{2.41}$$

The zeroth-order Hamiltonian is further split into the Fock operator and the intermolecular interaction and intramolecular electron correlation operators

$$H = F + V + W \tag{2.42}$$

where in the two-body case

$$F = F_A + F_B \quad V = V_{AB} \quad W = W_A + W_B \tag{2.43}$$

and in the three-body case

$$F = F_A + F_B + F_C \quad V = V_{AB} + V_{BC} + V_{AC} \quad W = W_A + W_B + W_C \tag{2.44}$$

F represents the Fock operator, inter-monomer interaction is added via Coulomb terms V whereas a correction to the Hartree-Fock description (intra-monomer correlation) is captured by W. In the simplest approximation SAPT0, W is neglected and thus we have the Hamiltonian as $H = F + V$. From this, we obtain the simple SAPT0 method that provides energy decomposition but does not include intra-molecular electron correlation.

$$\begin{aligned} E_{int}^{\text{SAPT0}} =& E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} \\ &+ \delta E_{HF}^{(2)} \end{aligned} \tag{2.45}$$

where $\delta E_{HF}^{(2)}$ accounts for higher-order induction and exchange-induction effects from a super-molecular HF calculation.[8,31]

$$\begin{aligned} E_{int}^{\text{SAPT2}} =& E_{elst}^{(10)} + E_{elst,resp}^{(12)} \\ &+ E_{exch}^{(10)} + E_{exch}^{(11)} + E_{exch}^{(12)} \\ &+ E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{ind}^{(22)} + E_{exch-ind}^{(22)} + \delta E_{HF}^{(2)} \\ &+ E_{disp}^{(20)} + E_{exch-disp}^{(20)} \end{aligned} \tag{2.46}$$

$$E_{int}^{\text{SAPT2+}} = E_{int}^{\text{SAPT2}} + E_{disp}^{(21)} + E_{disp}^{(22)} \tag{2.47}$$

$$E_{int}^{\text{SAPT2+(3)}} = E_{int}^{\text{SAPT2+}} + E_{elst,resp}^{(13)} + E_{disp}^{(30)} \tag{2.48}$$

$$\begin{aligned} E_{int}^{\text{SAPT2+3}} =& E_{int}^{\text{SAPT2+(3)}} + [\delta E_{HF}^{(3)} - \delta E_{HF}^{(2)} + E_{ind,resp}^{30} + E_{exch-ind,resp}^{30}]_{ind} \\ &+ [E_{exch-disp}^{(30)} + E_{ind-disp}^{(30)} + E_{exch-ind-disp}^{(30)}]_{disp} \end{aligned} \tag{2.49}$$

The gradual improvements to the SAPT0 approximation are presented in 2.46 to 2.49. With regard to intramonomer electron correlation SAPT2 includes electrostatics, exchange, and induction up to second order. In SAPT2+ and higher, the intramonomer corrections to dispersion are accumulated via second-order.[8]

Alternatively, for SAPT(DFT) based on the Kohn-Sham description of the monomers, the Fock operator is replaced by the Kohn-Sham operator such that $K = K_A + K_B$ or $K = K_A + K_B + K_C$ for dimers and trimers, respectively.

$$\begin{aligned} E_{int}^{SAPT(DFT)} &= E_{elst}^{(1)}(KS) + E_{exch}^{(1)}(KS) + E_{ind}^{(2)}(CKS) + E_{exch-ind}^{(2)}(CKS) \\ &\quad + E_{disp}^{(2)}(CKS) + E_{exch-disp}^{(2)}(CKS) + \delta E_{int,2}^{HF} \end{aligned} \tag{2.50}$$

Since standard DFT fails to account for the correct asymptotic behavior of the exchange-correlation potential, the thoroughly benchmarked PBE0 functional with asymptotically corrected exchange-correlation potential is employed.[55] In addition, the induction and dispersion components are computed with orbital relaxation through coupled frequency-dependent density susceptibility (FDDS).[32,56,57] The exchange counterparts to dispersion and induction are estimated by scaling the uncoupled dispersion and induction respectively. For us, scaling occurs only in 3-body

SAPT(DFT).

$$\widetilde{E}^{(2)}_{exch-ind,2}(CKS) = E^{(2)}_{exch-ind,2}(KS)\frac{E^{(2)}_{ind,2}(CKS)}{E^{(2)}_{ind,2}(KS)} \tag{2.51}$$

$$\widetilde{E}^{(2)}_{exch-disp,2}(CKS) = E^{(2)}_{exch-disp,2}(KS)\frac{E^{(2)}_{disp,2}(CKS)}{E^{(2)}_{disp,2}(KS)} \tag{2.52}$$

The 3-body SAPT(DFT) framework solely captures the nonadditive interaction energy as shown in equation 2.53. The components comprise induction, dispersion, and exchange effects that arise solely due to the presence of all three monomers.[58]

$$\begin{aligned} E^{\text{SAPT(DFT)}}_{int,3} =& E^{(1)}_{exch,3}(\text{KS}) + E^{(2)}_{ind,3}(\text{CKS}) + \widetilde{E}^{(2)}_{exch-ind,3}(\text{CKS}) \\ &+ E^{(3)}_{disp,3}(\text{CKS}) + \widetilde{E}^{(2)}_{exch-disp,3}(\text{CKS}) + \delta E^{\text{HF}}_{int,3} \end{aligned} \tag{2.53}$$

The $\delta E^{\text{HF}}_{int,3}$ term takes into account higher-order induction effects that are computed in the supermolecular Hartree-Fock computed in the dimer centered basis set (DCBS)

$$\delta E^{\text{HF}}_{int,3} = E^{\text{HF}}_{int,3} - E^{(10)}_{exch,3} - E^{(20)}_{ind,3} - E^{(20)}_{exch-ind,3} \tag{2.54}$$

The extended variant of SAPT (XSAPT), as the name suggests, allows for the extension of SAPT to an arbitrary number of monomers. This is through he many body explicit polarization formalism, XPol,[59,60] that works by partitioning the system to a number of fragments which are completely independent of each other. In SAPT, the electrostatics, exchange and second-order dispersion decay much faster with intermolecular distance and therefore pairwise additivity serves as a good approximation. Polarization on the other hand exhibits strong nonadditive character and cannot be assumed to be just the sum of pairwise effects. Therefore, we introduce the many body polarization effects variationally through XPol method:

$$E_{XPol} = \sum_{A=1}^{n}[2\sum_{a} c_a^{\dagger}(h^A + J^A - \frac{1}{2}K^A)c_a + E^A_{nuc}] + E_{embed} \tag{2.55}$$

where the $[2\sum_a c_a^\dagger(h^A + J^A - \frac{1}{2}K^A)c_a + E_{nuc}^A]$ is the HF energy for monomer A and E_{embed} is the electrostatic embedding potential.

In combining SAPT with XPol, we take the zeroth-order wave functions to be direct products of XPol monomer wave functions. In that equation 2.55 for a dimer is incorporated into SAPT.

$$E_{XSAPT} = \sum_A (\sum_a [2\varepsilon_a^A - c_a^\dagger(J^A - \frac{1}{2}K^A)c_a] + E_{nuc}^A) + \sum_A \sum_{B>A} (E_{RSPT}^{[0;1_{AB}]} + E_{exch}^{[0;1_{AB}]} + E_{RSPT}^{[0;2_{AB}]} + E_{exch}^{[0;2_{AB}]}) + E_{3B} \tag{2.56}$$

The $[0;1_{AB}]$ notations shows that its is zeroth order in monomer fluctuation potential while products[61] of monomer wavefunctions computed at the XPol level are used as zeroth-order wavefunctions in the SAPT treatment. Moreover, the SAPT dispersion[62] and exchange−dispersion terms are replaced by an empirical force field term fitted from ab initio SAPT data[63] for small molecules.[64–66] More recently, there has been a dispersion correction based on the many body dispersion(MBD) formalism to incorporate the nonpairwise-additive dispersion with the goal of significantly improving accuracy for larger systems.[67] Alternatively, the nonadditive dispersion can be incorporated through the Axilrod-Teller-Muto(ATM) correction which provides a simplified approximation that is computationally efficient in capturing the three-body dispersion effects.[68]

2.4. Noncovalent Interaction Energy Computations

The accurate modeling of noncovalent interactions presents a fundamental challenge. These subtle interactions of small magnitudes relative to total molecular energies require accurate treatment of many body effects; ranging from pairwise interactions, three-body nonadditive interactions, up to the complex n-th body interactions in larger systems.[3,69] The quantum mechanical methods presented in this section tackle this challenge through two fundamental approaches/frameworks. The first of them is the supermolecular framework where the interaction energy in a dimer and trimer respectively is given by:

$$\begin{aligned} E_{int} &= E_{AB} - E_A - E_B \\ E_{int} &= E_{ABC} - E_A - E_B - E_C \end{aligned} \tag{2.57}$$

and in a non-additive context can be decomposed as

$$E_{int} = E_{int}^2 + E_{int}^3 \tag{2.58}$$

where the two-body part is the sum of pairwise interactions:

$$E_{int}^2 = (E_{AB} - E_A - E_B) + (E_{BC} - E_B - E_C) + (E_{AC} - E_A - E_C) \tag{2.59}$$

An advantage of the canonical supermolecular approach is that it is conceptually simple and easy to apply at any distance between interacting molecules.[48] Although quantum mechanics is riddled with many methods to directly compute the r.h.s of equations 2.57, E_{int} is 4-7 orders of magnitude smaller than the E_{ABC}, E_A, E_B, E_C and its accuracy relies on meaningful error cancellations. Therefore, the careful selection of a model with the appropriate theory and basis sets is necessary to guarantee accuracy and efficiency. Another drawback of the supermolecular approach arises from basis set incompleteness called basis set superposition error (BSSE). This is the artificial lowering of the monomer energies due to "borrowing" of basis functions from the other interacting monomer. The standard solution to remedy this is through the counterpoise (CP) correction.[70]

$$\begin{aligned} E_{int} &= E_{AB}^{AB} - E_A^{AB} - E_B^{AB} \\ E_{int} &= E_{ABC}^{ABC} - E_A^{ABC} - E_B^{ABC} - E_C^{ABC} \end{aligned} \tag{2.60}$$

This procedure involves performing all calculations in a dimer-centered basis set (DCBS) (a trimer-centered basis set for a trimer) where each monomer calculation uses "ghost" basis functions where the partner's atoms would be. This way, the artificial lowering of energies cancels out during subtraction, diminishes with an increase in basis and completely vanishes in the infinite basis.[70]

While the supermolecular approach with the CP correction provides a practical framework, the resultant interaction energy is only one number that gives no physical insight to the nature of

the interaction. Consequently, the perturbative approach, often through symmetry-adapted perturbation theory (SAPT), provides rigorous means for computing such quantities directly. This allows for the decomposition of the interaction into physically meaningful components.

2.5. Basis Sets

The foundation of quantum chemistry (as seen in section 2.1) relies on the explicit modeling of the electronic distribution in molecules. This is done by solving the Schrödinger equation to varying levels of approximations. For systems in the gas phase, these approximations can be grouped into two, the level of theory (discussed in sections 2.2 and 2.3) and the choice of basis set used. The former simplifies the Schrödinger equations while the latter provides a mathematical framework with which the electron distribution in the system of interest can be expressed.[71,72] Together, these two form the model for determining the accuracy and computational cost of describing the electronic structure of molecular and extended systems. While the choice of method is constrained by several factors such as computational cost, accuracy and systems size and nature, the choice of basis relies on the need to efficiently compute the daunting two-electron integrals that describe electron-electron correlation. Electronic structure makes use of two types of basis, Slater and Gaussian type orbitals, with the latter being preferred because it offers better computational efficiency.[72]

$$\phi_i = \sum_j c_{ij} \chi_j(r) \tag{2.61}$$

where:

$$G_\alpha(r) = N e^{-\alpha r^2} \tag{2.62}$$

corresponds to the radial form of the primitive gaussian which represents the s-type function with zero angular momentum. A more general form to describe the functions with non-zero angular momentum (p,d,f,.. functions) takes the form:[73,74]

$$G_{l,m}(r,\theta,\phi) = N \cdot r^l \cdot Y_m^l(\theta,\phi) e^{-\alpha r^2} \tag{2.63}$$

where $Y_m^l(\theta,\phi)$ introduces angular dependence and

$$\chi_j(r) = \sum_j d_{jk} G_{\alpha k}(r) \tag{2.64}$$

where d_{jk} is the contraction coefficients with k being the number of primitives within the contracted basis function and r is the distance between the electron and the atom in question.

Our work extensively makes use of Dunning correlation-consistent bases aug-cc-pVXZ $\equiv$ aXZ (X = D, T, Q, 5).[75,76] To allow for the systematic convergence of the correlation energy, we extrapolate the correlation energy to the CBS-limit. This is done by taking into account the assumption that the correlation energies follow the X^{-3} convergence pattern in the Dunning basis sets and that the CBS-limit correlation energy can be computed from the results of two consecutive basis sets.

$$E_\infty^{corr} = E_X^{corr} + \frac{(1-1/X)^3}{1-(1-1/X)^3}(E_X^{corr} - E_{X-1}^{corr}) \tag{2.65}$$

Furthermore, to achieve efficiency without sacrificing accuracy, we employ density fitting (DF). This approach involves complementing the standard orbital basis set with an auxiliary basis set denoted by adding -RI/-MP2FIT and -JKFIT.

Chapter 3

Accurate three-body noncovalent interactions: the insights from energy decomposition

This chapter was published as: Noncovalent intermolecular interactions, besides determining spectroscopic and scattering data for bimolecular complexes, are critical for the properties of clusters, liquids, and solids. In these cases, the structure and properties of the system at hand is determined by the cooperative effect of all interactions,[4] not just the two-body forces between pairs of molecules. While the nonadditive three-body interactions are individually much smaller in comparison to their two-body counterparts, their importance increases rapidly with the cluster size, thus very subtly influencing the electronic structure of the interacting species.[3,77,78]

The magnitude of interactions and the importance of three-body nonadditive effects influences the choice of methods used to describe them either theoretically or experimentally. In principle, the most stringent test of a theoretical model is its comparison with experimental data. However, the interaction potentials between molecules are usually measured indirectly through various kinds of gas-phase spectroscopy.[79,80] This, however, comes with a limitation as the experimental data samples a range of intermolecular configurations simultaneously, and one also has to subtract the zero-point vibrational energy (ΔZPVE), making it hard to obtain interaction energies at the desired level of accuracy.[81] These limitations have made it a common practice to test the interaction energies of more approximate theoretical methods against accurate benchmark calculations. Therefore, benchmark databases such as S22,[12] S66,[13] X40,[14] or S12L[15,23] have been vital in quantifying the accuracy of computational methods as well as optimizing new density functionals, semiempirical methods, or machine learning approaches.[48,81]

Assessing the accuracy, reliability, and transferability of an electronic structure method requires broad testing on systems that are chemically diverse.[48,82] A well balanced benchmark set should aim to provide a statistically meaningful overview of the performance and allow for the

comparison of various methods. For example, the X40 dataset[14] provides a focused description of halogenated complexes including hydrogen and halogen bonding, while the S66 dataset[13] includes a broader description of bio-relevant dimers, including varied motifs of hydrogen bonding and $\pi - \pi$ interactions. While both of these databases focus on the van der Waals minimum geometries, they also have their X40x10 and S66x8 counterparts that include points across the entire dissociation curve.[13,14] Such databases involving both equilibrium and non-equilibrium points provide data for the parameterization of empirical and semiempirical methods towards the description of attractive and repulsive noncovalent interactions.[81]

In recent years, the number and breadth of benchmark databases of bimolecular complexes has greatly expanded. Burns et al.[63] sampled the entire Protein Data Bank of crystal structures identifying 3380 different conformations of interacting amino acid side chains. Řezač et al.[17] are in the process of developing an atlas of noncovalent interactions aimed at mapping the entire chemical space, one specific interaction type at a time. The first datasets of the project describe hydrogen bonding in neutral and ionic organic dimers,[17] also those including sulfur, phosphorus or halogens.[18] The next dataset[19] focuses on repulsive contacts in the same kinds of organic molecules. Additional datasets have been constructed to populate the atlas, focusing on sigma-hole interactions (halogen, chalcogen, and pnictogen bonds) and London dispersion interactions, sampling an extensive range of the chemical space.[20,21] Also recently, DiStasio and coworkers developed an extensive benchmark dataset of $\sim$8000 non-equilibrium geometries of bimolecular complexes,[83] aiming to adequately describe noncovalent interactions across the entire potential energy surfaces. The use of such broad databases ensures that the observed performance of approximate methods is transferable between systems and between different regions of the potential energy surface.

Compared to this extensive coverage of two-body complexes, there is limited benchmark data available to describe diverse types of nonadditive three-body interactions relevant to molecular clusters and condensed-phase systems. The most extensive dataset available is 3B-69[24] which

comprises accurate interaction energies in 69 trimers extracted from 23 molecular crystal structures (thus, every trimer has three identical molecules). Accurate data is also available for numerous structures of specific systems, such as the water trimer[84] and the benzene trimer.[85,86] Gordon and coworkers[25] constructed a number of trimers from the dimers in the S22 set[12] to create the S22(3) dataset. Their study focuses on investigating and quantitatively estimating the many-body dispersion effects for different bonding motifs in molecular clusters, and the trimers have 1–2 types of nonequivalent monomers (thus, they are of either the AAA or AAB type). More recently, Low et al.[26] combined the 3B-69 and S22(3) datasets with a larger, 509-trimer collection extracted from an enzyme-inhibitor complex to develop a machine learning approach to predict three-body interaction energies. This new collection involves trimers that are heteromolecular and also larger (up to 77 atoms), but the reference interaction energies were only computed with a tailored version of the spin-component-scaled MP2 approach[87] and it is not entirely clear if they all are of benchmark CCSD(T) quality.

While the 3B-69 and S22(3) datasets provide valuable reference energies and establish the performance of various electronic structure methods, primarily for complexes of identical molecules, they provide limited data for modelling different kinds of interactions simultaneously. Thus, one often needs a broader set of reference values, including also heteromolecular trimers. In the long run, such a broad dataset can establish a "dictionary" of three-body interaction types, translating between intuitive (but "fuzzy") classifications of the constituent dimers (hydrogen bonding, $\pi-\pi$ stacking, halogen bonding, $\cdots$) and quantitative numerical data obtained from benchmark supermolecular calculations and symmetry-adapted perturbation theory (SAPT) energy decomposition. This allows one to provide a chemically based classification of different interactions present in trimers. As a step towards this goal, we construct a new dataset of accurate three-body intermolecular interactions in ten heteromolecular trimers (20 structures) computed at the CCSD(T) level of theory at the complete basis set (CBS) limit. The complexes include up to 3 distinct monomers and represent a range of interactions including hydrogen and halogen bonding, π-π

stacking, cation/anion-π, CH-π, and OH-π bonds. The dataset is a mix of global and local minima.

A wavefunction-based, CCSD(T)-level benchmark reference is essential for quantifying nonadditive interactions and assessing the performance of more approximate methods. Density functional theory (DFT), using multiple exchange-correlation functionals, has failed to predict accurately three-body interaction energies, especially dispersion.[88] According to Jankiewicz et al.,[88] the main reason for the failure is the inability to remove the deficiencies of the functional using various dispersion corrections. This issue persists even with the many-body dispersion (MBD) extension by Tkatchenko and coworkers,[89] which formally incorporates correct many-body physics but does not necessarily predict more accurate three-body energies than the bare functionals. It should be noted that, in the MBD formalism, a "body" is an atom rather than an entire molecule, so this approach captures, to a reasonable approximation, all-order long-range correlation effects of both intra- and intermolecular character. The three-body dispersion (where each “body” is a different molecule) can be extracted from the MBD calculations in a standard supermolecular manner, by subtracting the pairwise terms from the interaction energy of the trimer.

To quantitatively analyze the diversity of interactions in our dataset, we employ the energy decomposition analysis provided by SAPT at both the two-body and three-body level. We compare the decompositions predicted by different SAPT variants including SAPT(DFT),[56,57] SAPT2+(3),[8] and XSAPT,[66] allowing us to assess the strengths and weaknesses of each approach (SAPT2+(3) is restricted to pairwise additive interactions while both SAPT(DFT) and XSAPT account for some nonadditive effects). To visualize and interpret these interactions, we expand upon the ternary diagram representation commonly used for bimolecular complexes.[82,90] In this extended representation, we incorporate the nonadditive three-body effects for both SAPT(DFT) and XSAPT. This approach enables us to analyze and interpret the complex interplay between different interaction components in our dataset, providing valuable insights into the nature and strength of these interactions.

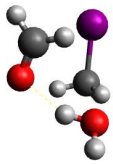
3.1. Computational Methods

3.1.1. Composition of the Data Set

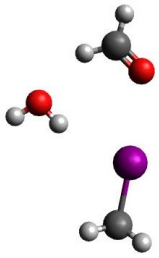
The goal of this work is to design a small but diverse database of heteromolecular trimers with an equally diverse range of noncovalent interaction types. The 3BHET dataset constructed here consists of 10 trimers, with the starting geometries obtained from dimers in the BEGDB database[12] by adding the third molecule and reoptimizing the geometry. The species included (see Figures for pictures) were specifically chosen to represent a variety of interactions relevant to bioorganic chemistry that include hydrogen and halogen bonding, π-π stacking, cation/anion-π, CH-π, and OH-π interactions. For each system, two geometries are studied, with one being the global minimum of the trimer, and the other one a local minimum of a different binding character than the global one (whenever possible).

3.1.2. Geometries

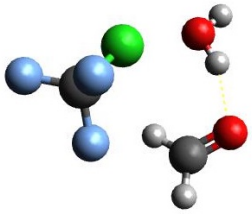
Geometry optimizations were carried out using MP2/aug-cc-pVDZ with counterpoise correction and tight convergence criteria. Frequency calculations were also carried out to ascertain that the geometries obtained were true global or local minima. Out of the 20 structures considered, 19 turned out to correspond to actual minima (no imaginary frequencies) while the last one (complex 5.2) was a high-symmetry saddle point. Both global and local minima obtained were used for the benchmark calculations and SAPT analysis, and several of them were later selected as starting points for radial potential energy curves. In such a case, the geometries and mutual orientations of all molecules are conserved as the intermolecular center-of-mass distance is scaled. This scaling is done by shifting one molecule away from the other two molecules in the complex, thereby creating a 1D potential curve. The minimum-energy intermolecular distance is scaled along this direction by factors ranging from 0.7 to 3.0. Note that three curves of this kind are obtained for each trimer as any one molecule can be translated relative to the other two.



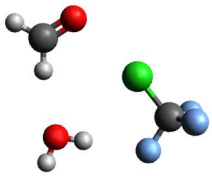
1.1 Iodomethane-Formaldehyde-Water (hydrogen bonds)



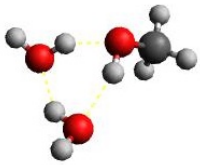
1.2 Iodomethane-Formaldehyde-Water (hydrogen and halogen bonds)



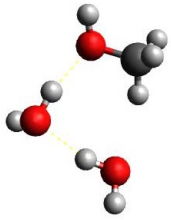
2.1 Trifluorochloromethane-Formaldehyde-Water (hydrogen bonds)



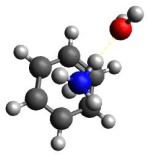
2.2 Trifluorochloromethane-Formaldehyde-Water (halogen and hydrogen bonds)



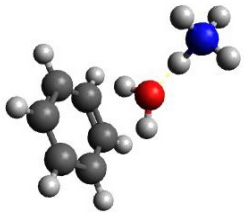
3.1 Methanol-Water-Water (hydrogen bonds)



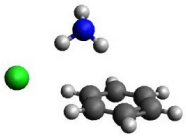
3.2 Methanol-Water-Water (hydrogen bonds)



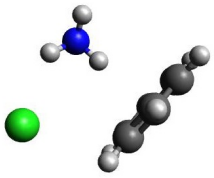
4.1 Ammonium-Benzene-Water (cation-π and hydrogen bonds)



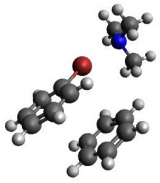
4.2 Ammonium-Benzene-Water (OH-π and hydrogen bonds)



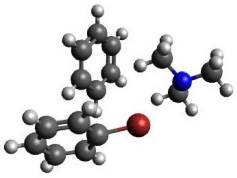
5.1 Chloride-Ammonia-Benzene (NH-π and hydrogen bonds)



5.2 Chloride-Ammonia-Benzene (NH-π, anion-π, and hydrogen bonds)



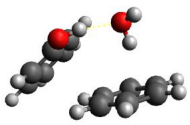
6.1 Bromobenzene-Trimethylamine-Benzene (halogen and $\pi-\pi$ bonds)



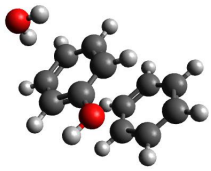
6.2 Bromobenzene-Trimethylamine-Benzene (CH-π bonds)

Figure 3.1: Global and local-minimum geometric arrangements of trimers present in the dataset.

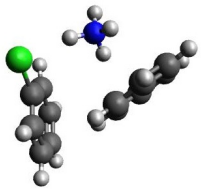
We performed the benchmark and two-body SAPT(DFT) calculations using the MOLPRO2012.1 program.[91] For SAPT2+(3) calculations, the Psi4 code was used.[92] Nonadditive SAPT(DFT) results were obtained using the three-body module of the SAPT package,[93] importing the DFT quantities from DALTON 2.0.[94] Finally, the XSAPT computations utilized the Q-Chem code.[95]



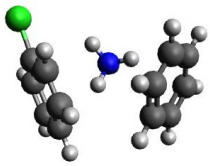
7.1 Phenol-Benzene-Water ($\pi-\pi$, OH-π, and hydrogen bonds)



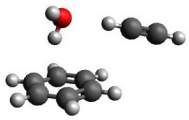
7.2 Phenol-Benzene-Water ($\pi-\pi$, OH-π, and hydrogen bonds)



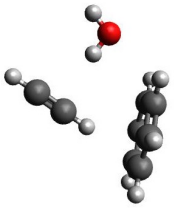
8.1 Chlorobenzene-Benzene-Ammonium (cation-π, CH-π, and hydrogen bonds)



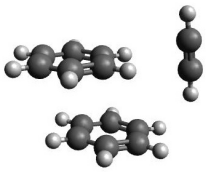
8.2 Chlorobenzene-Benzene-Ammonium (cation-π and hydrogen bonds)



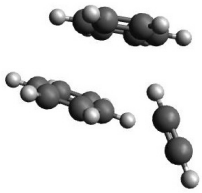
9.1 Acetylene-Benzene-Water (OH-π and CH-π bonds)



9.2 Acetylene-Benzene-Water (OH-π and CH-π bonds)



10.1 Acetylene-Benzene-Benzene (CH-π and π-π bonds)



10.2 Acetylene-Benzene-Benzene (CH-π and π-π bonds)

Figure 3.2: Global and local-minimum geometric arrangements of trimers present in the dataset.

3.1.3. Benchmark Two-Body and Three-Body Interaction Energies The nonadditive three-body interaction energy E_{int}^3 is defined as the difference between the trimer interaction energy and the sum of the pair interaction energies.[31] The total interaction energy in a trimer

$$E_{int} = E_{ABC} - E_A - E_B - E_C \tag{3.1}$$

can be decomposed as

$$E_{int} = E_{int}^2 + E_{int}^3 \tag{3.2}$$

where the two-body part is the sum of pairwise interactions:

$$E_{int}^2 = (E_{AB} - E_A - E_B) + (E_{BC} - E_B - E_C) + (E_{AC} - E_A - E_C) \tag{3.3}$$

The monomer and dimer calculations were performed in the full trimer basis set to eliminate basis set superposition error (BSSE). In addition, the interaction energies were computed using fixed monomer geometries ignoring all monomer deformation effects.

The benchmark interaction energies at both the two-body and three-body level were calculated using the composite CCSD(T)/CBS approach where correlation consistent basis sets fully augmented with diffuse functions were used; in this text, aug-cc-pVXZ (X = D,T,Q,5,...) are abbreviated as aXZ.[76] The composite interaction energy is constructed from the contributions of the Hartree-Fock method, the MP2 correlation energy, and a small-basis CCSD(T) correction:

$$E^{(\mathrm{CCSD(T)/CBS})} = E^{\mathrm{HF}} + E_{\mathrm{corr}}^{\mathrm{MP2/CBS}} + \delta E^{\mathrm{CCSD(T)}} \tag{3.4}$$

In this way, the MP2 term covers a large section of interaction energy at the CBS limit. However, MP2 is known to overestimate two-body dispersion effects in many systems[81] and it entirely misses three-body dispersion; therefore, a CCSD(T) correction $\delta E^{\mathrm{CCSD(T)}} = E^{\mathrm{CCSD(T)}} - E^{\mathrm{MP2}}$ in a moderate basis set is needed for benchmark accuracy. With the latter correction computed in a double-zeta basis set, this level of theory can be classified as the *silver standard* of accuracy, suitable for nearly all benchmark purposes.[96,97] As each trimer contains three pairwise interactions plus a comparatively small nonadditive effect, it is reasonable to expect a silver-standard approach to yield absolute interaction energy errors about three times larger than in the dimer case, which is still adequate for our benchmark purposes.

The Hartree-Fock energy converges fast with the basis set size. Therefore, a single HF calculation in a large enough basis set is sufficient, and we have used the a5Z basis set for this purpose. The MP2 correlation energy is extrapolated to the CBS limit from the aQZ and a5Z basis sets using Helgaker's X^{-3} formula.[98] The coupled cluster correction $\delta E^{\mathrm{CCSD(T)}}$ is calculated in aDZ for larger complexes and aTZ for a few smaller ones. Density fitting of two-electron integrals was employed in all MP2 computations in order to make calculations for larger complexes feasible.

3.1.4. SAPT Energy Decomposition

To adequately describe and interpret the supermolecular benchmark data obtained in this work, we need to know the nature of the three-body binding in each trimer in relation to its two-body counterpart. Therefore, the two-body and three-body interaction energies are decomposed into physically meaningful quantities using several SAPT variants. In the two-body and three-body SAPT based on the HF wavefunctions, the Hamiltonian is partitioned into the Fock operator F, the intermolecular interaction operator V, and the intramolecular correlation operator W encompassing all monomers:

$$H = F + V + W \tag{3.5}$$

where

$$F = F_A + F_B + F_C \qquad V = V_{AB} + V_{BC} + V_{AC} \qquad W = W_A + W_B + W_C \tag{3.6}$$

The simplest approximation that provides qualitatively reasonable total interaction energies is the SAPT0 one, which totally neglects the intramonomer electron correlation.[31] The two-body SAPT0 interaction energy is calculated as

$$\begin{aligned} E_{int}^{\text{SAPT0}} =& E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} \\ &+ \delta E_{HF}^{(2)} \end{aligned} \tag{3.7}$$

In a SAPT correction $E^{(kl)}$ in Eq. (4.7) and throughout the text, the subscripts k and l denote orders of perturbation theory with respect to V and W, respectively. SAPT0 provides the physical components that represent the electrostatic, first-order exchange, induction, exchange-induction (both with the inclusion of monomer relaxation/response effects as denoted by the additional subscript "resp"), dispersion, exchange-dispersion, and $\delta E_{HF}^{(2)}$ contributions. The $\delta E_{HF}^{(2)}$ term is meant to estimate the higher-order induction and exchange-induction effects from a supermolecular HF

calculation.[8,31] This energy decomposition allows one to classify complexes according to the relative importance of electrostatics, induction, and dispersion for their binding, for example, using a ternary diagram.[82]

To attain quantitative accuracy of total energies and their contributions, a higher-order SAPT level that includes both intra- and intermonomer electron correlation effects is required. In this study, the two-body interaction energy decomposition was performed by wavefunction-based SAPT at the SAPT2+(3) level of theory, simultaneously including the effects of the W operator on most SAPT0-level corrections and some terms that are third-order in V:[8]

$$\begin{aligned} E_{int}^{\mathrm{SAPT2+(3)}} =& E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} \\ &+ E_{disp}^{(20)} + E_{exch-disp}^{(20)} + \delta E_{HF}^{(2)} \\ &+ E_{elst,resp}^{(12)} + E_{exch}^{(11)} + E_{exch}^{(12)} + E_{ind}^{(22)} + E_{exch-ind}^{(22)} \\ &+ E_{disp}^{(21)} + E_{disp}^{(22)} + E_{disp}^{(30)} + E_{elst,resp}^{(13)} \end{aligned} \tag{3.8}$$

3.1.4.1. SAPT(DFT)

An alternative, Kohn-Sham based variant of SAPT (SAPT(DFT)) was also used to decompose the two-body and three-body interaction energies.[56,58,99] In SAPT(DFT), the Fock operator F in the partitioned Hamiltonian is replaced by the Kohn-Sham operator $K = K_A + K_B + K_C$. Because the Kohn-Sham operator includes effective electron correlation, the intramonomer electron correlation operator W can be neglected altogether, making SAPT(DFT) scale much better with system size as compared to wavefunction-based SAPT.[99] In the early studies of the two-body SAPT(DFT) method, the observed inaccurate interaction energies were attributed to the common DFT functionals being unable to provide exchange-correlation potentials with the correct asymptotic behavior.[56,99] This problem can be remedied by performing the SAPT(DFT) calculations (using the PBE0 functional as recommended in the literature[55]) with an asymptotically corrected exchange-correlation potential. To do this, the difference between the highest occupied molecular

orbital (HOMO) energies of the individual monomers and their true ionization potentials was used. However, the asymptotic correction alone is still insufficient to obtain an accurate description of the complete interaction energy. The induction and dispersion SAPT(DFT) terms, both two-body and three-body, need to be computed from monomer frequency-dependent density susceptibilities (density response functions) within the coupled KS (CKS) formalism.

Moving onto the SAPT(DFT) description of nonadditive three-body effects,[58] it is important to note that electrostatics and second-order dispersion are rigorously pairwise additive and do not constitute part of the three-body interaction energy; thus, one needs to go to the third order of SAPT to estimate nonadditive dispersion. The bulk of the nonadditive interaction energy is typically accounted for by induction; additionally, all SAPT exchange corrections are nonadditive starting from the first-order one. The exchange counterparts of the three-body dispersion and induction terms are scaled according to the ratios of the CKS and KS values for the corresponding induction and dispersion corrections (equations (3.9) and (3.10), with the additional subscript ",3" denoting the nonadditive three-body correction):[58,99]

$$\widetilde{E}^{(2)}_{exch-ind,3}(\mathrm{CKS}) = E^{(2)}_{exch-ind,3}(\mathrm{KS})\frac{E^{(2)}_{ind,3}(\mathrm{CKS})}{E^{(2)}_{ind,3}(\mathrm{KS})} \tag{3.9}$$

$$\widetilde{E}^{(2)}_{exch-disp,3}(\mathrm{CKS}) = E^{(2)}_{exch-disp,3}(\mathrm{KS})\frac{E^{(3)}_{disp,3}(\mathrm{CKS})}{E^{(3)}_{disp,3}(\mathrm{KS})} \tag{3.10}$$

For two-body SAPT2+(3) and SAPT(DFT), the calculations used the aTZ basis set (with the aTZ-PP effective core potential (ECP) for the iodine atom). Furthermore, density fitting with the aTZ-RI (SAPT(DFT)) and def2-QZVPP-RI (SAPT2+(3)) auxiliary bases was used to reduce the computational cost. The three-body SAPT(DFT) calculations, performed with the three-body code[58] from the SAPT2008 suite,[93] are computationally demanding, making the aDZ basis the only practical option to accommodate larger complexes. The truncated jun-cc-pVDZ[100] basis set was used for one complex (**6**) as it was not feasible to run it in the fully augmented basis, and a fully decontracted aug-cc-pVTZ-DK3 set was used for the iodine atom in complex **1** as this

calculation does not support ECPsBoth the two-body and three-body SAPT(DFT) methods require an asymptotic correction to overcome the qualitative inaccuracy of the exchange-correlation potential and kernel. In the three-body SAPT(DFT) method, we employed the Fermi-Amaldi asymptotic correction, which involves shifting the appropriate potential for the asymptotic region by $(I+\varepsilon_{HOMO})$ and smoothly splicing[101] it to match a short-range potential without derivative discontinuity.[99,102] The two-body SAPT(DFT) method employed the gradient-regulated asymptotic connection (GRAC) procedure. This scheme enables the construction of smooth asymptotically corrected potentials[103,104] by splicing the PBE0 functional with the asymptotically correct LB94 functional developed by van Leeuwen and Baerends.[105]

The characterization of each of the complexes in terms of two-body and three-body interaction types was done based on the effects that dominated the binding (attractive) contribution. The numerical values in two-body SAPT2+(3) and SAPT(DFT) calculations are obtained by summing the interaction energy components from each dimer in the trimer. The nonadditive interaction energy is entirely obtained from the three-body SAPT(DFT) calculation, with second-order exchange terms scaled according to Eqs. (3.9) and (3.10):

$$\begin{aligned} E_{int,3}^{\mathrm{SAPT(DFT)}} =& E_{exch,3}^{(1)}(\mathrm{KS}) + E_{ind,3}^{(2)}(\mathrm{CKS}) + \widetilde{E}_{exch-ind,3}^{(2)}(\mathrm{CKS}) \\ &+ E_{disp,3}^{(3)}(\mathrm{CKS}) + \widetilde{E}_{exch-disp,3}^{(2)}(\mathrm{CKS}) + \delta E_{int,3}^{\mathrm{HF}} \end{aligned} \tag{3.11}$$

In equation (3.11), higher-order terms derived for the three-body wavefunction-based SAPT, such as third-order induction and induction-dispersion corrections,[106] are not included as they are likely mostly quenched by their exchange counterparts which have not been derived. Therefore, the higher-order nonadditive induction and exchange-induction effects are estimated with supermolecular HF.[58,99] This correction covers the difference between the nonadditive HF interaction energy term and the sum of its SAPT contributions, the first-order exchange and second-order induction

and exchange-induction:

$$\delta E^{\mathrm{HF}}_{int,3} = E^{\mathrm{HF}}_{int,3} - E^{(10)}_{exch,3} - E^{(20)}_{ind,3} - E^{(20)}_{exch-ind,3} \tag{3.12}$$

Here, $E^{\mathrm{HF}}_{int,3}$ represents the nonadditive interaction energy from a supermolecular Hartree-Fock approach, and the other components are obtained from the three-body SAPT0 calculations.

3.1.4.2. Extended SAPT (XSAPT)

The computational cost of performing SAPT calculations quickly becomes prohibitive beyond simple molecular trimers. Employing a different flavor of SAPT (XSAPT) that incorporates many-body polarization into the monomer wavefunctions attempts to remedy this issue, also allowing the SAPT calculations to be extended to an arbitrary number of monomers.[65,66] This many-body polarization is included within the explicit polarization (XPol) scheme[59] by partitioning the system into noncovalently interacting fragments and performing DFT or HF calculations for each fragment in the embedding of point charges representing the electric field of its neighbors. This process is iterated until the atomic charges resulting from the DFT/HF electron density are the same as the atomic charges defining the embedding. At convergence, the densities and orbitals are used to compute standard pairwise-additive SAPT0 or SAPT(KS) energy corrections. Thanks to the XPol scheme, these corrections additionally contain both nonadditive (three- and higher-body) induction and some higher-order two-body induction effects, for example, the third-order term resulting from the density of molecule A polarized by unperturbed molecule B interacting with the density of molecule B polarized by unperturbed molecule A.[107] The dispersion and exchange-dispersion contributions in XSAPT are not calculated explicitly, but are instead taken from the many-body dispersion (MBD) model of Tkatchenko and co-workers.[89] This leads to significant computational savings (as dispersion and exchange-dispersion terms are by far the most time-consuming part of SAPT0 or SAPT(KS)), avoids the prevalent overestimation of dispersion by SAPT0 (note that the SAPT(KS) dispersion is no better unless it is computed in the CKS formalism like in SAPT(DFT)), and improves basis set convergence.[108] The resulting XSAPT method provides an estimate of the

nonadditive three-body induction (from the XPol embedding) and dispersion (from MBD) but neglects the nonadditive first-order exchange.

The total interaction energy in XSAPT is computed as

$$E_{int}^{\mathrm{XSAPT}} = E_{elst}^{(1)} + E_{exch}^{(1)} + E_{Disp}^{(2)} + [E_{ind}^{(2)} + E_{exch-ind}^{(2)} + \sum_{A}\sum_{B>A} \delta E_{AB}^{\mathrm{HF}} + \sum_{A}\sum_{B>A} (E_{AB}^{\mathrm{XSAPT}} - E_{AB}^{\mathrm{SAPT(KS)}}) + E_{int}^{\mathrm{MB}}] \tag{3.13}$$

where

$$\delta E_{AB}^{\mathrm{HF}} = E_{int}^{HF} - (E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)}) \tag{3.14}$$

with all terms on the r.h.s. computed for the dimer AB. The $\delta E_{AB}^{\mathrm{HF}}$ term takes into account higher-order induction effects for the A-B pair. Thus, it is assumed that, for systems with multiple bodies (molecules), this correction can be treated as a sum of pairwise contributions ($\delta E_{int}^{\mathrm{HF}} = \sum_A \sum_{B>A} \delta E_{AB}^{\mathrm{HF}}$). The terms in Eq. (3.13) were obtained from various stages of the XSAPT calculation. The first-order electrostatics and exchange terms, and a part of the overall (induction plus exchange-induction) effect, were obtained from SAPT(KS) calculations performed without electrostatic embedding. The $\delta E_{int}^{\mathrm{HF}}$ effect, which is classified as part of the overall induction energy, was computed from pairwise HF and SAPT0 calculations. The remaining parts of the XSAPT induction energy in Eq. (3.13) come from the effects of mutual polarization accounted for by the XPol embedding. Such effects can be partitioned into a pairwise additive and nonadditive part. The former is computed as a difference between the "total energies" of each pair computed using XSAPT (with the XPol embedding) and SAPT(KS) (without embedding) and thus takes into account some of the higher-order but two-body induction and exchange-induction effects. Note that the "total energies", the sums of monomer energies and the interaction term, are not commonly employed in the SAPT formalism, but are necessary in XSAPT as the XPol embedding effectively mixes terms of different orders of perturbation theory. The nonadditive part of the embedding effect, denoted E_{int}^{MB} in Eq. (3.13), contains the additional $E^{\mathrm{XSAPT}} - E^{\mathrm{SAPT(KS)}}$ "total energy" difference for the trimer beyond the differences resulting from the pairwise dimer calculations. Thus,

this effect approximates the actual induction nonadditivity in the trimer, in all orders of perturbation theory. Finally, the dispersion term $E^{(2)}_{Disp}$ (including, effectively, the exchange-dispersion contribution) was exclusively treated with the many-body dispersion model[89] (note that, in the MBD context, the bodies are individual atoms; in this sense, MBD accounts for both intra- and intermolecular dispersion nonadditivity).

The XSAPT approach is also capable of computing additional three-body induction couplings, derived and implemented in Ref. 109. These couplings do not occur in dimer calculations and can be exclusively attributed to the nonadditive part of the interaction energy. Note that the calculation of the couplings is quite cumbersome as they are not available in the newer and more efficient, atomic orbital-based implementation of XSAPT,[110] and one has to use the older molecular-orbital code. The XSAPT results were computed in the aTZ basis set for all atoms other than iodine and a fully decontracted aug-cc-pVTZ-DK3 set for iodine. In the XSAPT data below, the three-body induction couplings will be omitted unless explicitly stated otherwise. Overall, the XSAPT treatment includes the three-body nonadditivity in induction/exchange-induction (through the effects of the XPol embedding and possibly also the three-body induction couplings) and dispersion/exchange-dispersion (through the MBD formalism), but neglects the nonadditivity in the first-order exchange energy. At the two-body level, one might wonder if the addition of the $\delta E^{\mathrm{HF}}_{int}$ correction to XSAPT is rigorous, as this term accounts for some of the same higher-order induction effects as the XPol embedding. However, any possible double counting seems to be practically inconsequential and the addition of $\delta E^{\mathrm{HF}}_{int}$ substantially improves the XSAPT accuracy for hydrogen-bonded complexes.[108] At the nonadditive three-body level, the $\delta E^{\mathrm{HF}}_{int}$ term does not contribute so no double counting can occur.

3.2. Results and Discussion

3.2.1. Benchmark Interaction Energies

The construction of the 3BHET dataset was targeted at creating a dictionary of complexes that provides an example translation between the chemical, qualitative interaction types and the numerical, quantitative data. Such a dictionary tries to match the SAPT decompositions to the interaction

types present in the trimolecular complex. For this purpose, first, reference energies were established by performing single-point interaction energy calculations on two minimum geometries for each complex (in one case, only one minimum was found so we picked a high-symmetry saddle point geometry as the second one).

Table 3.1 provides a summary of the CCSD(T)/CBS benchmark total and nonadditive interaction energies. The complete interaction energies range from -36.2 kcal/mol down to -4.9 kcal/mol. Complexes with an ionic monomer exhibited the largest binding energies whereas $\pi-\pi$ stacked complexes showed interaction energies that were smaller in magnitude as compared to the rest of the dataset. The benchmark nonadditive contributions in Table 3.1 can be positive or negative and are up to 4–5 kcal/mol in magnitude, with the largest contributions observed for complexes where one of the monomers is ionic. The relative magnitude of the nonadditive terms ranges from nearly zero for complex **7.2** (in which two of the molecules are completely separated by the third one), to 1–5% for nonpolar systems **6** and **10**, to up to 14% for the hydrogen-bonded and/or ionic systems **3** and **4**.

The benchmark data in Table 3.1 allow us to estimate the accuracy of some approximate methods of calculating the total interaction energy and its nonadditive component. For this purpose, we use the mean unsigned errors (MUE) and mean unsigned relative errors (MURE) relative to the benchmark values. In this way, in Table 3.2 we quantify the accuracy of several approximate methods in reproducing the benchmark data. The MP2/CBS data include the correlation energy extrapolated from aQZ and a5Z, while the SAPT2+(3) and SAPT(DFT) calculations were performed in aTZ and aDZ, respectively. However, three-body SAPT(DFT) calculations of one of the larger complexes (number 6) was only feasible in the jun-cc-pVDZ basis set.

Table 3.2 shows the errors averaged over the van der Waals minimum geometries for the 20 trimers considered in this work (two structures for each system). The SAPT2+(3) approach is only available for two-body interaction energies, and we did not calculate three-body uncoupled dispersion needed for the SAPT0 E^3_{int}, so the three-body errors for these two methods are omitted.

Table 3.1: Benchmark two-body and three-body interaction energy data (kcal/mol). The CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. (4.4). The numbering of complexes is specified in Figs. 3.1 and 3.2.

Complex	MP2/CBS			CCSD(T)/CBS			Chemical Formula	Interaction Type[a]
	E_{int}^2	E_{int}^3	E_{int}	E_{int}^2	E_{int}^3	E_{int}		
1.1	−9.917	−0.563	−10.480	−9.772	−0.546	−10.317	$CH_3I-H_2CO-H_2O$	HB
1.2	−6.698	−0.577	−7.275	−6.466	−0.574	−7.041	$CH_3I-H_2CO-H_2O$	HB,XB
2.1	−8.306	−0.428	−8.734	−8.293	−0.405	−8.697	$CF_3Cl-H_2CO-H_2O$	HB
2.2	−8.079	−0.566	−8.645	−8.129	−0.559	−8.688	$CF_3Cl-H_2CO-H_2O$	HB,XB
3.1	−14.601	−2.491	−17.091	−14.573	−2.421	−16.994	$CH_3OH-H_2O-H_2O$	HB
3.2	−12.561	−1.619	−14.180	−12.523	−1.593	−14.115	$CH_3OH-H_2O-H_2O$	HB
4.1	−40.779	3.387	−37.391	−39.618	3.386	−36.232	$NH_4^+-C_6H_6-H_2O$	HB,cation-π
4.2	−28.656	−4.641	−33.297	−27.766	−4.623	−32.389	$NH_4^+-C_6H_6-H_2O$	HB,OH-π
5.1	−19.492	1.225	−18.267	−18.888	1.338	−17.550	$Cl^--NH_3-C_6H_6$	HB,NH-π
5.2 [b]	−13.798	1.127	−12.671	−12.974	1.171	−11.803	$Cl^--NH_3-C_6H_6$	HB,NH-π,anion-π
6.1	−11.479	0.232	−11.248	−7.206	0.347	−6.859	$C_6H_5Br-(CH_3)_3N-C_6H_6$	π-π, XB
6.2	−10.195	−0.227	−10.422	−7.211	−0.116	−7.327	$C_6H_5Br-(CH_3)_3N-C_6H_6$	CH-π
7.1	−15.593	−1.677	−17.271	−12.968	−1.559	−14.527	$C_6H_5OH-C_6H_6-H_2O$	HB,π-π,OH-π
7.2	−9.905	0.013	−9.891	−6.240	−0.009	−6.249	$C_6H_5OH-C_6H_6-H_2O$	HB,π-π,OH-π
8.1	−40.830	3.746	−37.084	−37.576	3.789	−33.788	$C_6H_5Cl-C_6H_6-NH_4^+$	HB,CH-π,cation-π
8.2	−40.069	4.084	−35.986	−37.114	4.152	−32.962	$C_6H_5Cl-C_6H_6-NH_4^+$	HB,cation-π
9.1	−8.193	−0.726	−8.919	−7.261	−0.669	−7.930	$HCCH-C_6H_6-H_2O$	OH-π,CH-π
9.2	−8.090	−0.550	−8.640	−6.883	−0.510	−7.392	$HCCH-C_6H_6-H_2O$	OH-π,CH-π
10.1	−9.655	−0.300	−9.955	−6.746	−0.200	−6.947	$HCCH-C_6H_6-C_6H_6$	CH-π,π-π
10.2	−8.452	−0.042	−8.494	−4.975	0.056	−4.919	$HCCH-C_6H_6-C_6H_6$	CH-π,π-π

[a]HB and XB stand for hydrogen and halogen bonding, respectively.
[b]This structure, unlike all the other ones, is a saddle point rather than a minimum.

Table 3.2: Error statistics of two-body, nonadditive three-body, and total interaction energies for different methods with CCSD(T)/CBS as the reference for 20 trimolecular complexes studied in this work. The MUE data are in kcal/mol and MURE in percent.

Method	E_{int}^2MUE	E_{int}^3MUE	E_{int}MUE	E_{int}MURE
HF/a5Z	10.154	0.153	10.030	104.318
SAPT0	2.818		2.957	27.361
XSAPT	2.008	0.205	1.873	19.240
XSAPT + $3b_{ind}$	2.008	0.206	2.059	19.954
MP2/CBS	1.613	0.055	1.665	18.600
SAPT(DFT)	1.720	0.085	1.710	12.814
SAPT2+(3)	0.280		1.391	7.823
CCSD(T)[a]	1.589	0.014	1.582	15.307

[a] Pure CCSD(T) in the same aDZ/aTZ basis as used for the $\delta E^{\mathrm{CCSD(T)}}$ term in the benchmark data.

Upon examining the error statistics in Table 3.2, it is evident that XSAPT outperforms SAPT0 (a 19% vs. 27% MURE for the total interaction energy). However, the performance of XSAPT

is inferior compared to SAPT(DFT), especially for the nonadditive three-body component. Moreover, the inclusion of three-body induction couplings actually makes the XSAPT nonadditive contributions and total interaction energies slightly worse (of course, this inclusion does not change the two-body XSAPT values). Among the approximate methods in Table 3.2, the best performer for the two-body interaction energy is by far SAPT2+(3). This excellent accuracy has to be to some extent coincidental, but it makes SAPT2+(3) give the lowest MUE (1.4% kcal/mol) and MURE (7.8%) for the total energy even though none of the nonadditive effects are recovered. Concerning the latter effects, even the simple HF method provides a decent description of the three-body term, with a MUE of 0.15 kcal/mol (although with a large MURE of 67% indicating that small nonadditive effects are not well described). Among the SAPT flavors studied, only SAPT(DFT) leads to more accurate three-body energies than HF, but the two correlated parts of the composite benchmark data, MP2/CBS and CCSD(T)/aDZ (aTZ for a few systems), each lead to even smaller three-body errors. The good performance of MP2/CBS for E^3_{int} indicates that nonadditive dispersion, missing at the MP2 level, is not particularly important for our systems.

3.2.2. Potential Energy Curves

Potential energy curves for three selected complexes were obtained through displacing a single molecule from the equilibrium geometry by scaling the intermolecular center-of-mass distance, keeping the separation and mutual orientation of the other two molecules unchanged. In this way, three 1D potential curves are generated passing through a particular global/local minimum. These curves have been computed at the same CCSD(T)/CBS level as all the minimum structures.

In Figures 3.3–3.5 , we highlight the potential energy curves of three complexes that showcase cation-π, halogen, and hydrogen bonding interactions. Furthermore, the interaction energy along these curves has been decomposed using two-body SAPT(DFT). The intermolecular distances R reported in these figures are scaled relative to the global-minimum separation: therefore, 1.0 on the horizontal scale always represents the minimum configuration which is the same for all three curves (which correspond to the displacements of a different monomer relative to the other two). The interactions tend to be relatively small at larger separations but do not tend to zero. This is

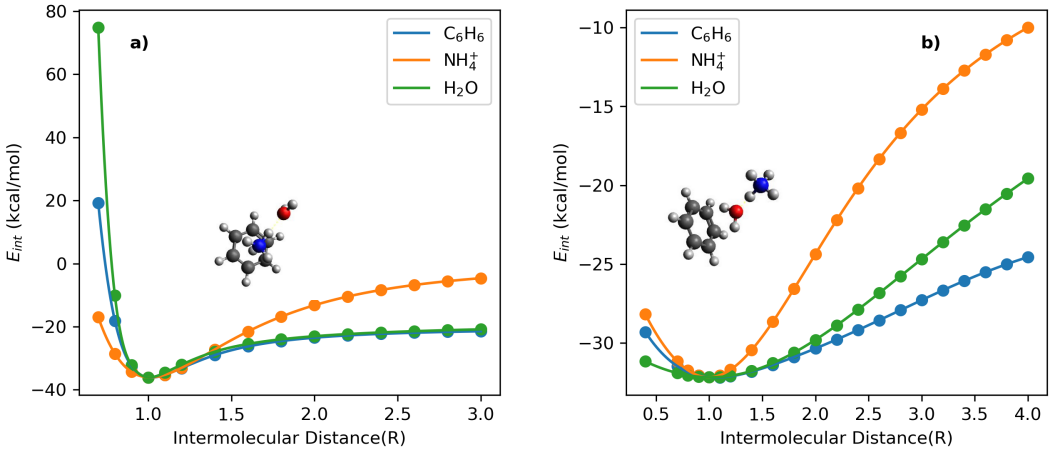


Figure 3.3: Radial benchmark potential energy curves (in kcal/mol) for the ammonium-benzene-water complex 4.1 **(a)** and 4.2 **(b)**. The curves are obtained by shifting C_6H_6, NH_4^+, and H_2O relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

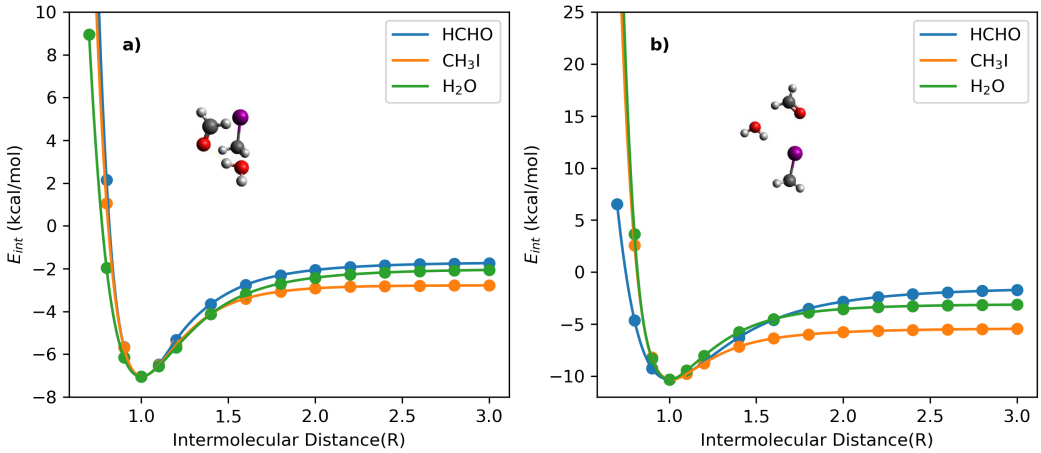


Figure 3.4: Radial benchmark potential energy curves (in kcal/mol) for the iodomethane-formaldehyde-water complex 1.1 **(a)** and 1.2 **(b)**. The curves are obtained by shifting HCHO, H_2O, and CH_3I relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

because the pairwise interaction between two monomers remains as only the third monomer is shifted away. At short-range, as expected, all interaction energies are seen to be positive due to the orbital overlap among the monomers and the resulting exchange repulsion.

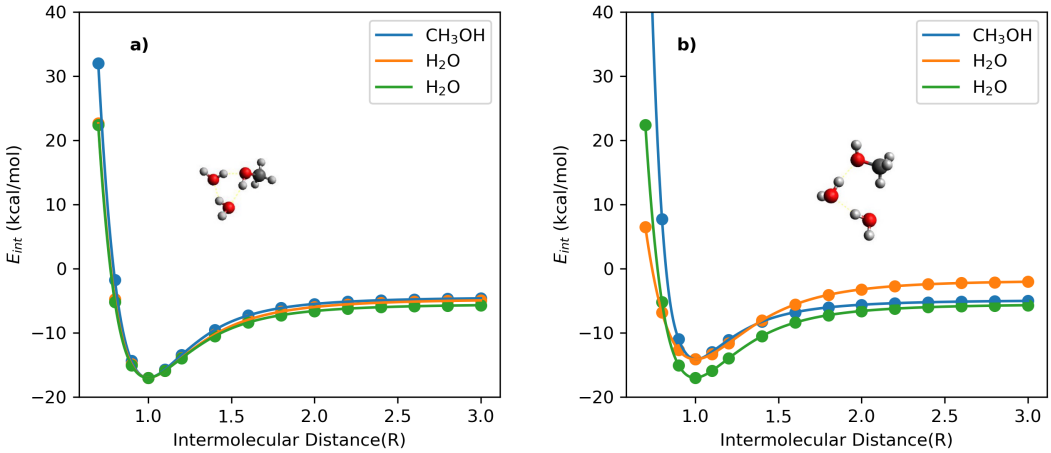


Figure 3.5: Radial benchmark potential energy curves (in kcal/mol) for the water-water-methanol complex 3.1 **(a)** and 3.2 **(b)**. The curves are obtained by shifting CH_3OH, H_2O, and H_2O relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

Figure 3.3 shows the 1D potential energy curves for the ammonium-benzene-water complex. The strongest repulsive interaction occurs when the water molecule is brought closer from the equilibrium distance. When the NH_4^+ ion is moved away from the benzene-water dimer, the interaction energy decays slowly, as $1/R^2$, as it is dominated by the charge-dipole interaction. The closeness of two curves in Fig. 3.3a is accidental and we verified that the results diverge more at larger distances: after all, the decaying term in the interaction is dominated by a charge-dipole contribution when the water molecule is moving away, and a charge-quadrupole one when the benzene molecule is shifted away.

At equilibrium, the most dominant two-body interaction in the iodomethane-formaldehyde-water complex (Fig. 3.4) is electrostatics, followed by dispersion and induction. This ordering holds for both the halogen-bonded structure 1.2 and the non-halogen-bonded configuration 1.1. The largest contributions to both electrostatics and dispersion come from the hydrogen and halogen bonds.

The water-water-methanol complex (Fig. 3.5) resembles the well known water trimer. Each monomer simultaneously acts as a H-bond donor and an H-bond acceptor in the cyclic trimer minimum geometry, and the fact that the non-hydrogen-bonded H atom of one of the water molecules is

replaced by a methyl group brings a minimal perturbation to the bonding pattern of a water trimer. The interactions at long range are dominated by the dipole-dipole term, decaying as $1/R^3$, and this decay is very similar when each monomer, water or methanol, is shifted away from the other two, as indicated by the strong similarity of the three curves in Figure 3.5a. In the second, 3.2 structure of this complex (Fig. 3.5 b), one of the hydrogen bonds is replaced by a weaker C-H$\cdots$O contact; as a result, the long-range interaction energy is more sensitive to the displacement of the water molecule taking part in both hydrogen bonds.

3.2.3. SAPT analysis of two-body interaction energy

The two-body interaction energy components for the twenty 3BHET structures are presented in Table 3.3 for SAPT(DFT), SAPT2+(3), and XSAPT. The physical meaning of these two-body components can be interpreted precisely. In the 3BHET dataset, the electrostatic interactions are all attractive because we picked favorable global- and local-minima geometries. First-order exchange interactions in dimers are always repulsive while the net induction and dispersion interactions are attractive (that is, the positive exchange-induction and exchange-dispersion terms attenuate, but not cancel, the negative induction and dispersion contributions). The magnitudes of individual terms vary depending on the geometry and complex size. The aromatic complexes are largely dispersion-bound while the complexes involving ions and the systems with strong hydrogen bonding are primarily electrostatics-bound. In addition, the systems 4, 5, and 8 exhibit large induction interaction energies due to the presence of charged species. The induction energy computed with all SAPT variants does include an appropriate δ^{HF} term, which accounts for higher-order induction and exchange induction interactions along with some small contributions from first-order effects.[111]

As far as the three SAPT flavors in Table 3.3 are concerned, the best agreement between the first-order components occurs between SAPT2+(3) and XSAPT (with mean absolute deviations (MAD) of 0.3 kcal/mol for electrostatics and 0.7 kcal/mol for exchange), but the agreement with SAPT(DFT) is not much worse (with MAD values up to 0.5 kcal/mol for electrostatics and 1.6 kcal/mol for exchange). As far as the three SAPT flavors in Table 3.3 are concerned, the

SAPT(DFT) and SAPT2+(3) first-order components are much closer to each other (with mean absolute deviations (MAD) of 0.4 kcal/mol for electrostatics and 1.7 kcal/mol for exchange) than either method is to XSAPT (the corresponding MAD are 2.5–2.7 kcal/mol for electrostatics and 3.1–3.6 kcal/mol for exchange). This observation likely means that the XSAPT first-order terms, computed at the uncoupled SAPT(KS) level, are not very accurate.The second-order contributions deviate between the methods to a similar extent, with MAD values in the range 0.8–1.3 kcal/mol for induction (the largest discrepancy occurs between XSAPT and SAPT(DFT)) and 1.3–2.7 kcal/mol for dispersion (with the largest discrepancy between XSAPT and SAPT2+(3)).

Table 3.3: Two-body SAPT(DFT), SAPT2+(3), and XSAPT interaction energy contributions (kcal/mol). The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Complex	2-body SAPT(DFT)				SAPT2+(3)				XSAPT			
	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$
1.1	−14.245	17.860	−4.143	−7.776	−14.223	18.468	−5.024	−8.823	−14.529	18.474	−4.803	−6.569
1.2	−9.408	11.778	−2.716	−5.679	−9.235	12.630	−3.361	−6.508	−9.671	13.651	−3.160	−5.130
2.1	−11.377	13.732	−3.086	−6.017	−11.572	13.865	−3.663	−6.811	−11.535	14.128	−3.325	−4.651
2.2	−11.708	13.867	−3.280	−5.559	−11.859	13.918	−3.879	−6.267	−11.907	14.323	−3.522	−4.167
3.1	−27.738	32.119	−7.842	−8.523	−27.965	32.758	−9.021	−9.672	−28.002	32.224	−8.817	−6.211
3.2	−22.049	25.231	−6.663	−6.996	−22.260	25.597	−7.632	−7.885	−22.311	25.509	−7.572	−4.904
4.1	−33.865	27.709	−21.270	−9.076	−34.310	28.422	−23.591	−9.915	−33.631	27.782	−26.704	−6.229
4.2	−35.256	36.690	−17.519	−9.568	−35.631	37.679	−19.441	−10.706	−35.237	37.004	−21.034	−6.723
5.1	−17.677	16.550	−11.656	−8.850	−18.967	21.906	−12.325	−10.125	−19.233	21.964	−11.350	−9.052
5.2	−9.315	11.925	−8.747	−8.278	−10.812	16.291	−9.424	−9.501	−11.158	16.295	−8.317	−7.821
6.1	−13.641	31.477	−3.521	−19.139	−13.752	32.472	−4.023	−21.895	−13.109	30.389	−3.320	−18.343
6.2	−11.752	27.635	−3.000	−17.704	−11.306	27.295	−3.367	−19.991	−11.161	26.238	−2.899	−16.683
7.1	−20.326	28.783	−6.928	−15.978	−21.207	33.784	−7.688	−17.921	−21.235	32.884	−7.632	−14.377
7.2	−8.657	18.712	−2.946	−14.438	−9.757	23.075	−3.143	−16.403	−9.350	21.333	−2.860	−13.356
8.1	−25.819	32.586	−24.645	−15.779	−26.335	34.119	−27.461	−17.623	−25.655	33.056	−29.759	−13.378
8.2	−23.020	27.384	−24.237	−13.816	−22.964	29.040	−26.997	−14.957	−22.587	27.557	−30.291	−10.975
9.1	−10.164	14.012	−3.090	−7.854	−10.158	14.860	−3.539	−8.645	−10.311	14.964	−3.464	−6.699
9.2	−9.059	12.928	−3.017	−8.175	−8.799	13.925	−3.308	−8.918	−8.826	13.976	−3.258	−7.489
10.1	−9.404	18.945	−2.956	−13.427	−9.150	20.161	−3.301	−14.804	−9.003	19.374	−3.226	−13.346
10.2	−8.083	19.684	−2.726	−13.799	−8.198	21.449	−2.983	−15.310	−7.896	19.989	−2.890	−13.488

The total two-body interaction energies from all three SAPT variants, shown in Table 3.4, exhibit varying magnitudes across the methods. The agreement between the total two-body energies is perhaps lower than expected, with an average absolute deviation between two "higher-level" methods, SAPT(DFT) and SAPT2+(3), amounting to 1.5 kcal/mol. The XSAPT two-body energies deviate on the average by 1.7 and 2.0 kcal/mol from SAPT(DFT) and SAPT2+(3), respectively.

3.2.4. SAPT analysis of three-body nonadditive interaction energy

For the decomposition of the three-body nonadditive interaction energy, we can no longer use the (purely two-body) SAPT2+(3) variant, so we now focus on SAPT(DFT) and XSAPT.

Table 3.4: Total two-body and three-body SAPT interaction energies (kcal/mol). The benchmark CCSD(T)/CBS values from Table 3.1 are copied here to facilitate comparisons.

Complex	SAPT(DFT)			SAPT2+(3)		XSAPT		CCSD(T)/CBS		
	E^2_{int}	E^3_{int}	E_{int}	E^2_{int}	E^2_{int}	E^3_{int}	E_{int}	E^2_{int}	E^3_{int}	E_{int}
1.1	−8.303	−0.582	−8.885	−9.603	−7.428	−0.505	−7.933	−9.772	−0.546	−10.317
1.2	−6.025	−0.538	−6.563	−6.474	−4.311	−0.396	−4.707	−6.466	−0.574	−7.041
2.1	−6.747	−0.408	−7.155	−8.182	−5.383	−0.381	−5.764	−8.293	−0.405	−8.697
2.2	−6.680	−0.533	−7.214	−8.086	−5.274	−0.473	−5.747	−8.129	−0.559	−8.688
3.1	−11.984	−2.547	−14.531	−13.899	−10.805	−2.264	−13.069	−14.573	−2.421	−16.994
3.2	−10.477	−1.708	−12.185	−12.180	−9.278	−1.649	−10.927	−12.523	−1.593	−14.115
4.1	−36.502	3.386	−33.116	−39.394	−38.782	2.682	−36.100	−39.618	3.386	−36.232
4.2	−25.653	−4.524	−30.176	−28.098	−25.989	−4.827	−30.815	−27.766	−4.623	−32.389
5.1	−21.632	1.694	−19.939	−19.511	−17.670	1.026	−16.645	−18.888	1.338	−17.550
5.2	−14.415	1.167	−13.248	−13.445	−11.001	0.795	−10.206	−12.974	1.171	−11.803
6.1	−4.824	0.436	−4.387	−7.198	−4.383	0.484	−3.900	−7.206	0.347	−6.859
6.2	−4.821	0.052	−4.768	−7.369	−4.506	−0.254	−4.760	−7.211	−0.116	−7.327
7.1	−14.448	−1.350	−15.798	−13.031	−10.360	−1.695	−12.055	−12.968	−1.559	−14.527
7.2	−7.329	0.004	−7.324	−6.228	−4.234	−0.033	−4.267	−6.240	−0.009	−6.249
8.1	−33.657	3.862	−29.795	−37.301	−35.737	3.234	−32.503	−37.576	3.789	−33.788
8.2	−33.689	4.254	−29.435	−35.878	−36.295	3.293	−33.001	−37.114	4.152	−32.962
9.1	−7.095	−0.618	−7.714	−7.482	−5.510	−0.637	−6.146	−7.261	−0.669	−7.930
9.2	−7.323	−0.491	−7.814	−7.099	−5.597	−0.536	−6.133	−6.883	−0.510	−7.392
10.1	−6.841	−0.133	−6.974	−7.093	−6.201	−0.193	−6.395	−6.746	−0.200	−6.947
10.2	−4.924	0.164	−4.761	−5.043	−4.285	0.009	−4.277	−4.975	0.056	−4.919

The individual three-body components for both methods are given in Table 24. Note that while the former method includes the leading-order nonadditivity in all primary components except for the (entirely pairwise additive) electrostatics, the XSAPT approach neglects the nonadditivity of first-order exchange and includes the nonadditive induction (with or without three-body induction couplings) and dispersion (the part of the MBD value for the trimer that cannot be explained by pairwise MBD terms). For the 20 structures in 3BHET, the three-body exchange effects are always attractive, slightly lowering the two-body exchange repulsion. The nonadditive dispersion term is repulsive (again compensating a small part of the two-body dispersion attraction) while the largest one of the three, the nonadditive induction, can be both attractive and repulsive, and sometimes changes sign between two configurations of the same complex. While the nonadditive induction term ranges from −3.7 to 4.1 kcal/mol for the trimers in our dataset, the magnitude of the three-body exchange and dispersion does not exceed 0.9 and 0.4 kcal/mol, respectively.

Figure 3.6 displays a more extensive comparison of the total nonadditive induction energy, including all available minor variants of XSAPT and SAPT(DFT) as well as SAPT0. For SAPT(DFT), one can include or omit the nonadditive $\delta E^{HF}_{int,3}$ term, Eq. (3.12). In the case of XSAPT, the HF delta term is pairwise additive and does not influence the three-body energy, but one can include

Table 3.5: Three-body SAPT(DFT) and XSAPT interaction energy contributions (kcal/mol). The capitalized component names indicate the inclusion of the corresponding nonadditive exchange effects: specifically, $E^{(2)}_{Ind,3} = E^{(2)}_{ind,3} + E^{(2)}_{exch-ind,3}$ and $E^{(3)}_{Disp,3} = E^{(3)}_{disp,3} + E^{(2)}_{exch-disp,3}$. Three-body XSAPT induction couplings are not included.

Complex	SAPT(DFT)			XSAPT	
	$E^{(1)}_{exch,3}$	$E^{(2)}_{Ind,3}$	$E^{(3)}_{Disp,3}$	$E^{(2)}_{Ind,3}$	$E^{MBD}_{Disp,3}$
1.1	−0.172	−0.562	0.151	−0.545	0.040
1.2	−0.055	−0.566	0.083	−0.420	0.024
2.1	−0.129	−0.398	0.118	−0.409	0.029
2.2	−0.051	−0.557	0.075	−0.486	0.013
3.1	−0.573	−2.208	0.233	−2.315	0.051
3.2	−0.253	−1.553	0.098	−1.657	0.007
4.1	−0.205	3.457	0.134	2.671	0.010
4.2	−0.854	−3.697	0.027	−4.779	−0.048
5.1	−0.560	1.870	0.385	0.929	0.097
5.2	−0.434	1.362	0.239	0.760	0.035
6.1	−0.112	0.293	0.255	0.340	0.144
6.2	−0.306	0.041	0.317	−0.346	0.091
7.1	−0.501	−1.151	0.302	−1.739	0.045
7.2	−0.105	0.100	0.009	−0.026	−0.007
8.1	−0.034	3.681	0.215	3.185	0.049
8.2	−0.114	4.112	0.256	3.237	0.056
9.1	−0.326	−0.455	0.163	−0.670	0.033
9.2	−0.152	−0.485	0.147	−0.569	0.033
10.1	−0.328	−0.102	0.298	−0.283	0.090
10.2	−0.119	0.051	0.231	−0.071	0.079

or omit the three-body induction couplings.[109] Figure 3.6 shows that the SAPT0 approach and the SAPT(DFT) method without $\delta E^{HF}_{int,3}$ give results quite similar to each other, and smaller in magnitude than the other variants. The addition of the $\delta E^{HF}_{int,3}$ correction makes the SAPT(DFT) nonadditive induction term larger in magnitude and likely more accurate. Notably, the SAPT(DFT) values with $\delta E^{HF}_{int,3}$ are quite similar to the XSAPT data without three-body induction couplings, while the inclusion of the latter increases the magnitude of nonadditive induction even further. It should be stressed that while the description of nonadditive induction by SAPT(DFT) and XSAPT may appear similar, the total nonadditive interaction energies from SAPT(DFT) are much closer to the benchmark, cf. Table 3.2; the likely reason is the neglect of the nonadditive first-order exchange within XSAPT. We conclude that the iterative embedding scheme incorporated by XSAPT leads to a good description of three-body induction effects even without including the explicit three-body induction couplings; actually, the addition of the latter does not seem to improve the overall accuracy any further.

A scatter plot of the total nonadditive interaction energy against its induction component, displayed in Fig. 3.7 , confirms that the latter accounts for the greatest percentage of the nonadditive

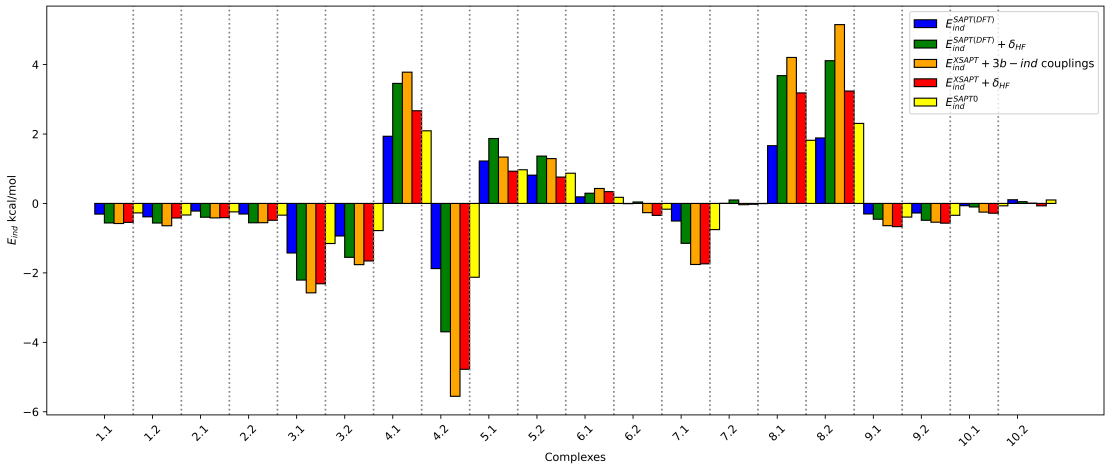
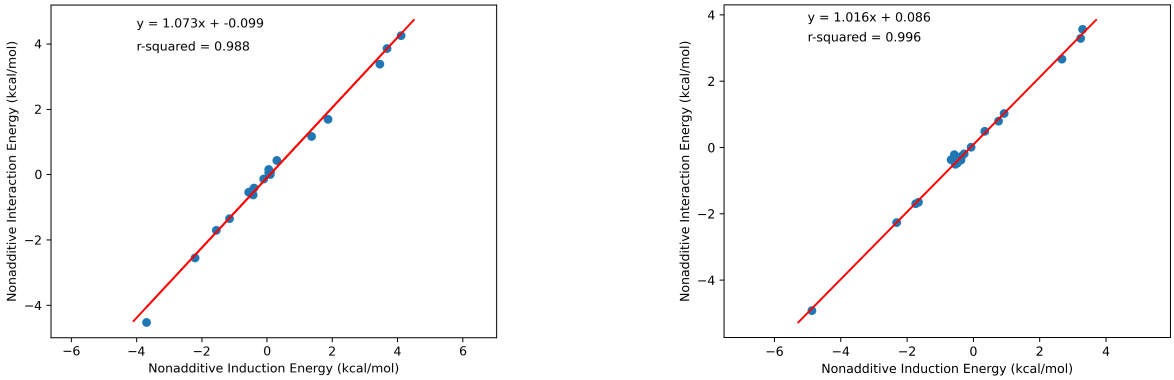


Figure 3.6: Comparison of the nonadditive induction energies for the 3BHET dataset predicted by SAPT0, XSAPT, and SAPT(DFT).

three-body effect. Thus, accounting reliably for the nonadditive induction effects is the key to providing an accurate description of the entire E_{int}^{3} for the systems in our dataset, and those effects can be recovered, to a reasonable extent, already with a highly approximate method like HF. Moreover, Fig. 3.7 indicates that the total three-body effect is nearly proportional to its induction component, suggesting that a suitable scaling of the latter term might be a good way to approximate the entire nonadditive energy.



SAPT(DFT) nonadditive induction **XSAPT nonadditive induction**

Figure 3.7: Total SAPT(DFT) and XSAPT nonadditive interaction energy plotted versus its respective induction contribution.

The different nonadditive dispersion estimates for the systems in our dataset are compared in Fig. 3.8. The calculations presented in this work provide three meaningful estimates of three-body dispersion: the sum $E^{(3)}_{disp,3}(\mathrm{CKS}) + \widetilde{E}^{(2)}_{exch-disp,3}(\mathrm{CKS})$ computed in three-body SAPT(DFT), the difference between the MBD dispersion energies for the trimer and dimers computed as part of XSAPT, and the difference between the supermolecular CCSD(T)/CBS and MP2/CBS nonadditive interaction energies. The latter difference is obviously not just due to three-body dispersion (for example, the improved multipole moments and polarizabilities in CCSD(T) will result in an altered value for the nonadditive induction), but since MP2 misses the pure three-body dispersion completely and CCSD(T) includes it, the difference has been used to estimate the nonadditive dispersion effects in the past.[85] Our results in Fig. 3.8 show that the nonadditive dispersion energy term, however small, is very hard to obtain accurately, as its three estimates computed in this work disagree with each other quite substantially. Most notably, the SAPT(DFT) estimate is consistently much larger than the other two, indicating that the CKS approach, very successful for computing two-body dispersion accurately and efficiently, might not be as reliable for three-body dispersion. This overestimation of nonadditive dispersion by SAPT(DFT) has been observed before[112] and might possibly be attributed to the lack of so-called "dispersion size-consistency" for many-molecule SAPT.[113] The fact that the SAPT(DFT) estimate requires scaling the exchange-dispersion term by the CKS/KS ratio[58] might lead to additional inaccuracies. The CCSD(T)$-$MP2 and XSAPT(MBD) estimates are more consistent with each other, but sometimes (e.g. for systems 4.2 and 7.1) also exhibit large relative differences. Note that the nonadditive (in the molecular sense) MBD three-body dispersion is obtained using a supermolecular approach, as the difference between the trimer energy and the sum of the pair energies, which might amplify residual errors. Overall, even though none of the methods presented in Fig. 3.8 can be considered an unambiguous benchmark measure of pure nonadditive dispersion effects, the agreement between CCSD(T)$-$MP2 and XSAPT(MBD) for most (but not all) systems suggests that both approaches are quite trustworthy for reproducing this effect, which is small for systems in our dataset but important for many condensed-phase systems.

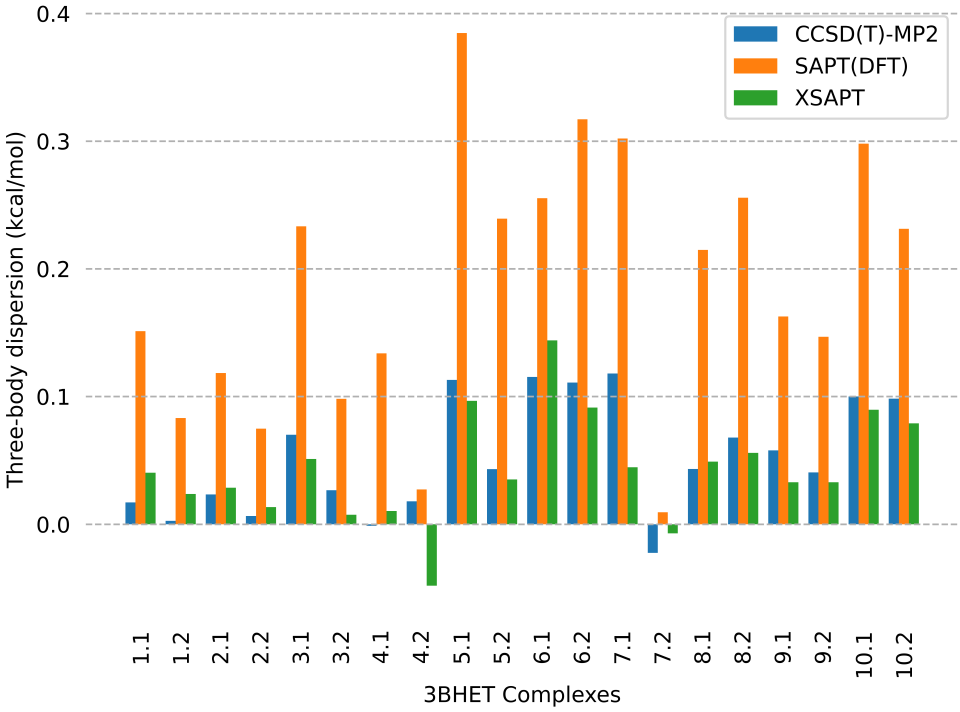


Figure 3.8: Three-body nonadditive dispersion energy for the systems in the 3BHET dataset, estimated as the supermolecular CCSD(T)/CBS−MP2/CBS difference as well as computed with SAPT(DFT) and XSAPT (the latter uses the MBD formalism for the dispersion part).

3.2.5. Total SAPT interaction energies

The total interaction energies from three SAPT variants (SAPT2+(3), SAPT(DFT), and XSAPT) for the trimers in our dataset are presented in Table 3.4. Overall, the discrepancies between variants can be quite substantial, up to over 8 kcal/mol (or roughly double the nonadditive three-body term) for the strongly bound ionic complex 8. When these total interaction energies are compared to the reference CCSD(T)-level data from Table 3.1, the two-body SAPT2+(3) method is observed to overestimate and underestimate subsets of data to a similar extent — the mean signed error (MSE) of SAPT2+(3) is only 0.01 kcal/mol. More specifically, SAPT2+(3) exhibits too strong binding (by $\geq$ 1.5 kcal/mol) for systems 4.1, 5.1, 5.2, 8.1, and 8.2, and too weak binding (also by $\geq$ 1.5 kcal/mol) for 3.1, 3.2, and 4.2. Notably, for each of the structures listed in the previous sentence, the nonadditive three-body term is sizeable and of opposite sign to the two-body error, so an inclusion of even an approximate three-body term would substantially improve

the accuracy of SAPT2+(3) total interaction energies. In contrast, the SAPT(DFT) and XSAPT variants underestimate the binding in the 3BHET complexes quite systematically, with an MSE of 1.05 kcal/mol for SAPT(DFT) and 1.87 kcal/mol for XSAPT. However, the main source of this underestimation is quite different in each case, with SAPT(DFT) leading to largest errors for systems involving the ammonium cation (4 and 8), while XSAPT gives the largest errors for the strongly hydrogen-bonded system 3, followed by halogen-containing (and halogen-bonded in some configurations) systems 2 and 6. On the positive side, the errors from SAPT(DFT), as well as those from XSAPT, are generally quite consistent between the two configurations of the same trimer, implying that both approaches might be better suited to examining relative energies between such configurations.

SAPT provides physical insights to examine the nature of the two- and three-body interaction energies, and a ternary diagram[82,90] provides a simple way to graphically assign a two-body interaction type. Such a SAPT(DFT) ternary diagram for the 20-point 3BHET dataset is presented in Figure 3.9 (blue dots), showing that the majority of the 3BHET trimers are mainly electrostatics and dispersion-bound. Using the SAPT(DFT) decomposition of the nonadditive three-body energy into induction, dispersion, and first-order exchange, we can enhance the ternary diagram by combining the respective SAPT contributions of a given type. The resulting relative importance of the electrostatic, induction, and dispersion two- plus three-body contributions is shown by red dots in Figure 3.9. We see that the blue and red dots clearly come in pairs and often overlap with each other, indicating that the three-body contribution provides only a minor modification of the two-body interaction type. However, the binding decomposition in Fig. 3.9 can be viewed as slightly incomplete, as the nonadditive first-order exchange is not displayed while it is attractive for the complexes studied here and provides an additional minor contribution to the binding. To account for all attractive two- and three-body interactions, Figure 3.10 contains a modified ternary diagram with the pure electrostatic data (which are pairwise additive) replaced by the entire first-order energies (electrostatics plus exchange, where the latter is not entirely additive). Compared to

Fig. 3.9 , the addition of exchange causes an obvious shift to more repulsive values, while the two-body (blue dots) and two- plus three-body (red dots data) still form closely lying pairs. Moreover, within each pair, the two points are displaced from each other in a direction roughly towards or away from the pure induction vertex of the diagram, confirming that the largest difference between the two-body and two- plus three-body binding character stems from nonadditive induction.

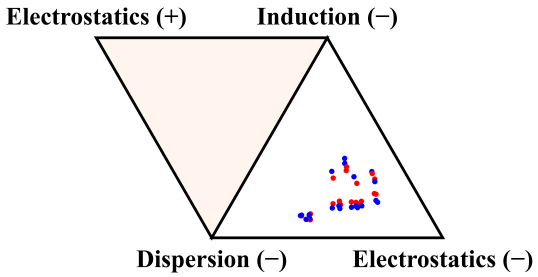


Figure 3.9: Two-body (blue dots) and two- plus three-body (red dots) SAPT(DFT) interaction energy distribution for trimers studied here, displayed on a ternary diagram.

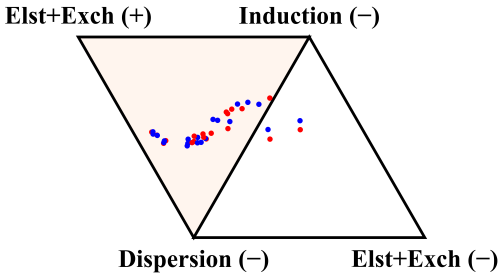


Figure 3.10: Two-body (blue dots) and two- plus three-body (red dots) SAPT(DFT) interaction energy distribution, including first-order exchange effects, for trimers studied here, displayed on a ternary diagram.

Finally, the information from two- and three-body SAPT decompositions allows us to make connections between the chemical interaction type and the interaction energy contributions. The complexes involving ions exhibit the strongest binding overall, dominated by the two-body electrostatic and induction components. The same complexes have the largest magnitude of the three-body energy thanks to a sizable nonadditive induction. The later term can be attractive or repulsive,

depending on whether the polarization of some monomer X by monomer Y makes the X–Z interaction more or less favorable. For example, the two structures of complex 4 (ammonium-benzene-water) exhibit opposite signs of nonadditive induction, which can be rationalized by looking at the most polarizable monomer (benzene) responding to the monomer producing the strongest electric field (the ammonium ion). When the π-electron density is pulled towards the positive ion, the change in the benzene-water interaction is unfavorable for 4.1 (where the negatively charged side of the water molecule points towards benzene) and favorable for 4.2 (where the positive side of water is closer to benzene), in accordance with the sign of the nonadditive induction effect. The next strongest binding occurs in neutral complexes that are strongly hydrogen-bonded, such as 3 (methanol-water-water) and 7 (phenol-benzene-water). The induction energy no longer plays the most dominant role, and the binding in complex 3 is primarily due to electrostatics. Since complex 7 also involves two aromatic rings, dispersion energy is just as important for its stability as electrostatic energy. Looking now at complexes 1, 2, and 6 that can form halogen bonds (with a halogen atom X connected to C), we picked one halogen-bonded geometry (with the C-X bond pointing toward an electronegative atom in another molecule) and one non-halogen-bonded geometry for each system. For our complexes, no clear signature of halogen bonding is observed in the two-body energy contributions. However, the halogen bonded structures 1.2, 2.2, and 6.1 exhibit a smaller magnitude of SAPT(DFT) nonadditive exchange and dispersion corrections than their non-halogen-bonded counterparts. This might be coincidental and requires calculations on a broader class of systems to confirm any general trend. System 6 and the remaining two complexes (9 and 10) are the least polar and do not have hydrogen bonds. These complexes exhibit the weakest binding in general, and this binding is driven by dispersion followed by electrostatics. At the three-body level, induction is no longer a clearly dominant term and all three SAPT(DFT) contributions are roughly of similar importance. While nonadditive dispersion never plays a very large role in our dataset, its effect is notable in complexes 6 and 10 which involve two aromatic monomers.

3.2.6. Conclusions

Effective treatment of noncovalent interactions in many-body systems is vital as computational studies are moving towards bigger clusters and condensed phase systems. Exploring the binding mechanisms in molecular aggregates, solids, and liquids enables one to design new materials with tailored properties. However, it is essential to gather data on noncovalent interactions in vast regions of chemical space, not limited to dimers and homotrimers present in organic crystals. The new heteromolecular trimer benchmark dataset 3BHET developed here aims to aid in parameterizing and refining approximate methods that could be universally applicable throughout the chemical space. This dataset presents a diverse mixture of small and medium sized heteromolecular trimers that features various interactions involving both neutral and charged molecules. The benchmark two-body and three-body interaction energies were calculated at the CCSD(T)/CBS level.

Symmetry-adapted perturbation theory aids in the understanding and interpretation of noncovalent interaction energies at both two-body and three-body levels. Here, an extensive set of SAPT variants was tested on our new dataset, including SAPT2+(3), SAPT(DFT), and XSAPT. Overall, at the two-body level, the complexes are bound together by a varying mixture of electrostatics and dispersion, with an sizable induction effect present in complexes involving ions. The nonadditive three-body term can be favorable or unfavorable and is dominated by three-body induction. The other nonadditive effects, attractive first-order exchange and repulsive dispersion, are small but not negligible. To further elucidate the binding patterns and the degree of diversity present in our dataset, we extend the commonly used ternary diagrams from interactions in dimers to the complete two- and three-body interactions in trimers, illustrating also the importance of different nonadditive effects including induction, dispersion, and (optionally) first-order exchange. For our dataset, the balance of different pairwise and nonadditive interactions present in each trimer, despite their large chemical variety, resulted in a ternary diagram that features relatively similar points. This shows that the diversity of a dataset should be assessed by several complementary measures. A detailed and nuanced analysis of interactions provided by SAPT and ternary diagrams, in conjunction with the chemical intuition and a suitable selection of minimum

and off-minimum separations and orientations, will be highly desired for modelling more diverse interactions in future benchmark studies. While the available mixed-trimer benchmark data are not yet sufficient for modelling a broad range of complexes, this study takes the first step forward by examining the relationship between the physical interaction types and quantitative two-body and three-body energy contributions on a reasonably broad range of small systems.

The different SAPT variants tested in this work exhibit quite different performance. The benchmark two-body interaction energies are very well recovered by SAPT2+(3) (with a MUE below 0.3 kcal/mol), however, no treatment of nonadditive three-body effects commensurate with SAPT2+(3) is available. The XSAPT approach is computationally efficient and does include nonadditive induction and dispersion effects (the latter through the MBD formalism), but its three-body accuracy is hampered by the lack of nonadditive first-order exchange effects. The three-body induction couplings may be included in XSAPT and are required for formal completeness of the method, however, the effect of these couplings on the accuracy of XSAPT interaction energies is minor and actually slightly detrimental. The SAPT(DFT) variant is the only one that provides estimates of all important two-body and three-body effects. For the latter, the dominant induction term agrees quite well between SAPT(DFT) and XSAPT as long as the HF delta term is included in the former. The nonadditive dispersion energy is consistently overestimated by SAPT(DFT), while the MBD three-body dispersion included in XSAPT is of the same order as the difference between nonadditive CCSD(T) and MP2 energies (this difference is primarily but not exclusively dispersion) but the agreement is far from perfect. Overall, the total nonadditive energies from SAPT(DFT) are significantly more accurate than those from XSAPT, but still less accurate than supermolecular MP2 at the CBS limit. When it comes to total interaction energies, SAPT2+(3) emerges as the most accurate variant for our dataset despite the complete lack of three-body terms. In contrast, the accuracy of SAPT(DFT) and XSAPT interaction energies is far from satisfactory due to their inferior performance on two-body terms. The simplest SAPT0 variant is also the least accurate one, and XSAPT provides a substantial improvement over SAPT0.

Chapter 4

Influence of Three-Body Effects on Halogen Bonding

Accurate computational data are vital in advancing chemical understanding in domains such as template-directed synthesis, binding mechanisms in enzyme-ligand binding, and in determining the structure and properties of materials. Its importance cannot be overemphasized, because many molecular and electronic properties responsible for a broad range of phenomena cannot be measured directly. Benchmarks, therefore, play a vital role in validation of the accuracy of approximate computational methods and the development of those that rely on empirical parameters. In the realm of noncovalent interactions, where experimental data does not typically offer a straightforward comparison, these datasets are vital for ensuring accuracy and facilitating method development. Many datasets addressing different kinds of noncovalent interactions have been introduced. Examples include S22,[12] S66,[13] X40,[14] S12L,[15] NENCI,[16] or more recently NCI atlas by Řezáč et al. with several datasets trying to map dimer interactions over the entire chemical space one interaction at a time.[17–21] Each dataset features a specific theme. For instance, X40, S22, S66 represent small to medium dimer systems with less than 40 atoms computed using a composite CCSD(T)/CBS scheme. The NENCI dataset by DiStasio and coworkers represents a large dataset ($\sim$8000) which features a set of non-equilibrium geometries starting from the S66 and S101[22] datasets.

In molecular clusters, liquids, and solids, interactions between pairs of molecules provide only an approximate description of the total binding energy. Nonadditive contributions from triplets (and sometimes even quadruplets) of molecules are essential for obtaining the correct energetic order of water hexamers[3] and energetically ranking different polymorphs of molecular crystals.[77] However, the existing benchmark datasets for molecular trimers (or larger clusters) are not nearly as broad and are dominated by data for water clusters[114] or for clusters involving an ion such as Cl^-

solvated by a few water molecules.[60] Only a handful of benchmark three-body datasets contains trimers of different chemical character. The 3B-69 dataset contains 69 AAA-type trimers from 23 different molecular crystals.[24] The S22(3) set[25] includes 24 AAA- and AAB-type trimers obtained by augmenting the complexes in the popular S22 database.[12] Finally, the 3BHET dataset from our earlier work[32] comprises 20 model AAB- and ABC-type trimers exhibiting diverse types of interactions. However, these are relatively small datasets built for specific purposes and only focus on fairly specific types of interactions. Therefore, the field would benefit from the development of benchmark nonadditive datasets that are larger, more general, and as accurate as possible.

In this work, we build on the SH250[20] dataset of halogen bonded complexes featuring Cl, Br, I to include trimers and study the effects of three-body interactions on the halogen bond (XB). In recent years, non-covalent interactions involving σ-hole interactions, including halogen and chalcogen bonding, have gained significant attention.[1,115,116] These interactions are characterized by the attractive forces between an electrophilic region on a halogen or chalcogen atom and a Lewis base. This is as a result of an anisotropic redistribution of the halogen electron density. The positive region that appears at the extension of a σ covalent bond is termed a "σ-hole".[117–119] Hydrogen and halogen bonding act as complementary design features for the design of functional molecules and materials, stabilization of drug-protein complexes, and self-assembly of supramolecular structures.[120–122] Consequently, the development of robust computational methods to describe halogen bonding is widely important. Ideal methods not only aid in accurately modeling and understanding these interactions, but also ensure reliability in predicting interaction energies. Additionally, when the accurate methods become prohibitively costly, it is common practice to test the interaction energies from approximate theoretical methods against accurate benchmark calculations.

Several data sets covering σ-hole interactions have been introduced previously but only focus on dimers. The extension of the SH250$\times$10 dataset to trimers presented here aims to unite the coverage of these interactions with a more comprehensive benchmark. Besides introducing the dataset, this work also focuses on its applications to benchmarking SAPT flavors. This study was motivated by our previous work on heteromolecular trimers which revealed that the XB complexes

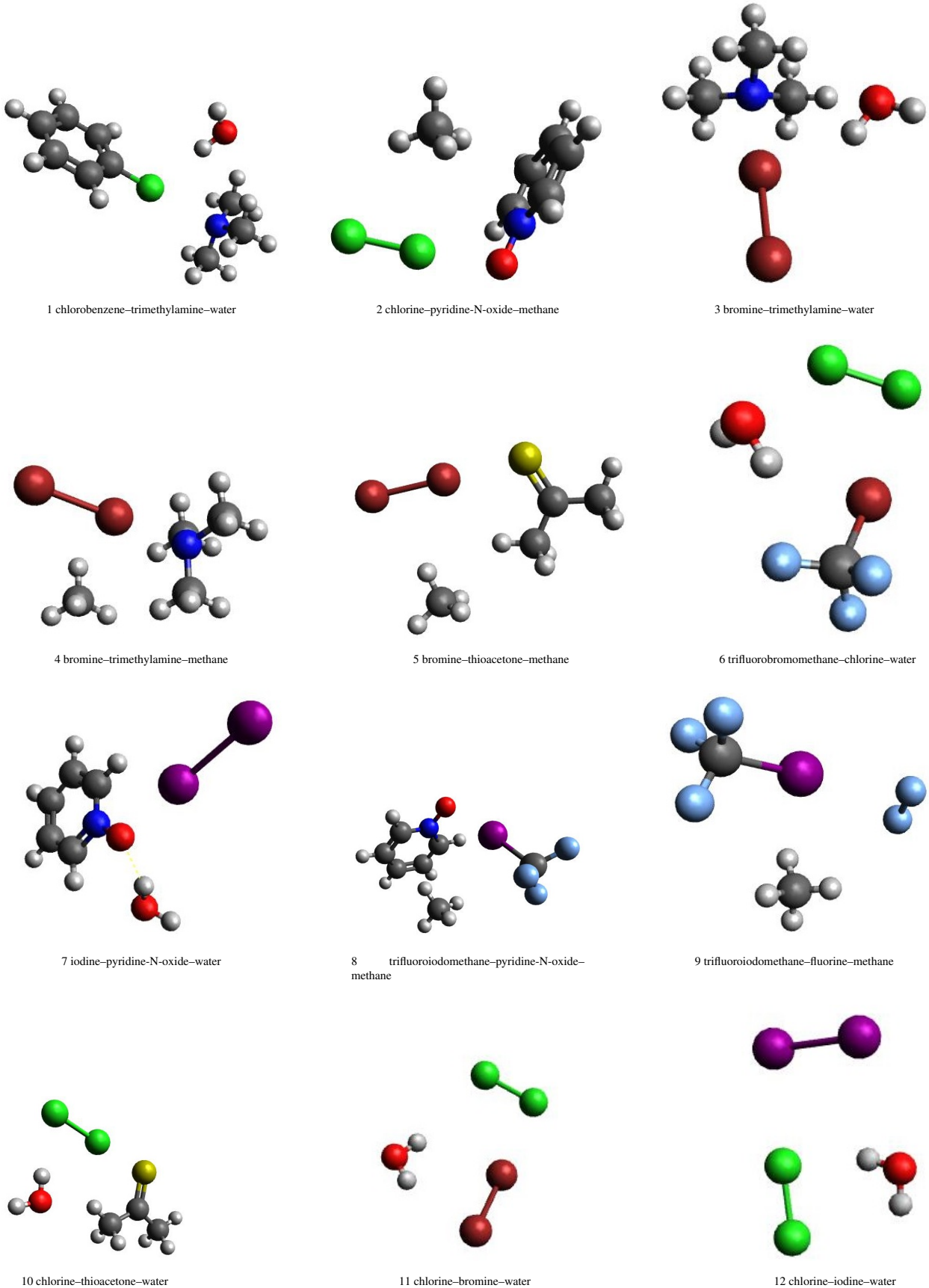
exhibited a comparatively smaller magnitude of the nonadditive exchange and dispersion contributions from SAPT(DFT) energy decomposition analysis.[32] With only three systems from a dataset with 20 trimers, it is too small to be useful in making predictions on general trends in specific interactions such as XB. Building on these findings, the focus on exclusively XB bonded trimers in this work seeks to attain a broader coverage with a more comprehensive benchmark of trimers in a larger chemical space than previously considered.

4.1. Computational Methods

4.1.1. Composition of the Data Set

The dataset introduced in this work, which will be denoted 3BXB, is an extension of the 3BHET dataset[32] with a focus on halogen bonding. The 3BXB data set is constructed on top of the SH250 database,[20] extending its coverage of halogen-bonded interactions to trimers. We examine the influence of the three-body non-additive interaction on the halogen bond, looking for common trends in the three-body contributions that might characterize halogen-bonded structures. 3BXB comprises 214 trimers (Fig. 4.1), 107 with water and 107 with methane, obtained by extending the halogen-bonded dimers featuring Cl, Br and I from the SH250 dataset.[20] We picked either water or methane as the third partner because the two molecules have the same number of electrons (10), but exhibit very different kinds of interactions, predominantly electrostatic/hydrogen bonded for water and mostly dispersive for methane. To ensure a seamless comparison with the previous 3BHET dataset,[32] the benchmark interactions are performed using the same composite CCSD(T)/CBS scheme. In this scheme, the reference interaction energy E_{int} and its non-additive contribution E^3_{int} are assembled from three components, the Hartree-Fock interaction energy, the MP2 correlation energy, and a small-basis CCSD(T) correction. The HF term converges faster with basis set, therefore a single calculation in a large enough basis set (in this case a5Z) is sufficient. The electron correlation energy from MP2 is extrapolated to the complete basis set limit from calculations in two basis sets (aQZ and a5Z). Lastly, the δCCSD(T) correction is computed in a smaller basis set, in this case aDZ.

4.1.2. Benchmark two-body and three-body interaction energies



1 chlorobenzene–trimethylamine–water

2 chlorine–pyridine-N-oxide–methane

3 bromine–trimethylamine–water

4 bromine–trimethylamine–methane

5 bromine–thioacetone–methane

6 trifluorobromomethane–chlorine–water

7 iodine–pyridine-N-oxide–water

8 trifluoroiodomethane–pyridine-N-oxide–methane

9 trifluoroiodomethane–fluorine–methane

10 chlorine–thioacetone–water

11 chlorine–bromine–water

12 chlorine–iodine–water

Figure 4.1: Representative trimer structures considered in the 3BXB dataset.

The interaction energies of interest are the two-body and nonadditive three-body terms, where the latter is defined as the difference between the trimer interaction energy and the sum of the pair interaction energies.[31] Specifically, the complete interaction energy in the trimer

$$E_{int} = E_{ABC} - E_A - E_B - E_C \tag{4.1}$$

can be decomposed as

$$E_{int} = E_{int}^2 + E_{int}^3 \tag{4.2}$$

where the two-body part is the sum of pairwise interactions:

$$E_{int}^2 = (E_{AB} - E_A - E_B) + (E_{BC} - E_B - E_C) + (E_{AC} - E_A - E_C) \tag{4.3}$$

The monomer and dimer calculations were performed in the full trimer basis set to eliminate basis set superposition error (BSSE). In addition, the interaction energies were computed using fixed monomer geometries ignoring all monomer deformation effects.[76] The benchmark interaction energies at both the two-body and three-body level were calculated using the composite CCSD(T)/CBS approach where correlation consistent basis sets fully augmented with diffuse functions were used; in this text, aug-cc-pVXZ (X = D, T, Q, 5, ...) are abbreviated as aXZ. The benchmark calculations employ the well-kown composite CCSD(T)/CBS scheme that has contributions of the Hartree-Fock method(HF), the MP2 correlation energy, and a small-basis CCSD(T) correction.

$$E^{(\mathrm{CCSD(T)/CBS})} = E^{\mathrm{HF}} + E_{\mathrm{corr}}^{\mathrm{MP2/CBS}} + \delta E^{\mathrm{CCSD(T)}} \tag{4.4}$$

In this formalism, a single HF calculation in a large basis set, a5Z basis for this purpose, is sufficient as it converges fast with the basis set size. The MP2 correlation energy is extrapolated to the CBS limit from the aQZ and a5Z basis sets using Helgaker's X^{-3} formula[98] and the coupled cluster correction $\delta E^{CCSD(T)}$ is calculated in the aDZ basis set. Thus, the benchmark interaction energies produced in this work correspond to the so-called silver standard of theory.[96,97]

4.1.3. SAPT interaction energy decomposition

For appropriate interpretation of the supermolecular data obtained from the benchmark calculations, we need to know the nature of both the two and three body binding energies in each trimer. This is efficiently achieved through energy decomposition analysis. In this study, we employ several Symmetry Adapted Perturbation Theory (SAPT) formalisms[8,30,31] to decompose the interaction energy into physically meaning components. While we utilize various flavors of SAPT, the formalisms described below follow a set of general rules that apply across these methods. These rules provide a consistent framework for interpreting the interaction energies and their components, regardless of the specific SAPT variant employed.

In the two-body and three-body SAPT based on the HF wavefunctions, the Hamiltonian is partitioned into the Fock operator F, the intermolecular interaction operator V, and the intramolecular correlation operator W encompassing all monomers:

$$H = F + V + W \tag{4.5}$$

where

$$F = F_A + F_B + F_C \qquad V = V_{AB} + V_{BC} + V_{AC} \qquad W = W_A + W_B + W_C \tag{4.6}$$

The simplest approximation that provides qualitatively reasonable total interaction energies is the SAPT0 one, which totally neglects the intramonomer electron correlation.[31] The two-body SAPT0 interaction energy is calculated as

$$\begin{aligned} E_{int}^{\mathrm{SAPT0}} =& E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} \\ &+ \delta E_{HF}^{(2)} \end{aligned} \tag{4.7}$$

In a SAPT correction $E^{(kl)}$ in Eq. (4.7) and throughout the text, the subscripts k and l denote orders of perturbation theory with respect to V and W, respectively. SAPT0 provides the physical components that represent the electrostatic, first-order exchange, induction, exchange-induction

(both with the inclusion of monomer relaxation/response effects as denoted by the additional subscript "resp"), dispersion, exchange-dispersion, and $\delta E_{HF}^{(2)}$ contributions. The $\delta E_{HF}^{(2)}$ term is meant to estimate the higher-order induction and exchange-induction effects from a supermolecular HF calculation.[8,31] This energy decomposition allows one to classify complexes according to the relative importance of electrostatics, induction, and dispersion for their binding, for example, using a ternary diagram.[82]

To attain quantitative accuracy of total energies and their contributions, a higher-order SAPT level that includes both intra- and intermonomer electron correlation effects is required. In this study, the two-body interaction energy decomposition was performed by wavefunction-based SAPT up to the SAPT2+3 level of theory , simultaneously including the effects of the W operator on most SAPT0-level corrections and some terms that are third-order in V. Beyond SAPT0, the following levels of SAPT have been established:[8]

$$\begin{aligned} E_{int}^{\text{SAPT2}} =& E_{elst}^{(10)} + E_{elst,resp}^{(12)} \\ &+ E_{exch}^{(10)} + E_{exch}^{(11)} + E_{exch}^{(12)} \\ &+ E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{ind}^{(22)} + E_{exch-ind}^{(22)} + \delta E_{HF}^{(2)} \\ &+ E_{disp}^{(20)} + E_{exch-disp}^{(20)} \end{aligned} \tag{4.8}$$

$$E_{int}^{\text{SAPT2+}} = E_{int}^{\text{SAPT2}} + E_{disp}^{(21)} + E_{disp}^{(22)} \tag{4.9}$$

$$E_{int}^{\text{SAPT2+(3)}} = E_{int}^{\text{SAPT2+}} + E_{elst,resp}^{(13)} + E_{disp}^{(30)} \tag{4.10}$$

$$E_{int}^{\mathrm{SAPT2+3}} = E_{int}^{\mathrm{SAPT2+(3)}} + [\delta E_{HF}^{(3)} - \delta E_{HF}^{(2)} + E_{ind,resp}^{30} + E_{exch-ind,resp}^{30}]_{ind} + [E_{exch-disp}^{(30)} + E_{ind-disp}^{(30)} + E_{exch-ind-disp}^{(30)}]_{disp} \quad (4.11)$$

Equations (4.8)–(4.11) show the progressive improvements to the SAPT0 approximation. In addition, for levels of SAPT beyond SAPT2, we may include additional corrections, denoted δE_{int}^{MP2} and (CCD). The $\delta E_{int}^{MP2} = E_{int}^{MP2} - E_{int}^{SAPT2}$ term is based on a supermolecular MP2 interaction energy and may account for higher-order couplings between induction and dispersion as well as for some exchange effects beyond the commonly used single exchange approximation.[8] The (CCD) notation in turn signifies that the second-order dispersion energy has been computed nonperturbatively, using the CCD+ST(CCD) dispersion introduced by Williams et al.[123] in place of the perturbative sum $E_{disp}^{(20)} + E_{disp}^{(21)} + E_{disp}^{(22)}$.

Alternative variants of SAPT, SAPT(DFT) and XSAPT, were also used to perform energy decomposition calculations on both two and three-body interaction energies. This follows the same workflow as highlighted in our previous work for both SAPT(DFT) and XSAPT.[32] For SAPT(DFT) based on Kohn Sham description of the monomers, the Fock operator is replaced by the Kohn Sham operator such that $K = K_A + K_B$ or $K = K_A + K_B + K_C$ for dimers and trimers respectively.

$$\begin{aligned} E_{int}^{SAPT(DFT)} &= E_{elst}^{(1)}(KS) + E_{exch}^{(1)}(KS) + E_{ind}^{(2)}(CKS) + E_{exch-ind}^{(2)}(CKS) \\ &\quad + E_{disp}^{(2)}(CKS) + E_{exch-disp}^{(2)}(CKS) + \delta E_{int,2}^{HF} \end{aligned} \quad (4.12)$$

the $\delta E_{int,2}^{\mathrm{HF}}$ takes into account higher-order induction effects that are computed in the supermolecular Hartree-Fock computed in dimer centered basis sets(DCBS)

$$\delta E_{int,2}^{\mathrm{HF}} = E_{int,2}^{\mathrm{HF}} - E_{elst}^{(10)} - E_{exch}^{(10)} - E_{ind,resp}^{(20)} - E_{exch-ind,resp}^{(20)} \quad (4.13)$$

Since DFT fails to account for exchange-correlation potential with correct asymptotic behavior, the thoroughly benchmarked PBE0 functional with an asymptotically corrected exchange-correlation potential is employed.[55] In addition, the induction and dispersion components are computed with orbital relaxation through coupled frequency-dependent density susceptibility (FDDS).[32,124] The exchange counterparts to dispersion and induction are evaluated within the time-dependent HartreeâĂŞFock (TDHF) or TDDFT and coupled perturbed Hartree-Fock or Kohn-Sham (CPHF/CPKS) respectively.[124] These are often calculated using the S^2 approximation which ensures the exchange effects are captured via overlap-based terms.

The extended variant of SAPT (XSAPT), as the name suggests, allows for the extension of SAPT to an arbitrary number of monomers. This is through the explicit polarization formalism, XPol,[65] that works by partitioning the system to a number of fragments which are completely independent of each other. When combining SAPT with XPol, we obtain the the description to the total interaction energy in XSAPT as:

$$E_{int}^{\mathrm{XSAPT}} = E_{elst}^{(1)} + E_{exch}^{(1)} + E_{Disp}^{(2)} + [E_{ind}^{(2)} + E_{exch-ind}^{(2)} + \sum_{A}\sum_{B>A} \delta E_{AB}^{\mathrm{HF}} + \sum_{A}\sum_{B>A} (E_{AB}^{\mathrm{XSAPT}} - E_{AB}^{\mathrm{SAPT(KS)}}) + E_{int}^{\mathrm{MB}}] \tag{4.14}$$

and

$$\delta E_{AB}^{\mathrm{HF}} = E_{int}^{HF} - (E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)}) \tag{4.15}$$

The $\delta_{AB}^{\mathrm{HF}}$ term takes care of higher-order induction and, for systems with multiple bodies, the correction is obtained from the summation of pairwise contributions.

4.1.4. Three-body SAPT energy decomposition and its alternatives

While the previous section introduced the wavefunction SAPT, here we focus on the variant based on the density functional theory(DFT) description of monomers denoted as SAPT(DFT). It follows the same formalism of decomposing the interaction into electrostatics, exchange, induction, and dispersion and in its conventional implementation for a dimer interactions, the energy is

given by:

$$\begin{aligned} E_{int}^{SAPT(DFT)} &= E_{elst}^{(1)}(KS) + E_{exch}^{(1)}(KS) + E_{ind}^{(2)}(CKS) + \widetilde{E}_{exch-ind}^{(2)}(CKS) \\ &\quad + E_{disp}^{(2)}(CKS) + \widetilde{E}_{exch-disp}^{(2)}(CKS) + \delta E_{int,2}^{HF} \end{aligned} \tag{4.16}$$

where electrostatics and dispersion cover long-range interactions while exchange and induction account for short-range repulsion and polarization effects. Here, $E_{elst}^{(1)}$ and $E_{disp}^{(2)}$ are rigorously pairwise additive while $E_{disp}^{(3)}$ is a purely nonadditive many-body effect. In the three-body SAPT(DFT) framework, the nonadditive trimer interaction energy is given by:[58]

$$\begin{aligned} E_{int}^{SAPT(DFT)} &= E_{exch,3}^{(1)}(KS) + E_{ind,3}^{(2)}(CKS) + \widetilde{E}_{exch-ind,3}^{(2)}(CKS) \\ &\quad + \widetilde{E}_{exch-disp,3}^{(2)}(CKS) + E_{disp,3}^{(3)}(CKS) + \delta E_{int,3}^{HF} \end{aligned} \tag{4.17}$$

Since electrostatics can be approximated by sum over pairs, it is not included in three-body SAPT(DFT). On the contrary, exchange, induction and third-order dispersion introduce nonadditive terms that need explicit treatment in three-body SAPT(DFT). Exchange is computed in the Kohn-Sham formalism while induction and dispersion are computed from monomer frequency-dependent density susceptibilities (FDDS) with their exchange counterparts by scaling the uncoupled forms.[58]

$$\widetilde{E}_{exch-ind,3}^{(2)}(CKS) = E_{exch-ind,3}^{(2)}(KS)\frac{E_{ind,3}^{(2)}(CKS)}{E_{ind,3}^{(2)}(KS)} \tag{4.18}$$

$$\widetilde{E}_{exch-disp,3}^{(2)}(CKS) = E_{exch-disp,3}^{(2)}(KS)\frac{E_{disp,3}^{(3)}(CKS)}{E_{disp,3}^{(3)}(KS)} \tag{4.19}$$

Among the nonadditive components, the three-body dispersion is the most computationally demanding. Its electronic many-body quantum nature inherently makes it difficult to calculate accurately or generate appropriate approximations to. Previous studies have shown that these subtle forces are overestimated by popular *ab initio* quantum mechanical approximations such as DFT,[10] and they are completely missed by MP2. To reduce costs, it is routine to adopt simplified

approximations like ATM that effectively capture the leading physics of nonadditive three-body dispersion. Thus, the ATM leading-order asymptotic dispersion model used in this study accounts (or expresses) for the three-body dispersion energy as a summation over triplets of atoms in the system.[125,126] Thus the expression for intermolecular Axilrod-Teller-Muto triple-dipole dispersion for trimer ABC is

$$E_{ATM} = \sum_{a\in A}\sum_{b\in B}\sum_{c\in C} f_9^{abc}(\beta)E_{ATM}^{abc} \tag{4.20}$$

where:

$$E_{ATM}^{abc} = C_9^{abc}\frac{(1+3\cos\theta_a\cos\theta_b\cos\theta_c)}{(R_{ab}R_{ac}R_{bc})^3} \tag{4.21}$$

represents the triple-dipole ATM dispersion contribution. Here, R_{ab} is the distance between atoms a and b in monomers A and B, respectively (and similarly for R_{ac} and R_{bc}). The angles $\theta_a, \theta_b, \theta_c$ represent the triangle with each atom at the vertex, and C_9^{abc} is the dispersion coefficient for the atom triplet which can be obtained from Casimir-Polder integration of distributed frequency-dependent dipole-dipole polarizabilities from the coupled Kohn-Sham (CKS) formalism.[112]

$$C_9^{abc} = \frac{3}{\pi}\int_0^\infty d\omega\alpha^a(i\omega)\alpha^b(i\omega)\alpha^c(i\omega) \tag{4.22}$$

Because the total three-body contribution is around 5-20%[127] of the complete dispersion energy, it is reasonable to approximate C_9^{abc} from the C_6^{xy} coefficients as

$$C_9^{abc} \approx \sqrt{C_6^{ab}C_6^{ac}C_6^{bc}} \tag{4.23}$$

We obtain equation (4.20) from the summation of equation (4.21) over all atom triplets in the trimer. Each atom is from a different molecule, as terms involving atoms from the same molecule would contribute to the monomer or dimer dispersion term, therefore not forming a part of the non-additive three-body term. It is, however, important to note that at short-range (short interatomic distances), a damping is critical for the ATM dispersion model to prevent unphysical divergence to the three-body dispersion energy in addition to correcting for overlap effects. The damping

function, $f_9^{abc}(\beta)$ is written as a product of three two-body Tang-Toennies (TT) damping functions.[128,129]

$$f_9^{abc}(\beta) = f_6^{ab}(R_{ab}\beta) f_6^{ac}(R_{ac}\beta) f_6^{bc}(R_{bc}\beta) \tag{4.24}$$

where the two-body function $f_6(R\beta)$ takes the form:

$$f_6(R\beta) = 1 - \sum_{k=0}^{6} \frac{(R\beta)^k}{k!} e^{-\beta R} \tag{4.25}$$

and β is an empirical parameter that linearly depends on the sum of CCSD van der Waals radii r^{vdW} of the atoms

$$\beta = -0.31(r_a^{vdW} + r_b^{vdW}) + 3.43\,\text{bohr}^{-1} \tag{4.26}$$

The atomic van der Waals radii are obtained from the work of Tkatchenko *et al.* with no further reoptimizations.[126,128] The entire workflow for computing the TT-damped ATM dispersion term follows Ref. 126.

We performed the benchmark and two-body SAPT(DFT) calculations using the MOLPRO program.[91] For all wavefunction-based SAPT calculations (SAPT0–SAPT2+3(CCD)δMP2), the Psi4 code was used.[92] Nonadditive SAPT(DFT) results were obtained using the three-body module of the SAPT package,[93] importing the DFT quantities from DALTON 2.0.[94] The XSAPT computations utilized the Q-Chem code.[95] Finally, the ATM three-body dispersion terms were computed by a code written by Austin Wallace (Georgia Tech): see `https://github.com/Awallace3/dispersion`.

4.2. Results and Discussion

4.2.1. Benchmark interaction energies

Figures 4.2 and 4.3 provide a summary of the total benchmark interaction energy as well as its two-body and three-body contributions. The detailed benchmark values for each system are provided in Tables S1–S6 in the Supporting Information. The reference total interaction energy ranges from -2.190 to -21.755 kcal mol^{-1} for water complexes and from -1.476 to -17.588

kcal mol^{-1} for methane complexes.[14] The trimers with water show overall higher interaction energies than the trimers with methane, with I and Br showing the highest interaction and Cl the lowest interaction of the halogens. This can be linked to the size of the halogen atom involved in the interaction. The benchmark nonadditive contribution can be either attractive or repulsive and ranges between -2.87–0.69 kcal mol^{-1} for trimers with water, and only between -0.37–0.18 kcal mol^{-1} for trimers with methane. Complexes 5 and 10 in Fig. 4.1 show the systems that exhibit the largest nonadditive effects for methane and water complexes in the 3BXB dataset. Relative magnitudes of the nonadditive interaction energies are in the range 0.4–23 %.

In the next subsections, we will analyze the performance of various SAPT flavors in addition to MP2 and CCSD(T) methods applicable to our set of σ-hole noncovalent interactions. The MP2/CBS data includes the correlation energy extrapolated to the CBS limit from aQZ and a5Z. The wavefunction SAPT (SAPT0 to SAPT2+3(CCD)δMP2), two-body SAPT(DFT), and XSAPT calculations were performed in the aTZ basis while XSAPT and three-body SAPT(DFT) calculations were only feasible for complexes not involving iodine. The SAPT(DFT) calculations were performed in jun-cc-pVDZ for complexes with more than 25 atoms and aDZ otherwise. We will first investigate the individual interaction energy contributions, at the two-body and three-body levels, before moving on to examining the accuracy of the entire interaction energies using mean unsigned errors (MUE) relative to the CCSD(T)-level benchmark values.

4.2.2. SAPT two-body interaction energy contributions

Tables 1 to 6 in Appendix B show the two-body interaction energy components for the 3BXB structures from SAPT2+3(CCD)δMP2, SAPT(DFT), and XSAPT. The two-body components to the interaction energy are all attractive for those minimum structures, with the exception of exchange which is always repulsive. The magnitude of the interaction energy components depends on the halogen type with additional variation influenced by the presence of water or methane in the complex.[130] The trimers constructed by adding water to the pre-existing dimers are predominantly bound by electrostatics with the exception of the aromatic complexes that show larger dispersion magnitudes. Second-order dispersion is the leading term for complexes constructed with methane,

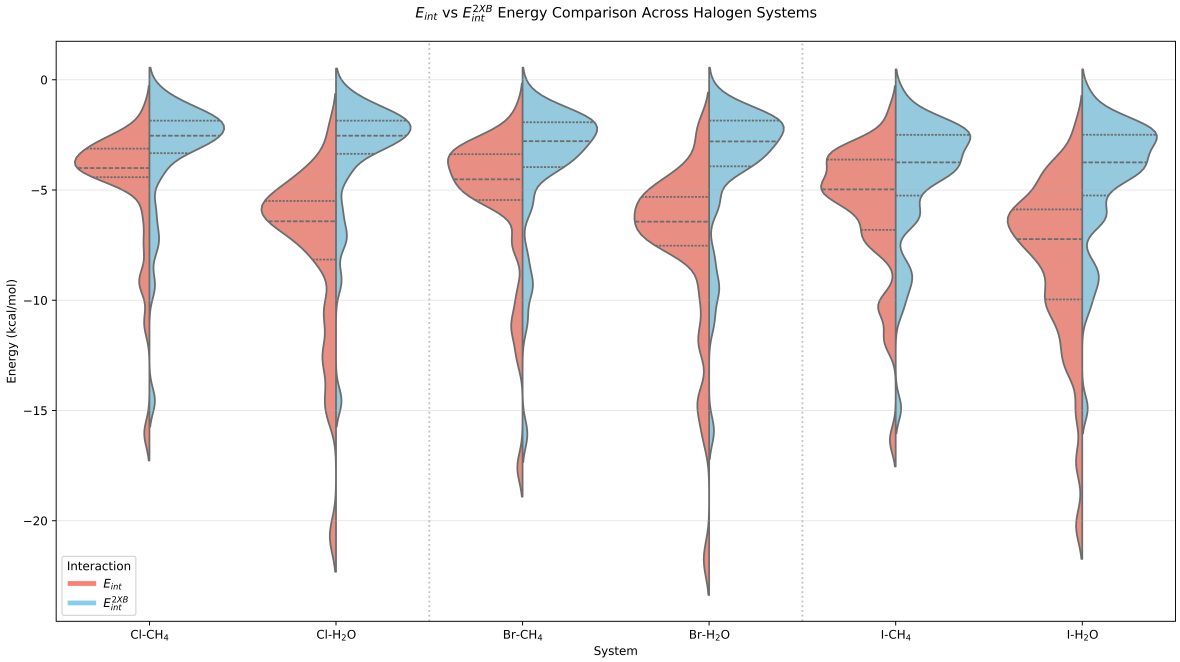


Figure 4.2: Violin plot highlighting differences in magnitude and variability between E_{int}^{3} and E_{int}^{2XB} across different halogens in the 3BXB dataset.

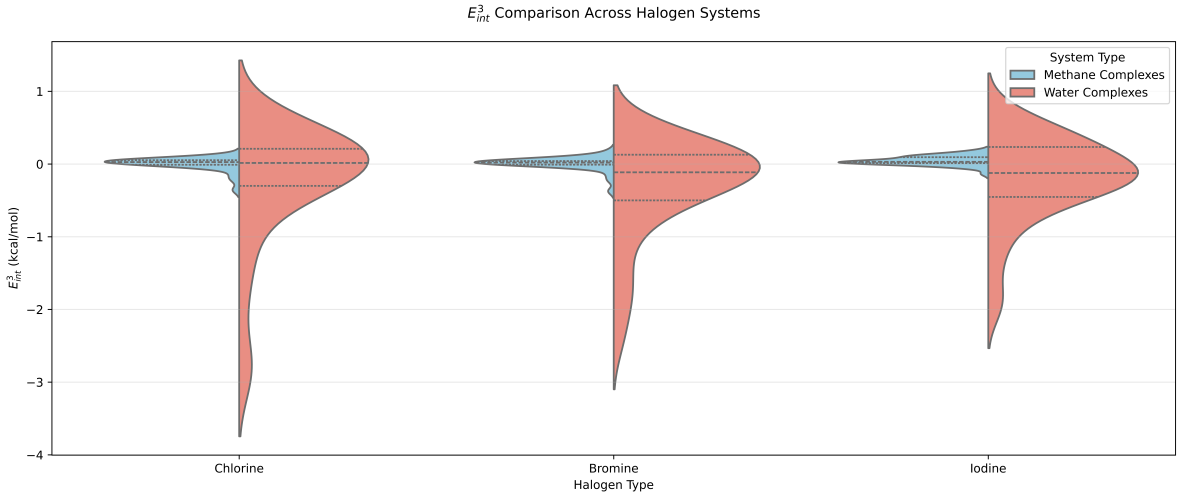


Figure 4.3: Distribution of Nonadditive E_{int}^{3} across water and methane complexes in the 3BXB dataset.

with a few exceptions for systems with Cl_2, Br_2 and I_2 as the donor. This behavior is expected because H_2O is a highly polar molecule, therefore significantly contributing to electrostatics whereas a nonpolar methane molecule drives dispersion forces. On the other hand, complexes with Cl_2, Br_2 and I_2 as illustrated by complexes 11 and 12 in Figure 4.1 exhibit strong electrostatic interactions because of the increased polarizability of the halogen atom.[14] With an increase in size, and thus polarizability, of the halogen atom, the largest contributors to the interaction, namely electrostatics and dispersion, also get stronger. In addition, the presence of electron withdrawing groups enlarges the σ-hole by pulling electron density away from the halogen, strengthening the interaction, while electron-donating groups act in the opposite direction.[14]

A detailed SAPT2+(CCD)δMP2 energy decomposition analysis of the 3BXB dataset shows that the interactions span the dispersion and electrostatics regions of the ternary diagram (Fig. 4.5).The two-body energy decomposition highlights that the methane complexes cluster towards the dispersion dominated region while the water complexes towards the electrostatics with stronger induction influence. Notably, chlorine complexes with methane and bromine complexes with water show the most pronounced induction effects. This is likely the result of the halogenated molecule inducing the electrostatic response in antoher nonpolar molecule.[14] To the contrary, iodine complexes, for both water and methane, are governed by a balance of dispersion and electrostatic interactions.

4.2.3. SAPT analysis of the three-body nonadditive interaction

The SAPT energy decomposition of the nonadditive interaction in the 3BXB dataset shows that nonadditive induction dominates while exchange and dispersion play smaller roles. Figures 4.3, 4.6 and 4.8 present an analysis of the three-body nonadditive interactions. Here, we can no longer use the two body SAPT2+3(CCD)δMP2 approach, but instead shift our attention to SAPT(DFT) and XSAPT. It should be noted that these nonadditive interaction energy calculations were not feasible for iodine complexes which were excluded. Consequently, the figures only include data from the bromine and chlorine complexes. As explained before, SAPT(DFT) includes all the leading order nonadditive effects while XSAPT neglects the nonadditive first-order

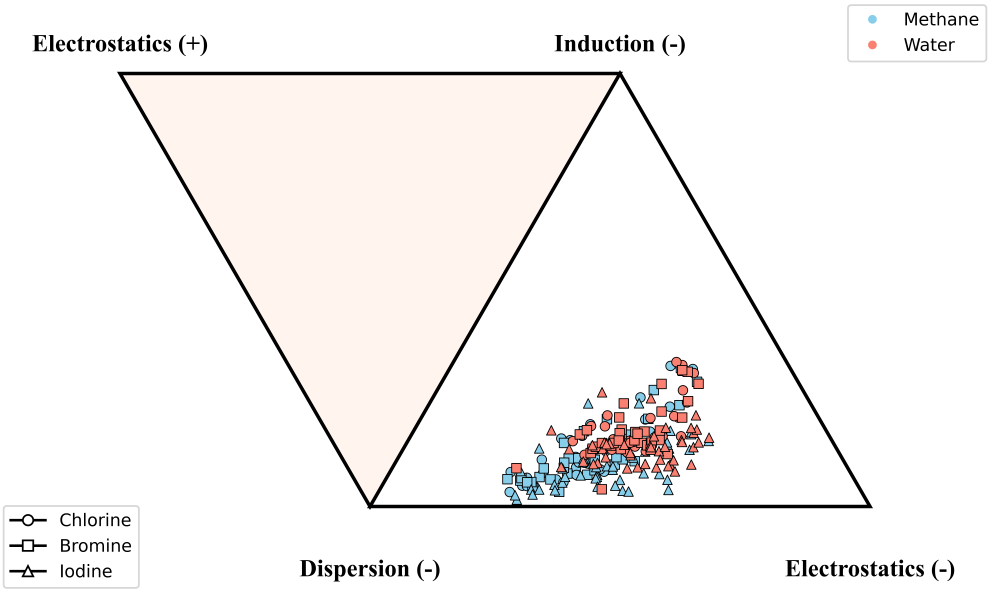


Figure 4.4: Two-body SAPT2+(CCD)δMP2 interaction energy distribution for both water and methane trimers studied here, displayed on a ternary diagram.

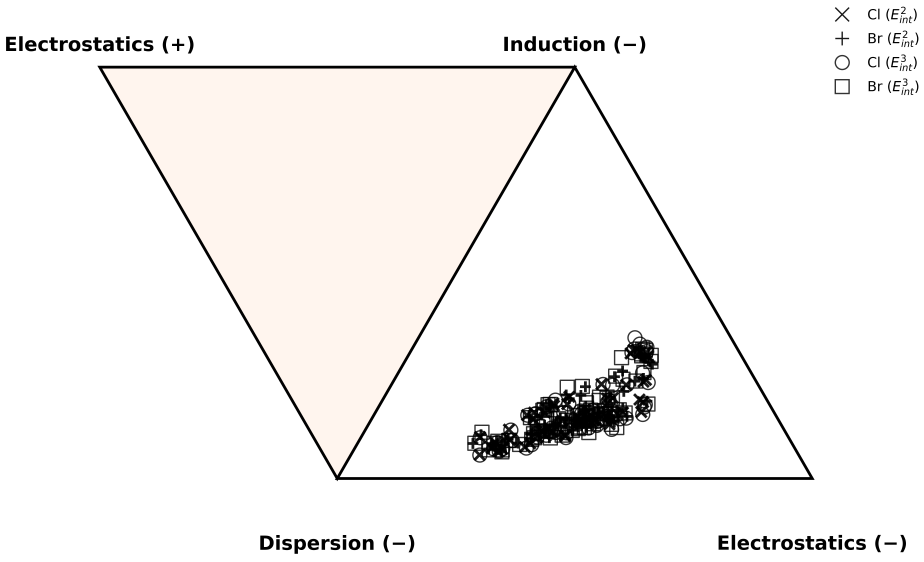


Figure 4.5: Two-body and two- plus three-body SAPT(DFT) interaction energy distribution for trimers in the 3BXB dataset.

exchange component. XSAPT includes, however, nonadditive induction through the explicit polarization (XPol) method, that is applicable to any number of molecules (not just three) and can be used with or without the higher order induction couplings.[67] As our earlier study[32] indicates that the latter couplings play a minor role, we neglect them here.

For the 3BXB dataset, the nonadditive interaction energies are relatively small in magnitude and range from 0.273 to -2.199 kcal mol^{-1} (can be either attractive or repulsive). The nonadditive exchange is predominantly attractive with the exception of a few complexes (chloromethane-trimethylamine-water/methane, bromine-thioacetone-water/methane and trifluorobromomethane-thioacetone-water/methane). The nonadditive dispersion in this dataset is repulsive as it accounts for a net reduction of the attraction of the system, remedying over-binding that might come from the two-body dispersion contribution. This component ranges from 0.020 to 0.288 kcal mol^{-1} at the SAPT(DFT) level. Nonadditive induction, conversely, can be both atrractive and repulsive and ranges from -2.309 to 2.142 kcal mol^{-1} for both water and methane complexes. An in-depth analysis of the nonadditive induction is shown in Fig. 4.8. It highlights this nonadditive component to the interaction energy from the minor variants of SAPT(DFT), XSAPT, and SAPT0. The SAPT0 variant does not comprise the δE^{HF} term while SAPT(DFT) can include or omit the nonadditive δE^{HF}. The δE^{HF} in XSAPT is assumed to be pairwise additive and thus has no effect on the nonadditive induction. Fig. 4.6 shows different dispersion estimates from the benchmark, XSAPT and SAPT(DFT) variants. Nonadditive dispersion is treated differently by all four methods. In the benchmark supermolecular approach, nonadditive dispersion can be approximated as a difference between CCSD(T) and MP2 as the latter entirely misses three-body dispersion. Note that this difference is not strictly due to dispersion energy as the molecular multipole moments, polarizabilities, and electron densities somewhat change between MP2 and CCSD(T), leading to differences in other contributions.

4.2.4. Total SAPT interaction energies

In this section, we will systematically evaluate the various SAPT flavors in reproducing the interaction energies relative to the benchmark. For clarity, this section is split into three parts where

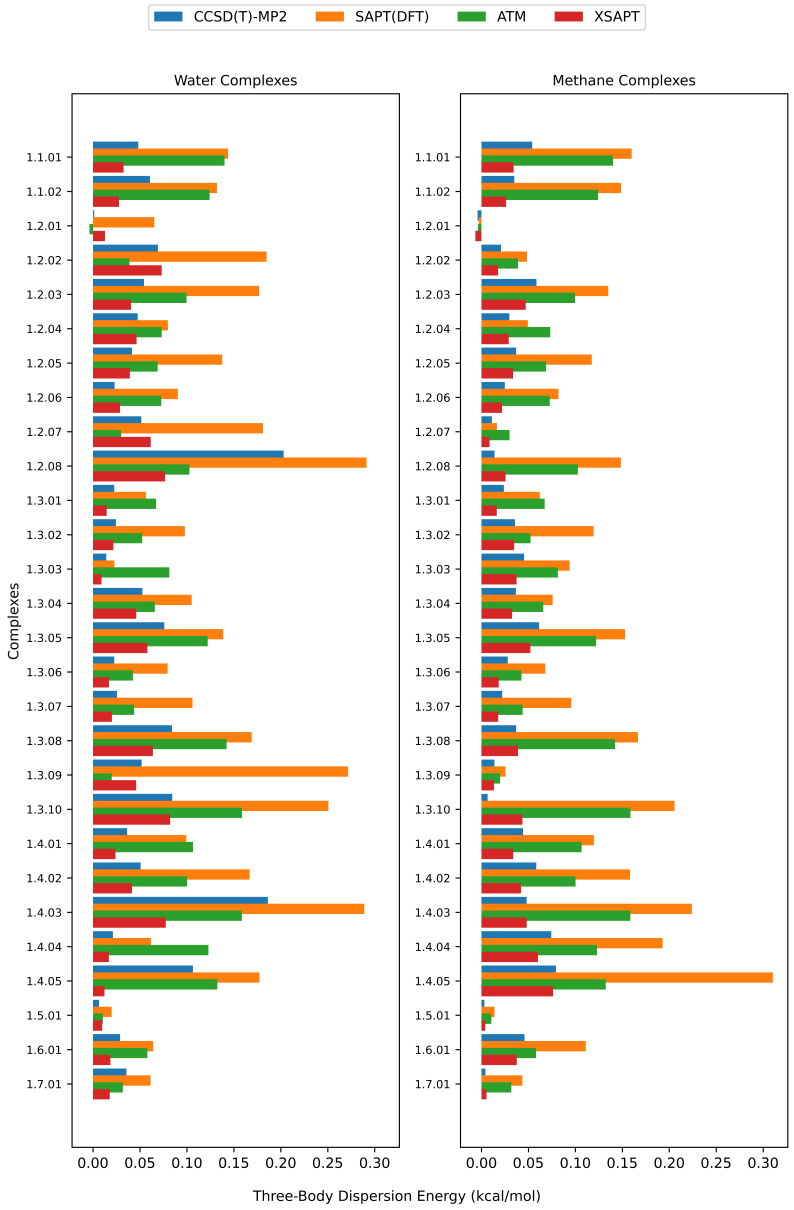


Figure 4.6: Three-body dispersion energy for Cl systems, estimated as the supermolecular CCSD(T)/CBS−MP2/CBS difference as well as computed with SAPT(DFT) and XSAPT (MBD), and ATM formalisms.

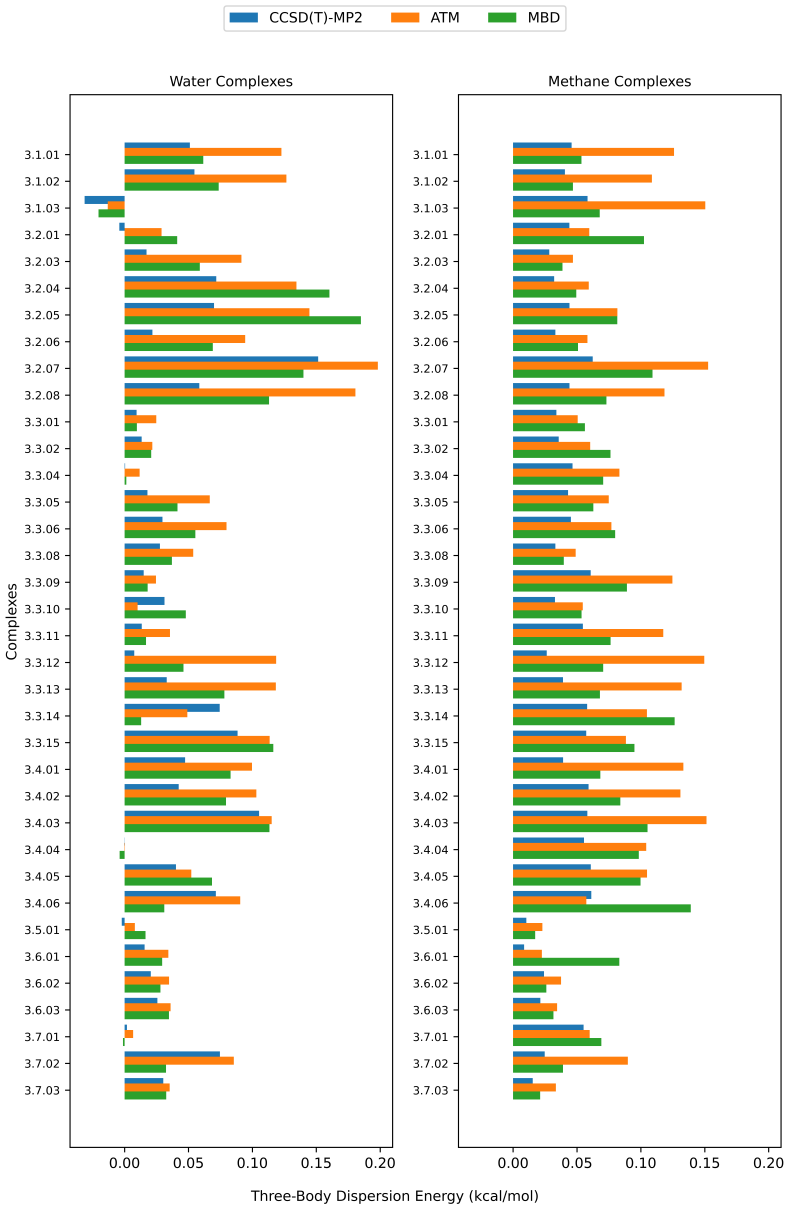


Figure 4.7: Three-body dispersion energy for iodine systems, estimated as the supermolecular CCSD(T)/CBS−MP2/CBS difference as well as computed with ATM and MBD.

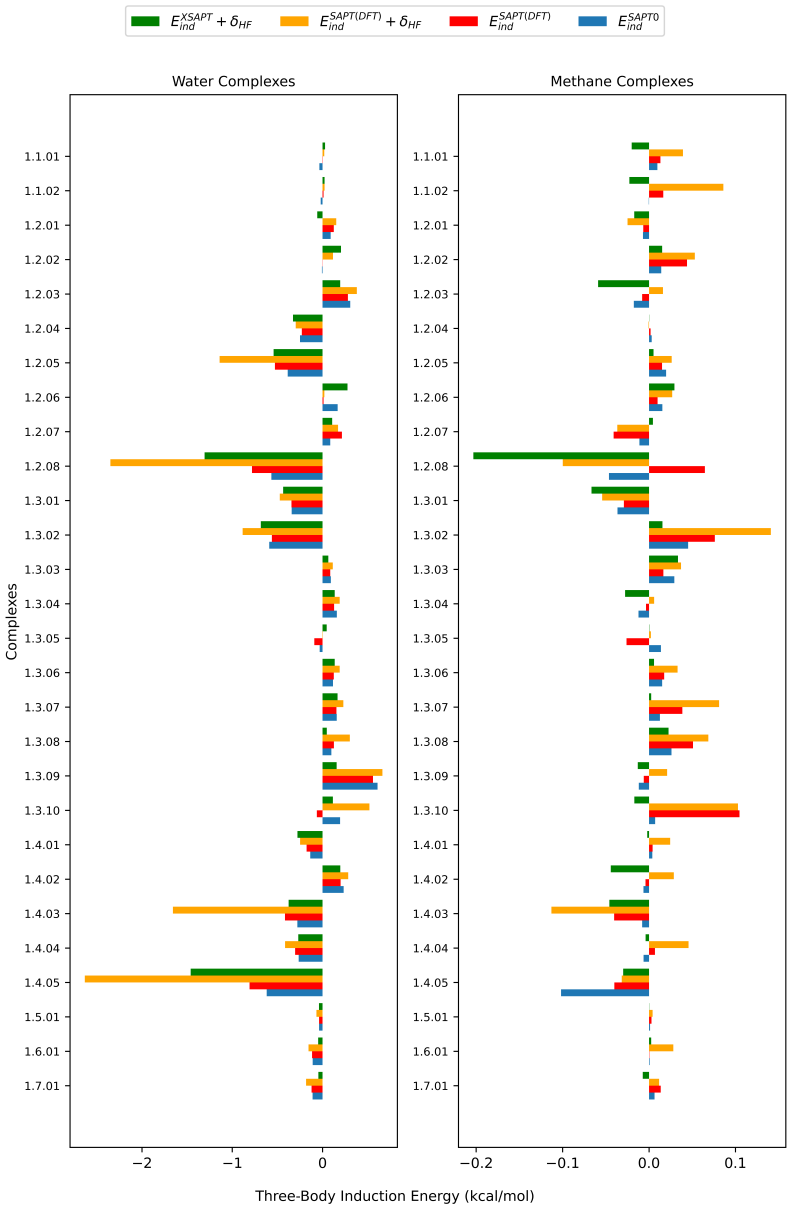


Figure 4.8: Comparison of the nonadditive three-body induction energies for the chlorine systems in the 3BXB dataset predicted by SAPT0, XSAPT, and SAPT(DFT).

we compare the SAPT variants against the two-body benchmark, nonadditive contributions and the three-body benchmarks. We also include hybrid methods

4.2.4.1. SAPT variants vs. 2-body benchmark

In this subsection we only evaluate the SAPT falvors that only have the pairwise interaction energies and compared against CCSD(T) benchmark that only includes the two body interaction energy. When compared against the two-body reference(Fig. 4.9), most methods(9 out of 16) generally perform better for methane complexes in contrast to the 6 out of 16 methods that exhibit lower MUEs the water complexes. Methane systems showing relatively lower MUEs likely comes from the fact that the systems are predominantly dispersion-bound and many of the methods, especially the SAPT methods are found to reliably describe all but the most difficult dispersion-bound systems.[8]However, the other methods that perform better for water systems, MP2, SAPT2, SAPT2+, SAPT2+(CCD), SAPT2+(3)(CCD)δMP2, SAPT2+3(CCD)δMP2 and SAPT2+(3)δMP2, likely benefit from how induction is treated. These systems exhibit induction and polarization effects and MP2 has been shown to perform well for hydrogen bonding type of interactions.[131] In addition, the δMP2 correction used for SAPT flavors works well for electrostatically dominated complexes by accounting for missing higher-order polarization in the perturbation series truncation.[8]

4.2.4.2. SAPT variants vs. nonadditive and 3-body benchmark

Here we analyze the methods that explicitly include the three-body term in the total energy and the MUE evaluated with respect to the three-body benchmark reference computed at the CCSD(T)/CBS level of theory. This allows us to evaluate the relative accuracy of these methods in capturing higher order effects that go beyond pairwise additivity. The MUEs presented in Fig 4.11 assess how accurately the different theoretical approaches describe the three-body interactions relative to the benchmark. We perform three sets of evaluations on three different classes of methods. First we consider two-body only methods that do not account for any nonadditive three-body effects. This is extremely useful in highlighting the errors as a result of ignoring completely the three-body nonadditive effects.

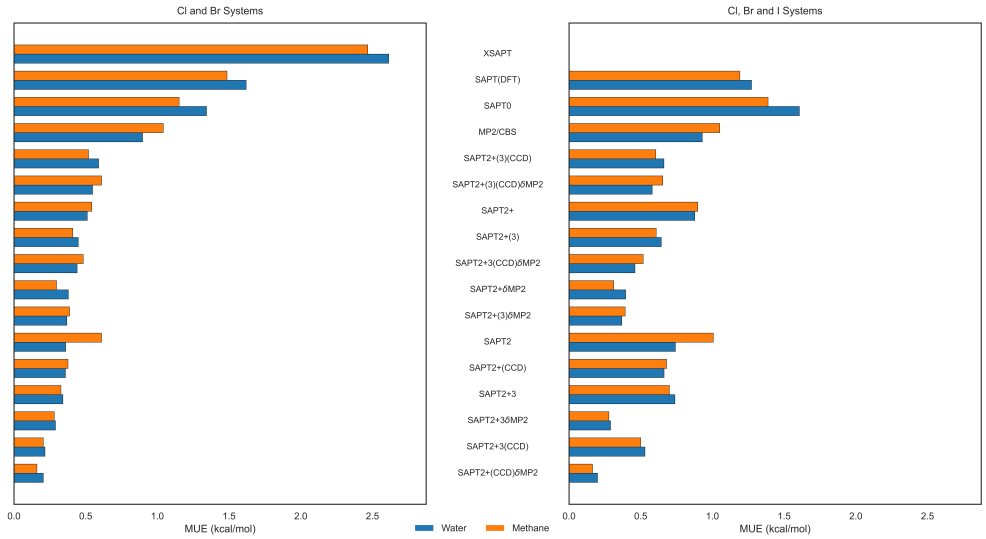


Figure 4.9: Mean unsigned errors (MUE, kcal/mol) of various two-body interaction energy estimates relative to the benchmark CCSD(T)-level E^2_{int} values.

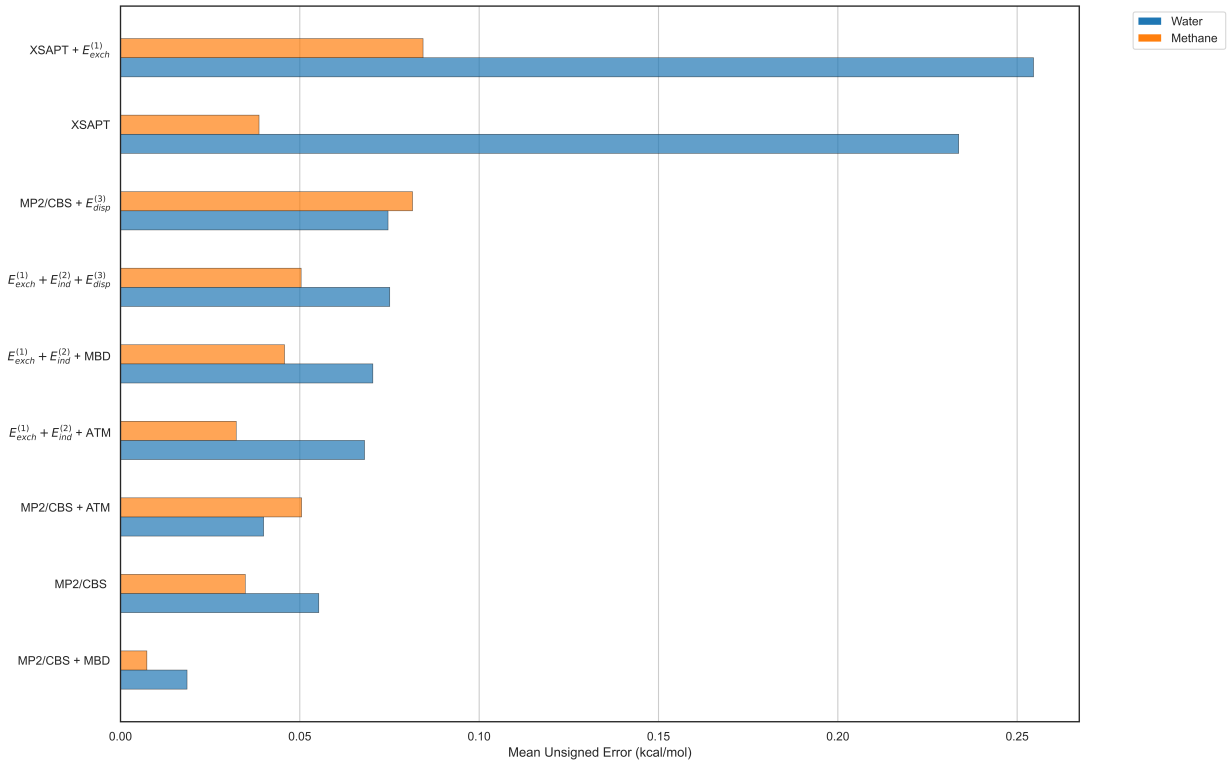


Figure 4.10: Mean unsigned errors (MUE, kcal/mol) of various nonadditive three-body interaction energy estimates relative to the benchmark CCSD(T)-level E^3_{int} values.

Similarly, we evaluate methods that include the nonadditive component either through energy decomposition like the SAPT formalism, or full calculation as in the canonical supermolecular approach. Lastly, we include a class of hybrid methods created by approximating the total interaction energy through the summation of independent components. This scheme assumes that the interaction components can be treated independently as in the many body expansion(MBE).[3] It is worth noting that despite lacking the explicit three-body term, the two-body methods exhibit lower MUE values in comparison to those that include the three-body nonadditive effects, i.e SAPT(DFT), XSAPT and MP2. These two-body methods show only a slight increase in the MUE values compared to their performance against the two-body reference. The performance of the hybrid methods illustrated in Fig.4.11, consistently show improved accuracy relative to their two and three-body counterparts in reference to the three-body benchmark. The addition of the individual three-body terms significantly reduces the MUE to below the range of the fully three-body-inclusive methods. Among all the tested methods, SAPT2+(CCD)δMP2 performs the best across the board. Even when combined with the nonadditive components computed differently, it consistently yields low MUE values.

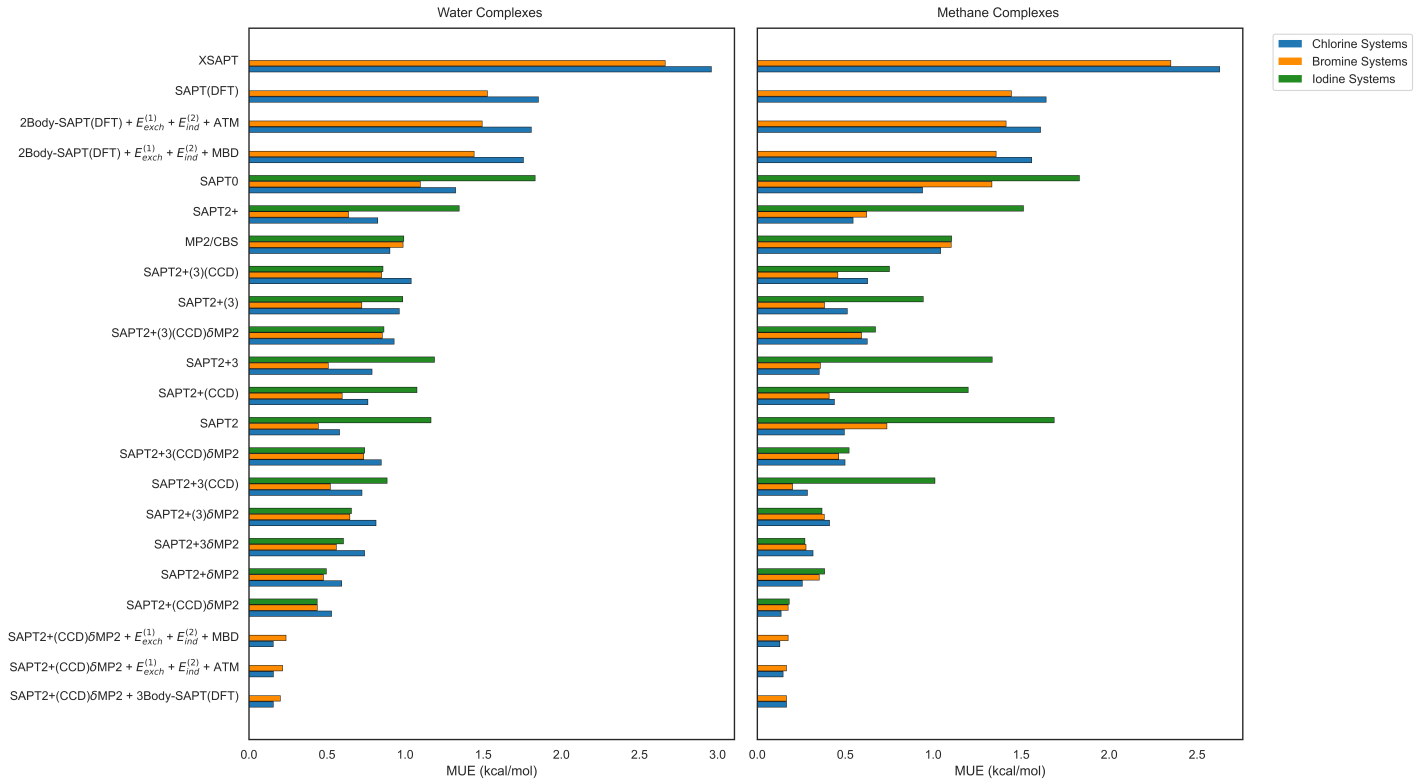


Figure 4.11: Mean unsigned errors (MUE, kcal/mol) of the total (two-body plus three-body) interaction energy estimates relative to the benchmark CCSD(T)-level complete E_{int} values.

4.2.5. Conclusions

The new 3BXB dataset introduced in this work represents a significant step towards the unification of the coverage of nonadditive noncovalent interactions. We built on the SH250 dataset to include trimers and assessed the performance of various SAPT flavors in reproducing the CCSD(T)-level benchmark data. This enhanced diversity of the dataset to include water and methane trimers was also useful in evaluating the nonadditive component of the interaction energy that comprises first-order exchange, dispersion, and induction. To this end, we performed an analysis of errors of the SAPT(DFT), XSAPT, MP2/CBS as well as hybrid models that integrate the nonadditive components from multiples sources. The new data obtained shows a trend of increasing errors when describing systems with iodine and chlorine. Bromine systems show the least for both systems with water and methane. Another important highlight is that, besides the different flavors and variants of SAPT providing a powerful tool in understanding the two-body, nonadditive and three-body interactions, SAPT2+(CCD)δMP2 emerges as the superior performer. The energy decomposition at the two and three-body levels, the complexes are bound by dispersion and electrostatics with significant influence of induction for the water complexes and increased electrostatic influence is also observed in Cl_2, Br_2 and I_2 systems due to increased polarizability of the halogen atom. For the nonaddtive interaction energy components, nonadditive induction dominates whereas nonadditive exchange and dispersion play relatively smaller roles but are system dependent. We extended this analysis to the hybrid methods applied to both nonadditive and total three-body interactions. We have demonstrated that the mix−and−match hybrid approaches generated by combining independent components offer a favorable balance of accuracy and computational cost.

Chapter 5

The Interactions of the OH Radical within Interstellar Objects

5.1. Introduction

The OH radical is central to both interstellar and atmospheric chemistry. In earth's chemistry, its abundance and readiness to interact with other species puts it at the core of water formation and polutant degradation.[33,132] In the interstellar medium, the OH radical offers a complementary view of oxygen chemistry in the gas phase through reactions in the interstellar medium[133] thereby contributing to our understanding of star and planet formation through the study of chemical and structural evolution of disks. Therefore, the first aspect of this project involves the study of interactions of the OH radical with the hydrogen molecule with the focus on vibrationally inelastic collisions. Our main task is to develop accurate OH-H_2 six dimensional potential energy surfaces (PES) in a bid to describe molecular structure and collisions. Previously, for low-energy scattering applications, rigid 4D PESs were used to describe this interaction.[134,135] The shortfall of the 4D PESs is that they do not describe vibrational transitions essential in the calculations of scattering cross-sections. Our development of the 6D PES and further improvements are meant to focus on the inclusion of residual basis set effects beyond the aug-cc-pVTZ/aug-cc-pVQZ level and provide a more accurate treatment of electron correlation through the inclusion of higher-order coupled-cluster excitations.[136] The first steps involve the description of specific OH-H_2 interacting geometries, and then we compare diagonal and off-diagonal interaction energies while enforcing a C_s symmetry restriction on the complex, where the two potential energy surfaces have distinct A' and A'' symmetries.

5.2. Computational Details

To describe geometry of the molecules (OH and H_2), we use R which is the distance between the centers of mass of both molecules, r_{HH} and r_{OH} that account for the bond lengths within each

molecule, the angle θ_H and θ_O between the molecular axes and the center-of-mass axis and a dihedral angle ϕ_H. The calculations were carried out in a 6 dimensional grid (R, θ_H, θ_O, ϕ_H, r_{HH}, r_{OH}) with about 101,500 grid points. The OH-H_2 distance is described by 18 values of R ranging from 4.0 to 20.0 bohr (R:

4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 10.0, 11.0, 12.0, 13.0, 14.0, 15.0, 20.0). The angular grid comprises 225 orientations which are combinations of 5 values of θ_H defined by $\cos\theta_H$ = 0 to 1 in steps of 0.25, 9 values of θ_O defined by $\cos\theta_O$ = -1 to 1 in steps of 0.25, and 5 values of ϕ_H from 0 to π in steps of 45^o. The grid values $r_{HH} \in \{1.1, 1.3, 1.448736, 1.6, 1.75\}$ bohr[137] and $r_{OH} \in \{1.55, 1.7, 1.8509, 2.05, 2.25\}$ bohr allow the molecules to undergo vibrations but not break apart. The r_{HH} = 1.448736 bohr and r_{OH} = 1.8509 bohr are the interatomic distances averaged over the ground vibrational states of the molecules H_2 and OH respectively ($v = 0$). The other inter-atomic distance points are obtained by first calculating the classical turning points on the 1D potential energy curves of the molecules. The classical turning points are where the molecule has zero kinetic energy while the potential and total energies are approximately equal. They also mark boundaries between regions where a classical particle is allowed and where it is not. Therefore, this is useful in obtaining points that adequately describe the vibrations without the bond breaking. The interaction energy for a given geometry would be given by the dimer energy minus the sum of the monomer energies as in a supermolecular approach including counterpoise (CP) correction for basis set superposition error.

The calculations involve RCCSD(T) at the aQZ basis with variable bond functions , along with a correction at CCSDT level of theory using aDZ basis with bond functions. Here, the variable bond functions are additional basis functions whose positions adjust dynamically based on the OH interatomic distance[138–141] and such midbonds have been shown to improve the description of weakly interacting systems than explicitly correlated methods.[142] All RCCSD(T) calculations and CCSDT corrections for C_{2v} symmetry have been completed at 1800 points. For the C_s symmetry

are finalized with corresponding CCSDT corrections are complete for both symmetries. Aditionally, the calculations for the C_1 symmetry were performed at he MRCI-F12 level of theory at the aQZ basis without any bond functions.

We eventually adopted an analytical fit that currently incorporates the radial functional form: [137]

$$V(R, r_{OH}, r_{HH}) = \sum_{k=0}^{K} (r_{OH} - r_{OH}^0)^k \sum_{l=0}^{L} (r_{HH} - r_{HH}^0)^l \sum_{m=0}^{M} A_{klm} R^m exp(-\alpha_{kl} R) - \sum_{k=0}^{K'} (r_{OH} - r_{0H}^0)^k \sum_{l=0}^{L'} (r_{HH} - r_{HH}^0)^l \left(\frac{C_{6kl}}{R^6} f6(dR) + \frac{C_{8kl}}{R^8} f8(dR) \right) \tag{5.1}$$

$$V(R, r_{OH}, r_{HH}, \theta_O, \phi) = \sum_{l, l_1, L_2} V^{l l_1 L_2}(R, r_{OH}, r_{HH}) \cdot A_{l l_1 L_2}(\theta_O, \theta_H, \phi) \tag{5.2}$$

where A_{klm} represent short-range coefficients fitted to *ab initio* data, α_{kl} are nonlinear parameters, C_{8kl} long-range constants, d is the damping parameter in the fitting process and C_6 is an asymptotic constant. The RCCSD(T) calculations provide data for the symmetric geometries(C_{2v} and C_s) and to account for higher order correlation effects, we applied a CCSDT$-$CCSD(T) correction.

5.2.1. Comparison of methods and summary of data

In an effort to determine the most accurate and computationally feasible method to generate the 6D OH-H_2 PES, we performed extensive preliminary calculations of interaction energies at different levels of theory. The comparisons include both diagonal(V_d) and off-diagonal(V_o) interaction energies at multiple intermolecular separations(R = 4.75, 5.0, 5.25, and 6.0) for C_1, C_{2v} and C_s symmetries. [38] Tables 5.1 and 5.3 summarize results obtained from RCCSD(T), MRCC and MRCI-based approaches evaluated at varied basis sets ranging from aDZ to a5Z. All the calculations are performed with bond functions and JK fitting was also performed where applicable. While the MRCC is primarily designed for multi-reference methods, here it was solely used for the restricted open-shell Hartree Fock(ROHF) to generate reference orbitals for the subsequent

CCSDT or CCSD(T) calculations. Tables 5.2 and 5.4 shows the CCSDT−CCSD(T) corrections for C_{2v} and C_s symmetries respectively that captures higher order correlation effects beyond perturbative triples. Finally table 5.5 only focuses on MRCI methods for the C_1 symmetry where single reference methods are insufficient.

The tables in this section allow us to asses the accuracy of our choice of methods(RCCSD(T) and MRCI) against benchmarks by Dagdigian.[38] We draw important conclusions from table 5.1. With accuracy and computational cost as the primary determinants to the choice of method, we tested a variety of the MRCI-based approaches. The MRCISD+Q method at AVTZ-JKFIT(3s,3p) and AVQZ-JKFIT(3s,3p) both yielded interaction energies in agreement with the CBS data.[38] While active spaces yield better descriptions of the multireference character, they bring additional cost and complexity[143] that was not justifiable for the C_1 symmetry. Thus MRCI-F12 at AQZ basis set was selected as the choice method. Evaluating the performance of RCCSD(T) method and basis set combinations, variants using the variable bond functions with the non-scaled triples show overall agreement with Dagdigian[38] and coworkers. At ATZ and AQZ levels, comparing them to the CBS data, the AVQZ + BF/non-scaled triples variant shows excellent agreement with CBS, with differences of $< 1cm^{-1}$ in the diagonal(V_d) and practically identical in the off-diagonal (V_0) elements. ATZ basis on the other hand overestimates the diagonal elements while calculations at the A5Z basis showed significant convergence problems at shorter intermolecular distances for both C_{2v} and C_s symmetries.

Table 5.1: Comparison of computed diagonal V_d and off-diagonal V_o interaction energies (cm^-1) at three values of R for the C_{2v} Symmetry. In the table AVnZ are shorthand notations for the aug-cc-p-VnZ basis sets

Method/basis set	V_d			V_o		
	5.0	5.25	6.0	5.0	5.25	6.0
MRCI						
AVQZ	272.29	4.48	−217.82	−35.50	−26.43	−12.33
AVQZ/3-state CAS	262.60	−2.35	−220.22	−35.36	−26.28	−12.21
AVTZ(3s,3p)	221.51	−43.50	−251.91	−38.11	−28.71	−13.66
AVQZ(3s,3p)	238.49	−20.77	−228.46	−36.11	−26.77	−12.32
AV5Z(3s,3p)	249.49	−10.42	−219.50	−34.60	−25.61	−11.74
AVQZ(3s,3p) 3-state CAS	238.51	−20.75	−228.44	−36.11	−26.77	−12.32
MRCISD+Q/ATZ-JKFIT	286.14	16.45	−208.22	−33.96	−25.24	−11.99
MRCISD+Q/AQZ-JKFIT	280.25	12.24	−212.15	−34.44	−25.71	−12.13
MRCISD+Q/ATZ-JKFIT(3s3p)	256.06	−5.64	−216.43	−34.17	−25.35	−12.11
MRCISD+Q/AQZ-JKFIT(3s,3p)	253.52	−6.63	−217.94	−34.06	−25.18	−11.47
RCCSD(T)						
F12 ADZ + Variable BF/non-sc	294.70	26.55	−200.03	−34.88	−25.82	−11.83
F12 AVTZ + Variable BF/sc	254.32	−5.55	−216.04	−34.42	−25.47	−11.66
F12 AVTZ + Variable BF/non-sc	261.51	0.10	−213.30	−34.28	−25.37	−11.62
F1 AVTZ Constant BF/sc	252.22	−7.41	−217.26	−34.26	−25.33	−11.59
F1 AVTZ Constant BF/non-sc	259.14	−1.93	−214.58	−34.13	−25.24	−11.56
F12 AVTZ/sc	260.07	−1.24	−214.64	−34.50	−25.52	−11.68
F12 AVTZ/non-sc	267.34	4.45	−211.91	−34.35	−25.42	−11.64
F12 AVQZ/sc	254.39	−5.23	−215.88	−34.14	−25.25	−11.55
F12 AVQZ/non-sc	257.84	−2.52	−214.58	−34.07	−25.21	−11.53
F12 AVQZ + Constant BF/sc	253.19	−6.34	−216.55	−34.10	−25.21	−11.52
F12 AVQZ + Constant BF/non-sc	256.48	−3.76	−215.31	−34.04	−25.17	−11.50
F12 AVQZ + Variable BF/non-sc	256.34	−3.80	−215.27	−34.06	−25.19	−11.52
F12 AVQZ + Variable BF/sc	253.07	−6.37	−216.52	−34.12	−25.23	−11.53
F12 AV5Z/sc	254.07	−5.51	−215.98	−34.04	−25.17	−11.50
F12 AV5Z/non-sc	255.92	−4.05	−215.28	−34.00	−25.14	−11.49
F12 AV5Z + Constant BF/sc	253.82	−5.75	−216.16	−34.03	−25.16	−11.50
F12 AV5Z + Constant BF/non-sc	255.61	−4.35	−215.48	−34.00	−25.14	−11.49
F12 AV5Z + Variable BF/sc	253.68	−5.85	−216.19	−34.03	−25.16	−11.50
F12 AV5Z + Variable BF/non-sc	255.44	−4.46	−215.53	−34.00	−25.14	−11.49
AVDZ/Non-F12	422.78	136.51	−140.23	−34.93	−25.64	−11.47
AVDZ/non-F12 + Constant BF	309.17	30.35	−210.48	−35.81	−26.52	−12.16
AVDZ/Non-F12 + Variable BF	360.98	66.97	−194.81	−35.07	−25.67	−11.38
AVTZ/non-F12	310.46	34.95	−200.31	−34.76	−25.72	−11.78
AVTZ/non-F12 + Constant BF	281.18	13.55	−209.56	−34.62	−25.58	−11.68
AVTZ/non-F12 + Variable BF	286.06	18.16	−206.05	−34.76	−25.71	−11.75
AVQZ/non-F12	272.59	8.10	−210.94	−34.26	−25.34	−11.59
AVQZ/non-F12 + Constant BF	266.98	3.50	−212.74	−34.26	−25.32	−11.57
AVQZ/non-F12 + Variable BF	266.45	3.43	−212.46	−34.29	−25.35	−11.59
AV5Z/non-F12	262.66	0.75	−213.54	−34.12	−25.22	−11.52
AV5Z/non-F12 + Constant BF	261.46	−0.24	−214.09	−34.11	−25.22	−11.52
AV5Z/non-F12 + Variable BF	260.49	−0.85	−214.16	−34.11	−25.22	−11.52
MRCC						
CCSD(T)/aDZ	422.24	136.13	−140.40	35.00	25.68	11.49
CCSD(T)/aDZ + BF (semicanonical)	308.42	29.33	−210.75	35.89	26.61	12.33
CCSDT/aDZ	302.07	23.05	−143.25	36.45	26.90	−11.57
CCSDT/aDZ (semicanonical)	414.92	130.39	−143.25	35.28	25.88	11.57
CCSDT/aDZ + BF (semicanonical)	299.34	22.11	−214.33	36.17	26.81	12.40
CCSDT(Q)/aDZ	414.94	130.33	−143.45	35.27	25.87	11.56
CCSDT(Q)/aDZ + BF	298.99	21.79	−214.50	36.20	26.82	12.40
CCSDTQ/aDZ (semicanonical)	414.21	129.77	−143.62	35.32	25.90	−11.57
CCSD(T)/aTZ (semicanonical)	309.85	37.38	−198.67	34.83	28.64	13.63
CCSDT/aTZ	299.87	26.66	−203.68	−35.00	−25.89	−11.85
CCSDT/aTZ (semicanonical)	301.93	28.16	−203.68	35.05	25.92	11.85
CCSD(T)/aTZ+BF (semicanonical)	280.54	12.69	−210.38	34.69	25.63	11.70
CCSDT/aTZ + BF (semicanonical)	272.95	6.62	−213.43	34.91	25.78	11.75
CCSD(T)/aQZ	271.97	7.64	−211.16	34.33	25.39	11.61
CCSD(T)/aQZ +BF	266.35	3.03	−213.05	34.32	25.37	11.58
CCSDT/aQZ (semicanonical)	264.90	1.96	−214.02	34.52	25.52	11.66
CCSDT/aQZ + BF (semicanonical)	259.51	−2.49	−215.84	34.53	25.50	11.63
CCSDT(Q)/aTZ	301.16	29.22	−204.20	35.07	25.95	−11.85

Table 5.2: Comparison of CCSDT − CCSD(T) interaction energy Corrections of computed diagonal V_d and off-diagonal V_o interaction energies (cm^-1) at three values of R for C_{2v} Symmetry.

Method/basis set	V_d			V_o		
	5.0	5.25	6.0	5.0	5.25	6.0
$\delta^{CCSDT}_{CCSD(T)}$/aDZ	−7.32	−5.74	−2.85	0.28	0.20	0.08
$\delta^{CCSDT}_{CCSD(T)}$/aDZ + BF	−9.09	−7.22	−3.58	0.28	0.20	0.07
$\delta^{CCSDT}_{CCSD(T)}$/aTZ	−7.92	−9.23	−5.00	0.22	−2.73	−1.78
$\delta^{CCSDT}_{CCSD(T)}$/aTZ + BF	−7.59	−6.07	−3.04	0.22	0.15	0.05
$\delta^{CCSDT(Q)}_{CCSDT}$/aDZ	0.02	−0.05	−0.20	−0.01	−0.02	−0.07
$\delta^{CCSDTQ}_{CCSDT(Q)}$/aDZ	−0.73	−0.57	−0.17	0.05	0.04	0.00

Table 5.3: Comparison of computed diagonal V_d and off-diagonal V_o interaction energies (cm^-1) at three values of R for the C_s Symmetry. In the table AVnZ are shorthand notations for the aug-cc-p-VnZ basis sets

Method/basis set	V_d				V_o			
	4.75	5.0	5.25	6.0	4.75	5.0	5.25	6.0
MRCI								
AVQZ	440.8313	250.2055	133.9893	2.6502	−358.33	−240.21	−161.54	−51.09
AVQZ/3-state CAS	444.5503	252.8468	135.9075	3.5061	−358.64	−240.49	−161.77	−51.20
AVTZ (3s,3p)	420.1567	231.8508	117.8195	−6.6896	−375.70	−251.68	−169.15	−53.56
AVQZ(3s,3p)	421.1345	235.2560	122.2353	−3.8013	−375.40	−251.60	−169.09	−53.12
AV5Z(3s,3p)	425.4757	239.8474	126.6884	−0.4258	−374.93	−251.49	−169.25	−53.47
AVQZ(3s,3p) 3-state CAS	421.5899	235.6302	122.5327	−3.6674	−375.36	−251.59	−169.11	−53.17
MRCISD+Q/ATZ-JKFIT	408.2272	226.1280	116.2217	−3.4359	−371.72	−249.80	−168.43	−53.80
MRCISD+Q/AQZ-JKFIT	404.1219	222.7415	113.2006	−6.6457	−368.72	−247.09	−166.05	−52.29
MRCISD+Q/ATZ-JKFIT(3s3p)	396.7180	217.3841	109.5200	−6.1552	−380.07	−255.22	−171.96	−55.04
MRCISD+Q/AQZ-JKFIT(3s,3p)	392.2330	213.6629	106.1028	−10.3757	−375.57	−251.43	−168.83	−53.00
RCCSD(T)								
F12 AVDZ + Variable BF/sc	403.9055	223.2065	113.8055	−6.5457	−380.17	−255.13	−171.81	−54.32
F12 AVTZ + Constant BF/sc	387.6975	210.0665	103.2533	−11.6335	−376.14	−251.84	−169.12	−52.99
F12 AVTZ + Constant BF/non-sc	392.2284	213.6159	106.0528	−10.2182	−375.58	−251.53	−168.96	−53.01
F12 AVTZ + Variable BF/sc	389.8065	211.9109	104.8223	−10.8032	−375.97	−251.72	−169.03	−52.99
F12 AVTZ + Variable BF/non-sc	394.5331	215.6021	107.7229	−9.3541	−375.38	−251.38	−168.86	−53.01
F12 AVTZ/sc	395.4018	215.6967	107.4638	−9.5473	−376.79	−252.34	−169.48	−53.16
F12 AVTZ/non-sc	399.9189	219.1919	110.1897	−8.2063	−376.17	−251.98	169.29	−53.16
F12 AVQZ/sc	391.5293	213.3471	106.0452	−9.9711	−375.37	−251.40	−168.90	−53.07
F12 AVQZ/non-sc	393.7490	215.0661	107.3860	−9.3133	−375.10	−251.25	−168.82	−53.08
F12 AVQZ + Constant BF/sc	388.2442	210.7087	103.9208	−11.1535	−375.19	−251.20	−168.70	−52.91
F12 AVQZ + Constant BF/non-sc	390.4102	212.3960	105.2446	−10.4932	−374.92	−251.05	−168.62	−52.92
F12 AVQZ + Variable BF/sc	388.3541	210.8207	104.0316	−11.0557	−375.17	−251.19	−168.68	−52.90
F12 AVQZ + Variable BF/non-sc	390.5355	212.5239	105.3697	−10.3873	−374.90	−251.03	−168.61	−52.91
F12 AV5Z/sc	390.6766	212.2192	105.2330	−10.3611	−374.40	−251.21	−168.75	−53.02
F12 AV5Z/non-sc	390.9067	212.9489	105.7967	−10.0906	−374.72	−250.92	−168.54	−52.93
F12 AV5Z + Constant BF/sc	388.4768	210.9420	104.1417	−10.9972	−374.81	−250.94	−168.52	−52.87
F12 AV5Z + Constant BF/non-sc	389.6633	211.8644	104.8638	−10.6392	−374.67	−250.86	−168.48	−52.87
F12 AV5Z + Variable BF/sc	388.3680	210.8718	104.0987	−11.0019	−374.81	−250.94	−168.52	−52.87
F12 AV5Z + Variable BF/non-sc	389.5361	211.7843	104.8158	−10.6441	−374.66	−250.86	−168.48	−52.87
AVDZ/Non-F12	534.1813	323.2224	190.0954	25.5340	−406.04	−275.03	−186.76	−60.14
AVDZ/Non-F12 + Constant BF	406.4566	222.0908	110.5355	−10.8147	−393.76	−265.03	−178.97	−57.42
AVDZ/Non-F12 +Variable BF	451.7816	255.8371	135.7819	0.2023	−401.46	−271.27	−183.90	−59.69
AVTZ/non-F12	433.2509	244.4238	129.3956	0.4732	−383.07	−257.29	−173.34	−55.02
AVTZ/Non-F12 + Constant BF	399.9588	218.6269	109.1675	−9.9062	−381.94	−256.18	−172.31	−54.11
AVTZ/non-F12 + Variable BF	406.1139	223.3212	112.7496	−8.2446	−381.63	−255.89	−172.07	−54.03
AVQZ/non-F12	405.8791	224.1528	114.2445	−6.2141	−378.43	−253.85	−170.83	−53.93
AVQZ/non-F12 + Constant BF	394.5959	215.1312	106.9767	−10.1836	−377.57	−253.02	−170.05	−53.41
AVQZ/non-F12 + Variable BF	395.1636	215.7250	107.5478	−9.7922	−377.56	−252.99	−170.03	−53.39
AV5Z/non-F12	396.5349	217.2440	109.1019	−8.5135	−376.34	−252.19	−169.53	−53.39
AV5Z/non-F12 + Constant BF	392.2681	213.5891	105.9826	−10.3891	−376.03	−251.88	−169.23	−53.14
AV5Z/non-F12 + BF	391.7566	213.3198	105.8635	−10.3399	−376.00	−251.85	−169.21	−53.13
MRCC								
CCSD(T)/aDZ	533.5445	322.8133	189.8217	25.4222	−406.63	−275.40	−187.00	−60.20
CCSD(T)/aDZ + BF	405.5094	221.4530	110.0869	−11.0081	−394.42	−265.43	−179.21	−57.47
CCSDT/aDZ (semicanonical)	528.7875	316.2808	186.9110	23.9612	−407.25	−278.57	−187.17	−60.18
CCSDT/aDZ	528.7842	319.1050	186.9104	52.7340	−407.27	−275.76	−187.19	−61.75
CCSDT/aDZ + BF	399.3594	216.5683	106.1878	−13.0364	−394.90	−265.62	−179.24	−57.34
CCSDT/aDZ + BF (semicanonical)	399.3632	216.5708	106.1893	−13.0360	−394.88	−265.60	−179.22	−57.33
CCSDT(Q)/aDZ	529.3757	319.5294	187.3470	24.1487	−407.76	−276.12	−187.39	−60.25
CCSD(T)/aTZ + BF	399.0345	218.0155	108.7471	−10.0800	−382.60	−256.58	−172.55	−54.15
CCSDT/aTZ + BF	393.7525	213.8520	105.4423	−11.7868	−382.98	−256.69	−172.52	−54.01
CCSDT/aTZ + BF (semicanonical)	393.7551	213.8534	105.4431	−11.7861	−382.96	−256.68	−172.51	−54.01
CCSD(T)/aTZ	432.4224	243.8835	129.0300	0.3208	−383.70	−257.67	−173.57	−55.07
CCSDT/aTZ	427.0920	239.6920	125.7161	17.0015	−384.07	−257.78	−173.54	−58.82
CCSDT/aTZ (semicanonical)	427.0947	237.5718	125.7169	−1.3546	−384.05	−259.89	−173.53	−54.93
CCSD(T)/aQZ	404.9728	223.5585	113.8398	−6.3766	−379.09	−254.25	−171.06	−53.97
CCSD(T)/aQZ +BF	393.6635	214.5174	106.5566	−10.3551	−378.24	−253.41	−170.28	−53.46
CCSDT/aQZ	399.8774	219.4209	110.6630	−7.9970	−379.45	−254.22	−171.02	−53.83
CCSDT(Q)/aTZ	426.8860	260.5265	146.3559	−1.4281	−384.48	−236.99	−194.54	−51.40

Table 5.4: Comparison of CCSDT − CCSD(T) interaction energy Corrections of computed diagonal V_d and off-diagonal V_o interaction energies (cm^-1) at three values of R for C_s Symmetry .

Method/basis set	V_d				V_o			
	4.75	5.0	5.25	6.0	4.75	5.0	5.25	6.0
$\delta^{CCSDT}_{CCSD(T)}$/aDZ	−4.7570	6.5326	2.9108	1.4610	−0.62	−3.17	−0.17	0.02
$\delta^{CCSDT}_{CCSD(T)}$/aDZ + BF	−6.1463	−4.8821	−3.8977	−2.0279	−0.46	−0.18	−0.01	0.13
$\delta^{CCSDT}_{CCSD(T)}$/aTZ	−5.3277	−6.3117	−3.3131	−1.6754	−0.35	−2.22	0.04	0.14
$\delta^{CCSDT}_{CCSD(T)}$/aTZ + BF	−5.2794	−4.1621	−3.3039	−1.7061	−0.36	−0.10	0.04	0.15
$\delta^{CCSDT(Q)}_{CCSDT}$/aDZ	0.5882	3.2486	0.4361	0.1875	−0.51	2.45	−0.22	−0.07

Table 5.5: Comparison of computed diagonal V_d and off-diagonal V_o interaction energies (cm^-1) at four values of R for the C_1 Symmetry. In the table AVnZ are shorthand notations for the aug-cc-p-VnZ basis sets .

Method/basis set	V_d				V_o			
	4.75	5.0	5.25	6.0	4.75	5.0	5.25	6.0
				MRCI				
AVQZ	255.7955	94.5102	2.4307	−77.6183	−350.22	−230.24	−151.53	−44.91
MRCISD+Q/ATZ-JKFIT	265.0255	102.8886	10.1364	−70.9803	−351.86	−232.12	−153.54	−46.74
MRCISD+Q/AQZ-JKFIT	261.3965	100.1803	7.8517	−73.3770	−349.17	−229.63	−151.24	−45.10
MRCISD+Q/ATZ-JKFIT(3s3p)	264.1191	101.6365	8.7592	−71.7605	−358.57	−236.43	−156.34	−47.79
MRCISD+Q/AQZ-JKFIT(3s3p)	259.7274	98.3499	6.0465	−74.7311	−354.33	−232.73	−153.14	−45.74

Chapter 6

Conclusions

In this work, in attempting to draw insights from energy decomposition from SAPT, we created two very meaningful datasets that aim to accurately describe the nonadditive three-body interactions in trimers and, by extension, extended systems. We used high accuracy *ab initio* methods to do this. Specifically the benchmark interactions were computed at the silver standard level of theory using the compsosite CCSD(T)/CBS method. The first dataset, 3BHET, comprises heteromolecular trimers containing neutral and charged species focused on showcasing a diverse set of interactions ranging from hydrogen bonding, $\pi - \pi$ stacking, OH$-\pi$ and so forth, whereas 3BXB is exclusively fo halogen bonded systems created by adding water and methane to existing halogen bonded dimers.

Similar trends were observed across both datasets energy decomposition analysis using SAPT methods. However in two-body interactions, 3BHET showed predominantly electrostatics while 3BXB showed a balance of both electrostatics and dispersion. The three-body nonadditive interactions showed similar patterns with nonadditive induction dominating and being either attractive or repulsive. Nonadditive exchange was always attractive while dispersion repulsive. The highest intermolecular interaction energies were observed in charged systems for 3BHET and while the systems with dihalogens in 3BXB showed the highest both two and three-body interactions. By evaluating a variety of methods, SAPT2+(CCD)δMP2 was found to be the most accurate in recovering both two and three-body interactions.

To balance cost and accuracy, we introduced a range of hybrid methods by mixing components from different sources. This resulted in considerably better predictions of both the nonadditive and total three body interactions for a fraction of the cost. Simultaneously, the OH-H_2 PES constructed to accurately describe rotational and vibrational transitions for calculation of scattering

cross-sections serves as a platform for ML potential training and testing which requires accurate data. Overall the datasets presented in this work span a range of interactions offering a robust testbed for testing and validating approximate methods.

6.1. Future Work

The chemical space is broad and would benefit from the population of various diverse datasets. Especially since theory is moving towards larger systems and condensed phase, it is prudent to have datasets that are low cost and high accuracy. With the hybrid approaches highlighted here, the first step is the extension to systems with more than fifty atoms at relatively more accurate approximations at lower computational cost. Another direction is the development of ML many-body potentials to describe molecular interactions in extended systems.

Bibliography

[1] Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. *Chemical Reviews* **2016**, *116*, 2478–2601.

[2] Domagała, M.; Lutyńska, A.; Palusiak, M. *International Journal of Quantum Chemistry* **2017**, *117*, e25348.

[3] Góra, U.; Podeszwa, R.; Cencek, W.; Szalewicz, K. *The Journal of Chemical Physics* **2011**, *135*, 224102.

[4] Mahadevi, A. S.; Sastry, G. N. *Chemical Reviews* **2016**, *116*, 2775–2825.

[5] Xantheas, S. S. *The Journal of Chemical Physics* **1994**, *100*, 7523–7534.

[6] Patkowski, K. *Annual Reports in Computational Chemistry* **2017**, *13*, 3–91.

[7] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chemical Physics Letters* **1989**, *157*, 479–483.

[8] Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. *Journal of Chemical Physics* **2014**, *140*, 094106.

[9] Antony, J.; Grimme, S. *Physical Chemistry Chemical Physics* **2006**, *8*, 5287–5293.

[10] Tkatchenko, A.; Von Lilienfeld, O. A. *Physical Review B - Condensed Matter and Materials Physics* **2008**, *78*, 045116.

[11] Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *Journal of the American Chemical Society* **2010**, *132*, 6498–6506.

[12] Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Physical Chemistry Chemical Physics* **2006**, *8*, 1985–1993.

[13] Řezáč, J.; Riley, K. E.; Hobza, P. *Journal of Chemical Theory and Computation* **2011**, *7*, 2427–2438.

[14] Řezáč, J.; Riley, K. E.; Hobza, P. *Journal of Chemical Theory and Computation* **2012**, *8*, 4285–4292.

[15] Grimme, S. *Chemistry âĂŞ A European Journal* **2012**, *18*, 9955–9964.

[16] Sparrow, Z. M.; Ernst, B. G.; Joo, P. T.; Lao, K. U.; DiStasioJr., R. A. *The Journal of Chemical Physics* **2021**, *155*, 184303.

[17] Řezáč, J. *Journal of Chemical Theory and Computation* **2020**, *16*, 2355–2368.

[18] Řezáč, J. *Journal of Chemical Theory and Computation* **2020**, *16*, 6305–6316.

[19] Kříž, K.; Nováček, M.; Řezáč, J. *Journal of Chemical Theory and Computation* **2021**, *17*, 1548–1561.

[20] Kříž, K.; Řezáč, J. *Physical Chemistry Chemical Physics* **2022**, *24*, 14794–14804.

[21] Řezáč, J. *Physical Chemistry Chemical Physics* **2022**, *24*, 14780–14793.

[22] Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J. P.; Ren, P. *Journal of Chemical Theory and Computation* **2015**, *11*, 2609–2618.

[23] Al-Hamdani, Y. S.; Tkatchenko, A. *The Journal of Chemical Physics* **2019**, *150*, 010901–010923.

[24] Řezáč, J.; Huang, Y.; Hobza, P.; Beran, G. J. *Journal of Chemical Theory and Computation* **2015**, *11*, 3065–3079.

[25] Alkan, M.; Xu, P.; Gordon, M. S. *Journal of Physical Chemistry A* **2019**, *123*, 8406–8416.

[26] Low, K.; Coote, M. L.; Izgorodina, E. I. *Journal of Chemical Theory and Computation* **2023**, *19*, 1466–1475.

[27] Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. *Chemical Reviews* **2016**, *116*, 2478–2601.

[28] Nunes, R. S.; Vila-Viçosa, D.; Costa, P. J. *Journal of the American Chemical Society* **2021**, *143*, 4253–4267.

[29] Wolters, L. P.; Schyman, P.; Pavan, M. J.; Jorgensen, W. L.; Bickelhaupt, F. M.; Kozuch, S. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 523–540.

[30] Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chemical Reviews* **1994**, *94*, 1887–1930.

[31] Patkowski, K. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2020**, *10*, 1–47.

[32] Ochieng, S. A.; Patkowski, K. *Physical Chemistry Chemical Physics* **2023**, *25*, 28621–28637.

[33] Price, P.; Bottorff, B.; Jenkins, J.; Brune, W. H.; Stevens, P. S. *Communications Earth and Environment* **2025**, *6*, 1–8.

[34] Marinakis, S.; Kalugina, Y.; Kos, J.; Lique, F. *Astronomy & Astrophysics* **2019**, *629*, A130.

[35] Kirste, M.; Scharfenberg, L.; Kłos, J.; Lique, F.; Alexander, M. H.; Meijer, G.; Van De Meerakker, S. Y. *Physical Review A - Atomic, Molecular, and Optical Physics* **2010**, *82*, 042717.

[36] Chen, J.; Xu, X.; Zhang, D. H. *Journal of Chemical Physics* **2013**, *138*, 154301.

[37] Dagdigian, P. J. *Journal of Chemical Physics* **2022**, *157*, 104305.

[38] Ma, Q.; Kłos, J.; Alexander, M. H.; Van Der Avoird, A.; Dagdigian, P. J. *The Journal of Chemical Physics* **2014**, *141*, 174309.

[39] Langhoff, S. R.; Davidson, E. R. *International Journal of Quantum Chemistry* **1974**, *8*, 61–72.

[40] Szabo, A.; Ostlund, N. *Modern quantum chemistry: introduction to advanced electronic structure theory*; 1996.

[41] Born, M.; Scientific, R. O. Q. C. C.; 2000, u. *World Scientific* **2000**, 1–24.

[42] Eisberg, R.; Resnick, R. *Quantum physics of atoms, molecules, solids, nuclei, and particles*; 1985.

[43] Magnasco, V. *Elementary Methods of Molecular Quantum Mechanics* **2007**, 255–361.

[44] Review, J. S. P.; 1930, u. *APS* **1930**, *35*, 210–211.

[45] Pacati, F. D.; Boffi, S. *Physical Review C* **1970**, *2*, 1205.

[46] Roothaan, C. C. *Reviews of Modern Physics* **1951**, *23*, 69.

[47] Hofmann, M.; Schaefer, H. F. *Encyclopedia of Physical Science and Technology* **2003**, 487–506.

[48] Patkowski, K. *Annual Reports in Computational Chemistry* **2017**, *13*, 3–91.

[49] Šimová, L.; Řezáč, J.; Hobza, P. *Journal of Chemical Theory and Computation* **2013**, *9*, 3420–3428.

[50] Gier, d.; Dorothea, H. *research.rug.nlHD de Gierresearch.rug.nl*

[51] Pople, J. A.; Binkley, J. S.; Seeger, R. *Wiley Online LibraryJA Pople, JS Binkley, R SeegerInternational Journal of Quantum Chemistry, 1976âĂćWiley Online Library* **1976**, *10*, 1–19.

[52] Manni, G. L.; Guther, K.; Ma, D.; Dobrautz, W. *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* **2020**, 132–203.

[53] Plasser, F.; Lischka, H. *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* **2020**, 277–297.

[54] Szalewicz, K. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 254–272.

[55] Korona, T. *Molecular Physics* **2013**, *111*, 3705–3715.

[56] Heßelmann, A.; Jansen, G.; Schütz, M. *The Journal of Chemical Physics* **2004**, *122*, 014103.

[57] Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *Journal of Chemical Physics* **2005**, *123*, 214103.

[58] Podeszwa, R.; Szalewicz, K. *Journal of Chemical Physics* **2007**, *126*, 194101–194114.

[59] Xie, W.; Song, L.; Truhlar, D. G.; Gâo, J. *The Journal of Chemical Physics* **2008**, *128*, 234108.

[60] Lao, K. U.; Herbert, J. M. *J. Phys. Chem. A* **2015**, *119*, 2024.

[61] Xu, P.; Alkan, M.; Gordon, M. S. *Chemical Reviews* **2020**, *120*, 12343–12356.

[62] Sengupta, A.; Li, B.; Svatunek, D.; Liu, F.; Houk, K. N. *Accounts of Chemical Research* **2022**, *55*, 2467–2479.

[63] Burns, L. A.; Faver, J. C.; Zheng, Z.; Marshall, M. S.; Smith, D. G.; Vanommeslaeghe, K.; MacKerell, A. D.; Merz, K. M.; Sherrill, C. D. *Journal of Chemical Physics* **2017**, *147*, 161727.

[64] Lao, K. U.; Herbert, J. M. *Journal of Chemical Theory and Computation* **2018**, *14*, 5128–5142.

[65] Lao, K. U.; Herbert, J. M. *Journal of Physical Chemistry A* **2015**, *119*, 235–252.

[66] Carter-Fenk, K.; Lao, K. U.; Herbert, J. M. *Accounts of Chemical Research* **2021**, *54*, 3679–3690.

[67] Lao, K. U.; Herbert, J. M. *Journal of Chemical Theory and Computation* **2018**, *14*, 5128–5142.

[68] Carter-Fenk, K.; Lao, K. U.; Liu, K. Y.; Herbert, J. M. *Journal of Physical Chemistry Letters* **2019**, *10*, 2706–2714.

[69] Metcalf, D. P.; Koutsoukas, A.; Spronk, S. A.; Claus, B. L.; Loughney, D. A.; Johnson, S. R.; Cheney, D. L.; Sherrill, C. D. *The Journal of Chemical Physics* **2020**, *152*, 074103.

[70] Boys, S. F.; Bernardi, F. *MOLECULAR PHYSICS* **1970**, *19*, 553–566.

[71] Pitman, S. J.; Evans, A. K.; Ireland, R. T.; Lempriere, F.; McKemmish, L. K. *Journal of Physical Chemistry A* **2023**, *127*, 10295–10306.

[72] Jensen, F. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 273–295.

[73] Hill, J. G. *International Journal of Quantum Chemistry* **2013**, *113*, 21–34.

[74] Nagy, B.; Jensen, F. **2017**, 93–149.

[75] Woon, D.; physics, T. D. J. T. J. o. c.; 1994, u. *pubs.aip.org* **1994**, *100*, 2975–2988.

[76] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *The Journal of Chemical Physics* **1992**, *96*, 6796–6806.

[77] Beran, G. J. *Chemical Reviews* **2016**, *116*, 5567–5613.

[78] Hapka, M.; Rajchel, Å.; Modrzejewski, M.; Schäffer, R.; Chałasiński, G.; SzczÈl'ś-niak, M. M. *Journal of Chemical Physics* **2017**, *147*, 084106–084117.

[79] Merritt, J. M.; Bondybey, V. E.; Heaven, M. C. *Science* **2009**, *324*, 1548–1551.

[80] Jankowski, P.; McKellar, A. R.; Szalewicz, K. *Science* **2012**, *336*, 1147–1150.

[81] Řezáč, J.; Hobza, P. *Chemical Reviews* **2016**, *116*, 5038–5071.

[82] Smith, D. G.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *Journal of Physical Chemistry Letters* **2016**, *7*, 2197–2203.

[83] Sparrow, Z. M.; Ernst, B. G.; Joo, P. T.; Lao, K. U.; DiStasioJr., R. A. *The Journal of Chemical Physics* **2021**, *155*, 184303.

[84] Babin, V.; Medders, G. R.; Paesani, F. *Journal of Chemical Theory and Computation* **2014**, *10*, 1599–1607.

[85] Kennedy, M. R.; McDonald, A. R.; Deprince, A. E.; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. *Journal of Chemical Physics* **2014**, *140*, 121104.

[86] Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K. L. *Science* **2014**, *345*, 640–643.

[87] Tan, S.; Barrera Acevedo, S.; Izgorodina, E. I. *Journal of Chemical Physics* **2017**, *146*, 064108.

[88] Jankiewicz, W.; Podeszwa, R.; Witek, H. A. *Journal of Chemical Theory and Computation* **2018**, *14*, 5079–5089.

[89] Tkatchenko, A.; Distasio, R. A.; Car, R.; Scheffler, M. *Physical Review Letters* **2012**, *108*, 236402.

[90] Singh, N. J.; Min, S. K.; Kim, D. Y.; Kim, K. S. *Journal of Chemical Theory and Computation* **2009**, *5*, 515–529.

[91] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 242–253.

[92] Smith, D. G. et al. *Journal of Chemical Physics* **2020**, *152*, 84108–84129.

[93] Bukowski, R. et al. **2009**, 02–093.

[94] Aidas, K. et al. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 269–284.

[95] Shao, Y. et al. *Molecular Physics* **2015**, *113*, 184–215.

[96] Burns, L. A.; Marshall, M. S.; Sherrill, C. D. *Journal of Chemical Physics* **2014**, *141*, 234111.

[97] Kodrycka, M.; Patkowski, K. *Journal of Chemical Physics* **2019**, *151*, 070901.

[98] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chemical Physics Letters* **1998**, *286*, 243–252.

[99] Misquitta, A. J.; Szalewicz, K. *The Journal of Chemical Physics* **2005**, *122*, 214109–214128.

[100] Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *Journal of Chemical Information and Modeling* **2007**, *47*, 1045–1052.

[101] Tozer, D. J.; Handy, N. C. *The Journal of Chemical Physics* **1998**, *109*, 10180–10189.

[102] Misquitta, A. J.; Szalewicz, K. *Chemical Physics Letters* **2002**, *357*, 301–306.

[103] Grüning, M.; Gritsenko, O. V.; Stan, .; Van Gisbergen, J. A.; Evert, .; Baerends, J.; Grü, M.; Van Gisbergen, S. J. A.; Baerends, E. J. *The Journal of Chemical Physics* **2002**, *116*, 9591–9601.

[104] Xie, Y.; Smith, D. G.; Sherrill, C. D. *Journal of Chemical Physics* **2022**, *157*, 24801.

[105] Van Leeuwen, R.; Baerends, E. J. *Physical Review A* **1994**, *49*, 2421.

[106] Lotrich, V. F.; Szalewicz, K. *The Journal of Chemical Physics* **1997**, *106*, 9668–9687.

[107] Patkowski, K.; Szalewicz, K.; Jeziorski, B. *Journal of Chemical Physics* **2006**, *125*, 154107.

[108] Gray, M.; Herbert, J. M. *Journal of Chemical Theory and Computation* **2022**, *18*, 2308–2330.

[109] Herbert, J. M.; Jacobson, L. D.; Un Lao, K.; Rohrdanz, M. A. *Physical Chemistry Chemical Physics* **2012**, *14*, 7679–7699.

[110] Lao, K. U.; Herbert, J. M. *Journal of Chemical Theory and Computation* **2018**, *14*, 2955–2978.

[111] Garcia, J.; Podeszwa, R.; Szalewicz, K. *Journal of Chemical Physics* **2020**, *152*, 184109.

[112] Huang, Y.; Beran, G. J. *Journal of Chemical Physics* **2015**, *143*, 044113.

[113] Nguyen, B. D.; Hernandez, D. J.; Flores, E. V.; Furche, F. *Electronic Structure* **2022**, *4*, 014003.

[114] Temelso, B.; Archer, K. A.; Shields, G. C. *Journal of Physical Chemistry A* **2011**, *115*, 12034–12046.

[115] Haberhauer, G.; Gleiter, R. *Angewandte Chemie - International Edition* **2020**, *59*, 21236–21243.

[116] Mahmudov, K. T.; Gurbanov, A. V.; Aliyeva, V. A.; Guedes da Silva, M. F. C.; Resnati, G.; Pombeiro, A. J. *Coordination Chemistry Reviews* **2022**, *464*, 214556.

[117] Stone, A. J. *Journal of the American Chemical Society* **2013**, *135*, 7005–7009.

[118] Davis, H. J.; Phipps, R. J. *Chemical Science* **2017**, *8*, 864–877.

[119] Politzer, P.; Murray, J. S.; Clark, T.; Resnati, G. *Physical Chemistry Chemical Physics* **2017**, *19*, 32166–32178.

[120] Metrangolo, P.; Neukirch, H.; Pilati, T.; Resnati, G. **2005**,

[121] Baldrighi, M.; Cavallo, G.; Chierotti, M. R.; Gobetto, R.; Metrangolo, P.; Pilati, T.; Resnati, G.; Terraneo, G. *Molecular Pharmaceutics* **2013**, *10*, 1760–1772.

[122] Bertolani, A.; Pirrie, L.; Stefan, L.; Houbenov, N.; Haataja, J. S.; Catalano, L.; Terraneo, G.; Giancane, G.; Valli, L.; Milani, R.; Ikkala, O.; Resnati, G.; Metrangolo, P. *Nature Communications 2015 6:1* **2015**, *6*, 1–9.

[123] Williams, H. L.; Szalewicz, K.; Moszynski, R.; Jeziorski, B. *The Journal of Chemical Physics* **1995**, *103*, 4586–4599.

[124] Heßelmann, A.; Jansen, G.; Schütz, M. *Journal of Chemical Physics* **2005**, *122*, 14103.

[125] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *Journal of Chemical Physics* **2010**, *132*, 154104.

[126] Xie, Y.; Glick, Z. L.; Sherrill, C. D. *Journal of Chemical Physics* **2023**, *158*, 94110.

[127] Gatica, S. M.; Calbi, M. M.; Cole, M. W.; Velegol, D. *Physical Review B* **2003**, *68*, 205409.

[128] Anatole Von Lilienfeld, O.; Tkatchenko, A. *Journal of Chemical Physics* **2010**, *132*, 234109.

[129] Tang, K. T.; Peter Toennies, J. *The Journal of Chemical Physics* **1984**, *80*, 3726–3741.

[130] Řezáč, J.; Riley, K. E.; Hobza, P. *Journal of Chemical Theory and Computation* **2012**, *8*, 4285–4292.

[131] Dunning, T. H. *Journal of Physical Chemistry A* **2000**, *104*, 9062–9080.

[132] Lamberts, T.; Fedoseev, G.; Kästner, J.; Ioppolo, S.; Linnartz, H. *Astronomy & Astrophysics* **2017**, *599*, A132.

[133] Wampfler, S. F.; Bruderer, S.; Karska, A.; Herczeg, G. J.; Van Dishoeck, E. F.; Kristensen, L. E.; Goicoechea, J. R.; Benz, A. O.; Doty, S. D.; McCoey, C.; Baudry, A.; Giannini, T.; Larsson, B. *Astronomy & Astrophysics* **2013**, *552*, A56.

[134] Alexander, M. H.; Rackham, E. J.; Manolopoulos, D. E. *The Journal of Chemical Physics* **2004**, *121*, 5221–5235.

[135] Ma, Q.; Kłos, J.; Alexander, M. H.; Van Der Avoird, A.; Dagdigian, P. J. *The Journal of Chemical Physics* **2014**, *141*, 174309.

[136] Thibault, F.; Patkowski, K.; Żuchowski, P. S.; Jóźwiak, H.; Ciuryło, R.; Wcisło, P. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2017**, *202*, 308–320.

[137] Bakr, B. W.; Smith, D. G.; Patkowski, K. *The Journal of Chemical Physics* **2013**, *139*, 144305.

[138] Tao, F. M.; Pan, Y. K. *The Journal of Chemical Physics* **1992**, *97*, 4989–4995.

[139] Tao, F. M. *The Journal of Chemical Physics* **1993**, *98*, 2481–2483.

[140] Tao, F.-M.; Chem Phys, J. *The Journal of Chemical Physics* **1993**, *98*, 3049–3059.

[141] Tao, F.-M.; Chem Phys, J. *The Journal of Chemical Physics* **1994**, *100*, 4947–4954.

[142] Patkowski, K. *THE JOURNAL OF CHEMICAL PHYSICS* **2013**, *138*, 154101.

[143] Pandharkar, R.; Hermes, M. R.; Cramer, C. J.; Gagliardi, L. *Journal of Chemical Theory and Computation* **2022**, *18*, 6557–6566.

Appendices

Appendix A

Supporting Information for Accurate three-body noncovalent interactions: the insights from energy decomposition

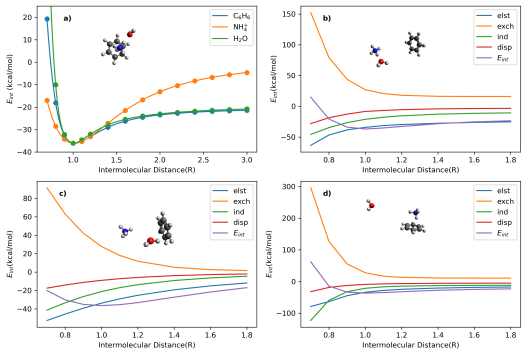


Figure 1: **a)** Radial benchmark potential energy curves (in kcal/mol) for the ammonium-benzene-water complex 4.1. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting C_6H_6 (**b**), NH_4^+ (**c**), and H_2O (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

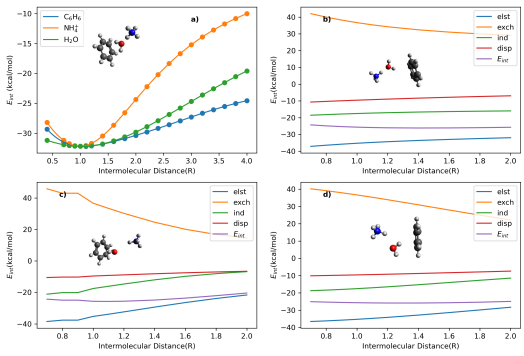


Figure 2: **a)** Radial benchmark potential energy curves (in kcal/mol) for the ammonium-benzene-water complex 4.2. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting C_6H_6 (**b**), NH_4^+ (**c**), and H_2O (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

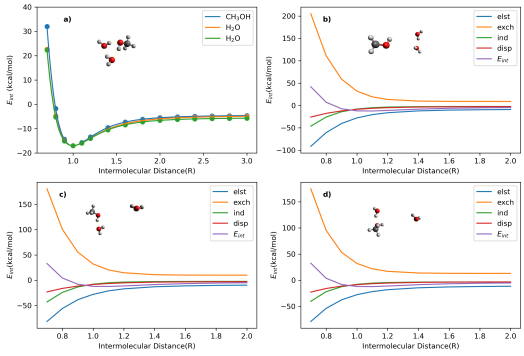


Figure 3: **a)** Radial benchmark potential energy curves (in kcal/mol) for the water-water-methanol complex 3.1. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting CH_3OH (**b**), H_2O (**c**), and H_2O (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

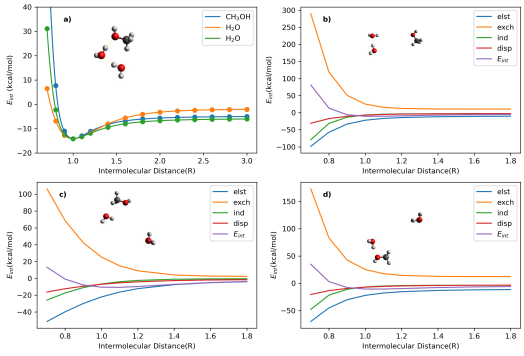


Figure 4: **a)** Radial benchmark potential energy curves (in kcal/mol) for the water-water-methanol complex 3.2. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting CH_3OH (**b**), H_2O (**c**), and H_2O (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

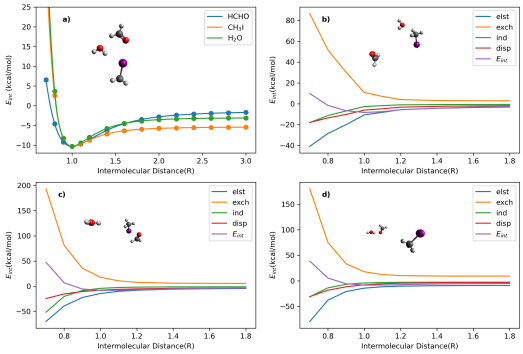


Figure 5: **a)** Radial benchmark potential energy curves (in kcal/mol) for the iodomethane-formaldehyde-water complex 1.1. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting HCHO (**b**), H_2O (**c**), and CH_3I (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

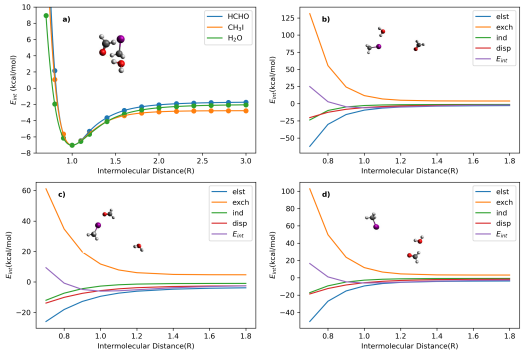


Figure 6: **a)** Radial benchmark potential energy curves (in kcal/mol) for the iodomethane-formaldehyde-water complex 1.2. **b)–d)** Two-body SAPT(DFT) energy decompositions along the curves obtained by shifting HCHO (**b**), H_2O (**c**), and CH_3I (**d**) relative to the center of mass of the entire complex. The relative geometry of the other two molecules stays unchanged for each panel.

Appendix B

Supporting Information for Influence of Three-Body Effects on Halogen Bonding

Table 1: Benchmark two-body and three-body interaction energy data for $Cl-H_2O$ systems(kcal mol^{-1}). The CCSD(T)/CBS estimate was obtained in a composite manner according to 2.57

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
1.1.01	−6.644	−0.033	−6.677	−3.223	−5.816	0.015	−5.800	−2.539
1.1.02	−8.205	−0.022	−8.226	−4.488	−6.462	0.039	−6.423	−2.984
1.2.01	−8.910	−0.008	−8.919	−4.019	−8.134	−0.007	−8.141	−3.360
1.2.02	−7.304	0.151	−7.153	−3.165	−6.720	0.220	−6.500	−2.899
1.2.03	−8.990	0.279	−8.710	−1.584	−8.489	0.334	−8.155	−1.387
1.2.04	−5.815	−0.348	−6.163	−2.102	−5.348	−0.301	−5.649	−1.858
1.2.05	−11.516	−1.276	−12.793	−7.708	−9.578	−1.235	−10.813	−5.873
1.2.06	−6.107	0.186	−5.920	−1.961	−5.746	0.209	−5.537	−1.705
1.2.07	−7.250	0.032	−7.218	−2.564	−6.871	0.084	−6.787	−2.263
1.2.08	−21.814	−2.937	−24.751	−18.332	−17.966	−2.734	−20.700	−14.563
1.3.01	−4.939	−0.465	−5.404	−2.907	−5.054	−0.443	−5.496	−2.875
1.3.02	−7.230	−1.006	−8.235	−4.915	−6.589	−0.981	−7.571	−4.260
1.3.03	−5.225	0.108	−5.116	−2.027	−4.845	0.122	−4.723	−1.925
1.3.04	−5.756	0.168	−5.588	−1.642	−5.666	0.221	−5.446	−1.467
1.3.05	−6.451	0.046	−6.405	−3.111	−6.128	0.122	−6.006	−2.636
1.3.06	−6.036	0.144	−5.892	−2.139	−6.163	0.166	−5.997	−2.142
1.3.07	−7.278	0.178	−7.101	−3.173	−7.005	0.203	−6.802	−2.819
1.3.08	−13.096	0.161	−12.935	−4.413	−12.941	0.245	−12.696	−4.129
1.3.09	−10.489	0.295	−10.194	−1.048	−10.550	0.346	−10.203	−0.929
1.3.10	−17.251	0.469	−16.782	−8.603	−15.714	0.553	−15.161	−6.954
1.4.01	−5.808	−0.281	−6.089	−3.401	−5.490	−0.245	−5.735	−2.958
1.4.02	−7.343	0.227	−7.116	−2.321	−6.963	0.277	−6.686	−1.979
1.4.03	−16.113	−1.967	−18.080	−12.608	−12.503	−1.781	−14.283	−9.100
1.4.04	−4.457	−0.356	−4.814	−2.628	−4.425	−0.335	−4.760	−2.383
1.4.05	−12.588	−2.979	−15.567	−10.399	−9.585	−2.873	−12.458	−7.366
1.5.01	−2.299	−0.058	−2.356	−0.738	−2.208	−0.052	−2.260	−0.660
1.6.01	−4.163	−0.147	−4.310	−2.009	−3.570	−0.118	−3.688	−1.507
1.7.01	−4.711	−0.177	−4.888	−2.411	−3.946	−0.141	−4.087	−1.711
1.8.01	−5.359	−0.322	−5.681	−3.018	−4.254	−0.283	−4.537	−2.017

Table 2: Benchmark two-body and three-body interaction energy data for $Cl-CH_4$ systems (kcal mol^{-1}). The CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. 2.57).

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
1.1.01	−4.960	−0.038	−4.997	−3.223	−4.104	0.016	−4.088	−2.539
1.1.02	−6.298	0.014	−6.283	−4.488	−4.468	0.049	−4.419	−2.984
1.2.01	−4.697	−0.029	−4.725	−4.019	−4.055	−0.033	−4.088	−3.328
1.2.02	−5.198	0.016	−5.182	−3.164	−4.420	0.037	−4.383	−2.894
1.2.03	−3.078	−0.032	−3.110	−1.584	−2.664	0.026	−2.637	−1.385
1.2.04	−4.530	0.009	−4.521	−2.102	−3.747	0.038	−3.708	−1.855
1.2.05	−9.850	−0.047	−9.898	−7.707	−7.462	−0.010	−7.472	−5.870
1.2.06	−3.360	0.036	−3.324	−1.961	−3.083	0.061	−3.022	−1.703
1.2.07	−4.775	−0.020	−4.795	−2.564	−3.868	−0.009	−3.877	−2.257
1.2.08	−19.494	−0.372	−19.865	−18.334	−15.688	−0.358	−16.045	−14.561
1.3.01	−3.815	−0.077	−3.892	−2.907	−3.877	−0.053	−3.930	−2.876
1.3.02	−5.792	0.024	−5.768	−4.913	−5.108	0.059	−5.049	−4.246
1.3.03	−4.054	0.027	−4.027	−2.028	−3.661	0.073	−3.588	−1.928
1.3.04	−3.625	−0.010	−3.636	−1.641	−3.150	0.026	−3.123	−1.464
1.3.05	−5.040	0.022	−5.019	−3.110	−4.252	0.083	−4.169	−2.630
1.3.06	−3.232	−0.006	−3.238	−2.139	−3.233	0.022	−3.210	−2.131
1.3.07	−4.571	−0.001	−4.572	−3.173	−4.116	0.021	−4.094	−2.800
1.3.08	−6.952	0.077	−6.875	−4.411	−6.225	0.114	−6.111	−4.126
1.3.09	−3.769	0.009	−3.760	−1.048	−3.084	0.023	−3.061	−0.921
1.3.10	−11.262	0.050	−11.212	−8.604	−9.079	0.056	−9.023	−6.955
1.4.01	−4.679	0.013	−4.667	−3.401	−4.276	0.057	−4.219	−2.973
1.4.02	−3.864	−0.024	−3.888	−2.320	−3.466	0.034	−3.432	−1.988
1.4.03	−14.229	−0.198	−14.427	−12.608	−10.899	−0.150	−11.048	−9.332
1.4.04	−4.386	−0.034	−4.421	−2.629	−4.043	0.040	−4.003	−2.406
1.4.05	−12.126	−0.299	−12.424	−10.398	−9.079	−0.219	−9.298	−7.631
1.5.01	−1.720	0.002	−1.718	−0.737	−1.512	0.005	−1.507	−0.633
1.6.01	−3.287	0.007	−3.279	−2.010	−2.649	0.053	−2.596	−1.507
1.7.01	−3.442	−0.015	−3.457	−2.410	−2.659	−0.011	−2.670	−1.732
1.8.01	−4.095	−0.036	−4.131	−3.019	−2.973	−0.035	−3.008	−2.014

Table 3: Benchmark two-body and three-body interaction energy data for $Br-H_2O$) systems (kcal mol^{-1}). The CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. 2.57)

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
2.1.01	−8.758	−0.584	−9.342	−5.802	−6.750	−0.486	−7.236	−3.861
2.1.02	−7.255	−0.085	−7.339	−3.944	−6.240	−0.032	−6.272	−3.089
2.1.03	−6.254	−0.141	−6.395	−2.589	−5.340	−0.098	−5.438	−1.788
2.2.01	−13.971	−0.785	−14.756	−10.372	−11.357	−0.680	−12.037	−8.097
2.2.02	−7.418	0.333	−7.085	−2.788	−6.727	0.417	−6.310	−2.429
2.2.03	−8.284	0.151	−8.133	−4.597	−7.463	0.225	−7.238	−4.128
2.2.04	−8.243	0.338	−7.905	−2.439	−7.610	0.403	−7.207	−2.082
2.2.05	−23.248	−2.642	−25.890	−19.725	−19.323	−2.432	−21.755	−15.929
2.2.06	−8.030	−0.136	−8.166	−4.246	−7.418	−0.097	−7.514	−3.703
2.2.07	−8.019	−0.189	−8.208	−3.569	−7.412	−0.151	−7.564	−3.043
2.3.01	−8.902	−1.202	−10.104	−6.479	−7.982	−1.165	−9.147	−5.581
2.3.02	−5.494	−0.232	−5.726	−2.323	−4.888	−0.231	−5.119	−2.137
2.3.03	−5.860	−0.572	−6.432	−3.952	−5.820	−0.542	−6.361	−3.857
2.3.04	−5.265	−0.102	−5.368	−1.806	−5.080	−0.066	−5.146	−1.636
2.3.05	−7.095	0.123	−6.973	−3.924	−6.538	0.212	−6.326	−3.288
2.3.06	−5.762	0.170	−5.592	−2.024	−5.622	0.233	−5.389	−1.806
2.3.07	−6.238	0.077	−6.161	−2.830	−5.699	0.100	−5.598	−2.509
2.3.08	−5.930	0.175	−5.755	−1.744	−5.751	0.227	−5.524	−1.510
2.3.09	−8.224	0.270	−7.954	−4.162	−7.702	0.299	−7.403	−3.627
2.3.10	−6.650	0.185	−6.465	−2.836	−6.727	0.209	−6.518	−2.810
2.3.11	−14.912	−2.127	−17.038	−11.366	−12.478	−2.008	−14.486	−9.300
2.3.12	−12.015	0.212	−11.802	−3.010	−11.911	0.219	−11.692	−2.785
2.3.13	−9.532	−0.934	−10.466	−5.991	−9.099	−0.890	−9.989	−5.599
2.3.14	−7.364	−0.224	−7.588	−3.426	−6.630	−0.194	−6.824	−2.741
2.4.01	−17.915	−1.907	−19.821	−14.401	−14.342	−1.746	−16.088	−10.930
2.4.02	−7.365	−0.723	−8.088	−4.782	−6.710	−0.661	−7.371	−4.137
2.4.03	−16.130	−1.677	−17.807	−12.468	−13.186	−1.527	−14.713	−9.637
2.4.04	−4.782	−0.154	−4.936	−3.812	−4.493	−0.142	−4.635	−3.426
2.4.05	−5.617	0.013	−5.605	−2.163	−5.190	0.052	−5.138	−1.775
2.5.01	−2.456	−0.061	−2.517	−0.835	−2.323	−0.056	−2.380	−0.724
2.6.01	−4.295	−0.085	−4.380	−2.478	−3.604	−0.074	−3.678	−1.796
2.6.02	−2.365	−0.040	−2.406	−1.717	−2.156	−0.034	−2.190	−1.424
2.7.01	−5.345	−0.289	−5.634	−3.072	−4.334	−0.255	−4.589	−2.173
2.7.02	−4.477	−0.151	−4.628	−1.998	−4.039	−0.130	−4.168	−1.643
2.8.01	−6.452	−0.475	−6.927	−4.129	−4.949	−0.425	−5.374	−2.707
2.8.02	−7.160	−0.308	−7.468	−2.329	−6.193	−0.316	−6.508	−1.866

Table 4: Benchmark two-body and three-body interaction energy data for $Br-CH_4$ systems (kcal mol^{-1}). The CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. 2.57)

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
2.1.01	−7.592	−0.034	−7.626	−5.801	−5.371	0.018	−5.353	−3.909
2.1.02	−5.553	−0.044	−5.597	−3.944	−4.521	0.009	−4.513	−3.089
2.1.03	−4.097	−0.067	−4.164	−2.588	−3.204	−0.042	−3.246	−1.947
2.2.01	−12.616	−0.064	−12.680	−10.372	−9.782	−0.013	−9.796	−8.069
2.2.02	−4.796	−0.004	−4.800	−2.787	−3.945	0.030	−3.915	−2.423
2.2.03	−6.640	0.007	−6.634	−4.596	−5.673	0.039	−5.634	−4.122
2.2.04	−4.273	0.004	−4.270	−2.437	−3.403	0.018	−3.385	−2.073
2.2.05	−21.097	−0.195	−21.292	−19.725	−17.446	−0.141	−17.588	−16.070
2.2.06	−6.512	−0.024	−6.536	−4.246	−5.338	−0.016	−5.354	−3.696
2.2.07	−4.703	0.012	−4.691	−3.569	−4.187	0.052	−4.135	−3.038
2.3.01	−8.273	−0.013	−8.286	−6.477	−7.330	0.039	−7.291	−5.615
2.3.02	−4.042	−0.021	−4.063	−2.324	−3.821	0.053	−3.768	−2.168
2.3.03	−5.705	0.037	−5.668	−3.952	−5.574	0.099	−5.475	−3.859
2.3.04	−3.082	−0.017	−3.100	−1.805	−2.912	0.020	−2.893	−1.644
2.3.05	−5.989	0.021	−5.967	−3.923	−5.001	0.095	−4.906	−3.287
2.3.06	−4.341	−0.001	−4.341	−2.023	−3.491	0.011	−3.479	−1.797
2.3.07	−4.486	0.053	−4.433	−2.830	−3.964	0.091	−3.873	−2.510
2.3.08	−3.649	−0.003	−3.652	−1.743	−3.109	0.027	−3.081	−1.515
2.3.09	−5.232	−0.023	−5.255	−4.161	−4.658	−0.013	−4.671	−3.655
2.3.10	−4.703	0.009	−4.695	−2.836	−4.553	0.033	−4.521	−2.789
2.3.11	−13.885	−0.004	−13.888	−11.366	−11.314	−0.002	−11.316	−9.300
2.3.12	−5.916	0.025	−5.891	−3.010	−5.167	0.067	−5.100	−2.787
2.3.13	−8.558	0.144	−8.414	−5.991	−7.694	0.175	−7.519	−5.601
2.3.14	−5.757	0.070	−5.687	−3.427	−4.711	0.102	−4.609	−2.740
2.4.01	−15.827	−0.283	−16.110	−14.402	−12.320	−0.223	−12.542	−10.928
2.4.02	−6.123	−0.003	−6.127	−4.781	−5.493	0.045	−5.448	−4.135
2.4.03	−13.728	−0.406	−14.134	−12.471	−10.642	−0.374	−11.016	−9.414
2.4.04	−4.872	−0.086	−4.958	−3.812	−4.565	−0.060	−4.624	−3.426
2.4.05	−4.023	−0.038	−4.060	−2.163	−3.413	0.021	−3.391	−1.752
2.5.01	−1.672	0.012	−1.661	−0.835	−1.511	0.021	−1.490	−0.725
2.6.01	−3.564	−0.026	−3.590	−2.477	−2.795	−0.007	−2.802	−1.821
2.6.02	−2.796	−0.030	−2.827	−1.717	−2.452	−0.004	−2.456	−1.430
2.7.01	−4.467	0.017	−4.450	−3.072	−3.384	0.032	−3.352	−2.141
2.7.02	−3.189	−0.042	−3.231	−1.997	−2.753	−0.019	−2.772	−1.638
2.8.01	−5.654	−0.040	−5.693	−4.130	−4.057	0.034	−4.023	−2.711
2.8.02	−3.743	−0.006	−3.750	−2.332	−3.197	0.036	−3.161	−1.877

Table 5: Benchmark two-body and three-body interaction energy data for $I-H_2O$) systems (kcal mol^{-1}). The CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. 2.57)

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
3.1.01	−8.294	−0.194	−8.488	−5.109	−7.025	−0.143	−7.168	−4.008
3.1.02	−7.485	−0.131	−7.616	−3.597	−6.308	−0.076	−6.384	−2.544
3.1.03	−9.898	0.294	−9.604	−6.528	−7.553	0.263	−7.290	−4.526
3.2.01	−11.251	−0.080	−11.331	−6.409	−9.938	−0.084	−10.023	−5.252
3.2.02	−6.862	0.275	−6.587	−2.812	−6.452	0.277	−6.175	−2.435
3.2.03	−10.700	−0.841	−11.541	−7.136	−9.792	−0.824	−10.616	−6.313
3.2.04	−8.410	0.254	−8.156	−4.011	−7.396	0.326	−7.071	−3.332
3.2.05	−14.653	−0.943	−15.596	−11.152	−12.013	−0.873	−12.886	−8.808
3.2.06	−8.917	−0.493	−9.410	−4.565	−8.138	−0.471	−8.609	−3.918
3.2.07	−21.541	−2.129	−23.670	−18.074	−18.254	−1.977	−20.231	−14.899
3.2.08	−10.911	−0.575	−11.486	−6.901	−9.948	−0.516	−10.464	−5.996
3.3.01	−9.936	0.291	−9.645	−3.044	−9.627	0.301	−9.326	−2.727
3.3.02	−13.782	0.604	−13.178	−7.328	−12.875	0.618	−12.257	−6.391
3.3.03	−10.385	0.341	−10.044	−3.543	−10.145	0.351	−9.794	−3.289
3.3.04	−4.810	−0.197	−5.007	−3.584	−4.626	−0.196	−4.823	−3.340
3.3.05	−6.306	−0.121	−6.427	−2.370	−5.926	−0.103	−6.029	−2.021
3.3.06	−7.975	−0.478	−8.454	−4.454	−7.188	−0.449	−7.637	−3.735
3.3.07	−6.178	−0.266	−6.444	−2.536	−5.874	−0.248	−6.121	−2.255
3.3.08	−7.505	0.264	−7.241	−4.015	−7.405	0.291	−7.114	−3.900
3.3.09	−6.458	0.036	−6.422	−2.904	−6.225	0.051	−6.175	−2.588
3.3.10	−8.816	0.392	−8.423	−4.832	−8.209	0.424	−7.786	−4.216
3.3.11	−6.579	−0.187	−6.765	−3.131	−5.952	−0.173	−6.125	−2.869
3.3.12	−11.906	−0.085	−11.992	−8.874	−11.068	−0.078	−11.145	−8.315
3.3.13	−9.086	−0.653	−9.739	−5.156	−8.026	−0.620	−8.646	−4.348
3.3.14	−19.363	0.617	−18.746	−10.949	−18.036	0.692	−17.344	−9.608
3.3.15	−12.628	0.421	−12.206	−4.782	−12.268	0.510	−11.758	−4.451
3.4.01	−10.229	−0.620	−10.849	−7.201	−9.224	−0.573	−9.797	−6.254
3.4.02	−8.444	−0.129	−8.573	−4.580	−7.556	−0.087	−7.642	−3.765
3.4.03	−16.557	−1.487	−18.043	−13.120	−13.693	−1.381	−15.075	−10.392
3.4.04	−6.954	−0.307	−7.261	−5.840	−6.420	−0.307	−6.727	−5.242
3.4.05	−6.694	0.328	−6.366	−3.425	−6.107	0.368	−5.739	−2.805
3.4.06	−13.742	−1.900	−15.643	−11.363	−11.437	−1.829	−13.266	−9.067
3.5.01	−2.236	−0.078	−2.313	−0.739	−2.148	−0.080	−2.227	−0.672
3.6.01	−4.354	−0.223	−4.577	−2.147	−3.875	−0.208	−4.082	−1.771
3.6.02	−4.841	−0.020	−4.861	−1.739	−4.423	0.000	−4.423	−1.364
3.6.03	−4.978	−0.248	−5.226	−2.734	−4.158	−0.223	−4.380	−2.020
3.7.01	−3.758	−0.073	−3.831	−2.564	−3.370	−0.071	−3.441	−2.080
3.7.02	−8.869	−1.221	−10.089	−5.890	−7.196	−1.146	−8.342	−4.293
3.7.03	−5.727	−0.324	−6.051	−3.420	−4.642	−0.294	−4.936	−2.453
3.8.01	−4.289	−0.178	−4.467	−3.107	−3.716	−0.172	−3.888	−2.478
3.8.02	−6.438	0.136	−6.303	−2.425	−5.791	0.150	−5.640	−1.841
3.8.03	−6.892	−0.491	−7.383	−4.576	−5.378	−0.453	−5.832	−3.187

Table 6: Benchmark two-body and three-body interaction energy data for $I-CH_4$) systems (kcal mol^{-1}). he CCSD(T)/CBS estimate was obtained in a composite manner according to Eq. 2.57)

Complex	MP2/CBS				CCSD(T)/CBS			
	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}	E_{int}^{2}	E_{int}^{3}	E_{int}	E_{int}^{AB}
3.1.01	−6.581	−0.013	−6.594	−5.109	−5.316	0.033	−5.283	−4.008
3.1.02	−5.005	−0.081	−5.086	−3.597	−3.825	−0.040	−3.866	−2.542
3.1.03	−8.163	−0.052	−8.216	−6.530	−5.878	0.006	−5.872	−4.533
3.2.01	−7.628	0.088	−7.540	−6.410	−6.355	0.132	−6.223	−5.259
3.2.02	−3.938	−0.001	−3.939	−2.812	−3.476	0.044	−3.432	−2.437
3.2.03	−9.166	−0.013	−9.179	−7.134	−7.829	0.015	−7.814	−6.307
3.2.04	−6.095	−0.004	−6.100	−4.011	−4.896	0.028	−4.868	−3.324
3.2.05	−13.245	−0.079	−13.324	−11.152	−10.362	−0.035	−10.397	−8.800
3.2.06	−6.478	−0.008	−6.485	−4.564	−5.350	0.025	−5.325	−3.912
3.2.07	−19.392	−0.198	−19.590	−18.074	−16.210	−0.135	−16.345	−14.895
3.2.08	−8.083	−0.026	−8.109	−6.901	−7.172	0.018	−7.154	−5.991
3.3.01	−4.324	0.052	−4.272	−3.044	−4.001	0.086	−3.915	−2.730
3.3.02	−8.396	0.079	−8.317	−7.331	−7.422	0.115	−7.308	−6.397
3.3.03	−5.360	−0.017	−5.377	−3.543	−5.038	0.045	−4.993	−3.291
3.3.04	−5.595	0.054	−5.541	−3.584	−5.057	0.101	−4.957	−3.346
3.3.05	−4.149	0.017	−4.132	−2.370	−3.505	0.060	−3.445	−2.022
3.3.06	−6.332	0.051	−6.281	−4.453	−5.323	0.097	−5.226	−3.735
3.3.07	−4.051	0.010	−4.041	−2.536	−3.614	0.075	−3.538	−2.256
3.3.08	−4.933	−0.009	−4.942	−4.015	−4.798	0.024	−4.773	−3.899
3.3.09	−4.761	0.072	−4.689	−2.904	−4.312	0.133	−4.178	−2.593
3.3.10	−5.860	−0.021	−5.880	−4.833	−5.175	0.012	−5.163	−4.214
3.3.11	−5.179	0.028	−5.151	−3.131	−4.836	0.083	−4.753	−2.877
3.3.12	−11.289	0.151	−11.138	−8.874	−10.265	0.177	−10.088	−8.315
3.3.13	−7.318	0.071	−7.247	−5.157	−6.051	0.110	−5.941	−4.345
3.3.14	−14.044	0.115	−13.929	−10.953	−12.161	0.173	−11.988	−9.612
3.3.15	−7.749	0.058	−7.691	−4.781	−6.912	0.115	−6.796	−4.447
3.4.01	−8.527	0.100	−8.427	−7.201	−7.584	0.139	−7.446	−6.251
3.4.02	−5.746	0.055	−5.690	−4.581	−4.942	0.115	−4.827	−3.763
3.4.03	−14.424	−0.013	−14.436	−13.121	−11.661	0.046	−11.616	−10.388
3.4.04	−7.663	−0.024	−7.686	−5.839	−6.844	0.032	−6.812	−5.244
3.4.05	−5.248	−0.045	−5.293	−3.425	−4.414	0.016	−4.398	−2.804
3.4.06	−12.722	−0.042	−12.764	−11.363	−10.276	0.019	−10.257	−9.064
3.5.01	−1.585	0.010	−1.575	−0.739	−1.496	0.021	−1.476	−0.674
3.6.01	−3.344	0.013	−3.331	−2.146	−2.927	0.022	−2.905	−1.767
3.6.02	−2.922	−0.029	−2.952	−1.740	−2.435	−0.005	−2.440	−1.362
3.6.03	−3.804	0.007	−3.797	−2.734	−2.994	0.029	−2.966	−2.019
3.7.01	−4.054	0.016	−4.038	−2.568	−3.434	0.071	−3.363	−2.091
3.7.02	−7.359	−0.080	−7.440	−5.890	−5.663	−0.056	−5.719	−4.291
3.7.03	−4.436	−0.021	−4.457	−3.421	−3.365	−0.005	−3.370	−2.451
3.8.01	−4.065	−0.011	−4.076	−3.108	−3.388	−0.002	−3.391	−2.484
3.8.02	−3.883	−0.014	−3.897	−2.427	−3.200	0.009	−3.191	−1.841
3.8.03	−5.788	−0.056	−5.844	−4.577	−4.251	−0.036	−4.287	−3.187

Table 7: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $Cl-H_2O$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−5.539	9.223	−1.794	−7.671	−5.780
1.1.02	−7.841	14.285	−3.853	−9.094	−6.503
1.2.01	−11.097	12.399	−3.326	−6.094	−8.118
1.2.02	−9.772	13.305	−2.942	−7.443	−6.853
1.2.03	−11.908	14.822	−3.827	−7.887	−8.800
1.2.04	−5.601	8.956	−1.959	−6.904	−5.508
1.2.05	−23.482	35.757	−10.996	−11.229	−9.950
1.2.06	−7.676	11.477	−2.629	−7.024	−5.851
1.2.07	−9.798	14.180	−2.946	−8.532	−7.096
1.2.08	−43.420	69.116	−27.123	−17.458	−18.885
1.3.01	−6.240	8.912	−1.728	−5.843	−4.898
1.3.02	−11.442	18.468	−5.097	−8.405	−6.477
1.3.03	−5.259	7.964	−1.998	−5.608	−4.901
1.3.04	−5.957	8.274	−2.009	−5.915	−5.607
1.3.05	−7.893	11.473	−2.608	−6.923	−5.952
1.3.06	−6.666	7.631	−1.748	−5.259	−6.042
1.3.07	−9.055	11.561	−2.961	−6.480	−6.935
1.3.08	−18.869	23.950	−7.526	−10.285	−12.730
1.3.09	−14.917	18.962	−6.177	−8.507	−10.640
1.3.10	−28.214	41.619	−15.418	−13.787	−15.800
1.4.01	−7.368	11.838	−2.339	−7.446	−5.315
1.4.02	−8.575	12.189	−2.830	−7.788	−7.004
1.4.03	−32.818	58.510	−23.255	−15.098	−12.660
1.4.04	−5.609	9.295	−1.764	−6.048	−4.126
1.4.05	−29.176	56.179	−21.832	−14.422	−9.252
1.5.01	−1.824	3.340	−0.836	−2.870	−2.189
1.6.01	−3.628	7.199	−1.730	−5.233	−3.392
1.7.01	−4.568	9.300	−2.434	−6.088	−3.791
1.8.01	−6.734	14.191	−3.832	−7.653	−4.028

Table 8: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $Cl-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−3.296	7.807	−0.568	−7.889	−3.946
1.1.02	−5.781	13.399	−2.735	−9.338	−4.455
1.2.01	−6.176	8.921	−2.256	−4.608	−4.119
1.2.02	−6.377	10.553	−1.200	−7.330	−4.354
1.2.03	−2.472	6.152	−0.585	−5.755	−2.661
1.2.04	−3.590	8.002	−0.542	−7.613	−3.743
1.2.05	−20.097	33.744	−9.575	−11.839	−7.767
1.2.06	−4.211	8.360	−1.033	−6.234	−3.118
1.2.07	−6.082	12.273	−1.182	−8.900	−3.891
1.2.08	−38.980	63.624	−25.419	−15.677	−16.451
1.3.01	−4.661	7.434	−0.962	−5.558	−3.747
1.3.02	−9.712	16.518	−4.342	−7.531	−5.068
1.3.03	−3.647	7.119	−0.672	−6.386	−3.586
1.3.04	−2.703	6.135	−0.440	−6.172	−3.181
1.3.05	−5.151	10.360	−1.668	−7.738	−4.197
1.3.06	−2.835	4.772	−0.442	−4.626	−3.131
1.3.07	−5.580	9.638	−1.761	−6.367	−4.071
1.3.08	−7.301	13.959	−2.022	−10.578	−5.943
1.3.09	−2.398	6.710	−0.792	−6.742	−3.222
1.3.10	−17.038	32.027	−10.359	−13.695	−9.064
1.4.01	−5.418	9.977	−1.401	−7.253	−4.095
1.4.02	−3.385	7.560	−0.855	−6.782	−3.462
1.4.03	−29.311	54.852	−21.683	−14.602	−10.743
1.4.04	−5.233	10.045	−1.399	−7.208	−3.796
1.4.05	−27.961	55.218	−20.700	−15.081	−8.524
1.5.01	−1.231	3.243	−0.157	−3.300	−1.445
1.6.01	−2.769	7.175	−1.037	−5.764	−2.394
1.7.01	−3.429	8.825	−1.772	−6.077	−2.453
1.8.01	−5.614	13.541	−3.084	−7.545	−2.701

Table 9: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $Br-H_2O$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
2.1.01	−9.587	17.245	−4.992	−9.673	−7.006
2.1.02	−6.228	10.193	−2.029	−8.160	−6.223
2.1.03	−5.489	9.047	−1.942	−7.083	−5.468
2.2.01	−27.278	40.868	−12.902	−12.855	−12.168
2.2.02	−9.864	14.038	−3.100	−8.070	−6.996
2.2.03	−12.353	16.951	−3.776	−8.463	−7.641
2.2.04	−10.683	14.224	−3.123	−8.282	−7.863
2.2.05	−45.771	67.875	−25.164	−17.876	−20.936
2.2.06	−12.344	17.553	−3.742	−9.207	−7.740
2.2.07	−13.157	19.149	−3.975	−9.761	−7.744
2.3.01	−14.950	23.492	−6.930	−9.700	−8.088
2.3.02	−5.741	9.866	−2.076	−7.194	−5.145
2.3.03	−7.888	11.723	−2.670	−6.887	−5.722
2.3.04	−6.510	10.234	−2.598	−6.417	−5.291
2.3.05	−8.885	13.512	−3.443	−7.670	−6.486
2.3.06	−5.958	8.885	−2.039	−6.483	−5.595
2.3.07	−5.985	9.600	−2.246	−7.094	−5.726
2.3.08	−6.251	8.645	−1.909	−6.208	−5.725
2.3.09	−10.840	14.231	−3.865	−7.330	−7.804
2.3.10	−7.650	8.837	−2.028	−5.761	−6.603
2.3.11	−23.248	39.827	−14.921	−14.502	−12.845
2.3.12	−18.450	23.309	−7.279	−9.663	−12.083
2.3.13	−12.201	17.580	−4.576	−9.628	−8.824
2.3.14	−7.894	12.268	−2.803	−8.387	−6.817
2.4.01	−35.852	59.869	−23.254	−15.641	−14.879
2.4.02	−11.297	18.411	−4.221	−9.544	−6.650
2.4.03	−34.606	61.187	−23.337	−16.182	−12.939
2.4.04	−7.295	13.316	−3.127	−7.177	−4.283
2.4.05	−7.072	11.764	−2.892	−7.099	−5.298
2.5.01	−2.094	3.892	−1.038	−3.096	−2.336
2.6.01	−4.107	8.425	−2.098	−5.684	−3.464
2.6.02	−1.347	3.419	−0.480	−3.595	−2.003
2.7.01	−6.842	11.984	−2.476	−6.922	−4.256
2.7.02	−4.834	6.937	−0.433	−5.628	−3.958
2.8.01	−9.110	18.586	−5.586	−8.757	−4.867
2.8.02	−10.939	14.328	−2.481	−7.037	−6.130

Table 10: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $Br-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}	Chemical formula
2.1.01	−7.611	16.890	−3.891	−10.843	−5.454	
2.1.02	−4.071	8.970	−0.910	−8.368	−4.380	
2.1.03	−2.639	7.792	−0.687	−7.531	−3.065	
2.2.01	−25.025	39.404	−11.682	−13.168	−10.471	
2.2.02	−5.860	10.727	−1.078	−7.791	−4.003	
2.2.03	−9.811	14.947	−2.158	−8.607	−5.628	
2.2.04	−4.698	9.760	−1.005	−7.481	−3.423	
2.2.05	−41.510	62.684	−23.440	−16.446	−18.712	
2.2.06	−10.118	17.460	−2.221	−10.595	−5.474	
2.2.07	−8.210	13.932	−1.988	−8.022	−4.288	
2.3.01	−13.688	22.710	−6.093	−10.296	−7.366	
2.3.02	−4.367	8.835	−1.142	−7.237	−3.911	
2.3.03	−7.149	10.991	−1.870	−7.385	−5.413	
2.3.04	−3.473	7.198	−1.130	−5.564	−2.968	
2.3.05	−6.885	13.154	−2.369	−8.896	−4.996	
2.3.06	−3.515	8.070	−0.491	−7.571	−3.507	
2.3.07	−3.941	8.135	−0.750	−7.271	−3.826	
2.3.08	−2.828	6.462	−0.545	−6.244	−3.155	
2.3.09	−7.070	11.806	−2.736	−6.687	−4.687	
2.3.10	−4.619	6.871	−0.949	−5.777	−4.473	
2.3.11	−21.275	37.542	−13.231	−14.645	−11.610	
2.3.12	−6.323	12.223	−1.960	−9.263	−5.323	
2.3.13	−10.370	18.065	−3.679	−11.373	−7.358	
2.3.14	−4.958	11.557	−1.437	−9.941	−4.779	
2.4.01	−32.577	56.676	−21.788	−15.113	−12.802	
2.4.02	−9.055	15.496	−2.975	−8.785	−5.319	
2.4.03	−31.443	57.478	−21.788	−14.872	−10.625	
2.4.04	−7.611	13.432	−2.453	−7.629	−4.261	
2.4.05	−4.670	10.434	−1.403	−7.766	−3.404	
2.5.01	−1.338	3.489	−0.296	−3.311	−1.457	
2.6.01	−3.421	8.266	−1.598	−5.858	−2.611	
2.6.02	−1.970	4.999	−0.480	−4.847	−2.297	
2.7.01	−6.040	12.035	−1.648	−7.585	−3.238	
2.7.02	−3.335	6.006	0.315	−5.586	−2.599	
2.8.01	−8.606	19.179	−4.656	−9.817	−3.900	
2.8.02	−2.911	7.151	−0.570	−6.689	−3.018	

Table 11: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $I-H_2O$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
3.1.01	−7.595	11.815	−1.891	−9.281	−6.952
3.1.02	−6.855	10.829	−1.692	−8.689	−6.406
3.1.03	−9.617	15.958	−3.445	−10.616	−7.720
3.2.01	−15.970	17.539	−3.786	−7.862	−10.078
3.2.02	−9.081	10.994	−2.263	−6.330	−6.679
3.2.03	−18.772	21.702	−3.061	−9.765	−9.895
3.2.04	−12.427	15.940	−2.131	−9.014	−7.633
3.2.05	−29.266	37.389	−7.432	−13.487	−12.796
3.2.06	−14.200	17.295	−2.496	−9.052	−8.453
3.2.07	−43.922	53.854	−11.661	−17.788	−19.516
3.2.08	−20.217	25.166	−2.881	−12.420	−10.353
3.3.01	−15.073	17.816	−4.237	−8.296	−9.790
3.3.02	−23.069	28.041	−6.993	−10.976	−12.997
3.3.03	−15.691	18.273	−4.266	−8.609	−10.294
3.3.04	−5.802	7.854	−1.085	−5.453	−4.487
3.3.05	−7.081	9.274	−1.814	−6.439	−6.060
3.3.06	−10.730	14.564	−3.064	−8.096	−7.326
3.3.07	−7.615	10.591	−2.058	−6.936	−6.019
3.3.08	−9.288	10.340	−1.916	−6.451	−7.314
3.3.09	−8.086	10.119	−2.248	−6.060	−6.274
3.3.10	−11.928	14.576	−3.162	−7.813	−8.327
3.3.11	−6.487	10.227	−2.167	−7.622	−6.049
3.3.12	−17.075	24.442	−5.694	−12.318	−10.644
3.3.13	−11.595	18.451	−3.785	−11.154	−8.082
3.3.14	−33.792	43.749	−13.036	−15.020	−18.099
3.3.15	−17.857	23.429	−6.502	−11.329	−12.258
3.4.01	−17.755	23.157	−3.425	−10.893	−8.915
3.4.02	−11.999	16.354	−2.357	−9.591	−7.593
3.4.03	−33.249	45.031	−10.476	−15.162	−13.856
3.4.04	−12.696	17.231	−2.153	−8.241	−5.859
3.4.05	−9.133	12.649	−1.691	−7.831	−6.006
3.4.06	−30.240	43.900	−9.960	−14.665	−10.965
3.5.01	−1.867	3.406	−0.745	−3.011	−2.216
3.6.01	−3.818	6.462	−1.058	−5.313	−3.727
3.6.02	−4.522	6.527	−1.242	−5.151	−4.390
3.6.03	−4.787	8.751	−1.813	−6.199	−4.048
3.7.01	−3.098	6.047	−0.842	−5.267	−3.160
3.7.02	−14.810	26.735	−8.477	−10.626	−7.177
3.7.03	−6.313	11.408	−2.365	−7.309	−4.580
3.8.01	−3.009	7.415	−1.932	−6.023	−3.550
3.8.02	−6.000	9.169	−2.157	−6.817	−5.805
3.8.03	−7.391	16.844	−5.875	−8.994	−5.416

Table 12: Total two-body SAPT2+(CCD)dMP2 interaction energy contributions (kcal/mol) for $I-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
3.1.01	−5.572	10.477	−0.867	−9.180	−5.141
3.1.02	−3.750	8.568	−0.486	−8.156	−3.824
3.1.03	−8.143	15.777	−2.319	−11.255	−5.939
3.2.01	−11.782	15.684	−2.973	−7.437	−6.508
3.2.02	−4.341	6.889	−0.762	−5.292	−3.505
3.2.03	−15.581	19.864	−1.708	−10.344	−7.769
3.2.04	−9.316	14.191	−0.697	−9.121	−4.943
3.2.05	−26.689	35.570	−6.106	−13.756	−10.980
3.2.06	−9.836	14.238	−0.679	−9.169	−5.446
3.2.07	−39.969	49.113	−10.063	−16.404	−17.323
3.2.08	−16.060	20.406	−1.071	−10.681	−7.406
3.3.01	−5.998	9.460	−1.151	−6.397	−4.087
3.3.02	−14.559	20.543	−4.156	−9.296	−7.468
3.3.03	−6.838	11.227	−1.122	−8.353	−5.086
3.3.04	−6.378	10.337	−0.992	−7.861	−4.893
3.3.05	−3.987	8.079	−0.582	−7.034	−3.524
3.3.06	−7.760	12.821	−1.695	−8.717	−5.350
3.3.07	−3.875	7.398	−0.414	−6.753	−3.643
3.3.08	−5.989	8.133	−0.784	−5.995	−4.635
3.3.09	−4.921	9.515	−0.951	−7.969	−4.326
3.3.10	−8.259	12.326	−1.987	−7.333	−5.252
3.3.11	−5.606	9.679	−1.045	−7.903	−4.876
3.3.12	−16.154	23.538	−4.218	−12.818	−9.653
3.3.13	−9.120	16.416	−2.050	−11.262	−6.016
3.3.14	−24.588	35.670	−8.380	−14.939	−12.236
3.3.15	−10.283	16.426	−2.170	−10.956	−6.983
3.4.01	−15.103	20.765	−2.092	−10.715	−7.145
3.4.02	−8.648	13.690	−0.895	−9.014	−4.866
3.4.03	−29.639	41.358	−8.863	−14.548	−11.691
3.4.04	−12.951	19.236	−1.951	−10.548	−6.215
3.4.05	−7.649	13.409	−0.813	−9.238	−4.292
3.4.06	−28.834	42.046	−8.919	−14.088	−9.795
3.5.01	−1.275	2.973	−0.073	−3.117	−1.492
3.6.01	−2.625	5.682	−0.233	−5.576	−2.751
3.6.02	−1.961	4.794	−0.183	−4.933	−2.284
3.6.03	−3.594	7.719	−0.929	−6.022	−2.826
3.7.01	−3.344	7.274	−0.411	−6.742	−3.223
3.7.02	−12.855	25.140	−7.318	−10.524	−5.556
3.7.03	−5.053	10.367	−1.510	−7.052	−3.248
3.8.01	−3.046	7.965	−1.507	−6.684	−3.273
3.8.02	−2.535	6.631	−0.799	−6.413	−3.115
3.8.03	−6.546	16.423	−4.918	−9.181	−4.223

Table 13: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $Cl-H_2O$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−5.312	8.944	−1.686	−6.787	−4.841
1.1.02	−7.429	13.340	−3.415	−8.028	−5.533
1.2.01	−10.877	12.428	−3.148	−5.432	−7.029
1.2.02	−9.121	13.363	−2.824	−6.555	−5.137
1.2.03	−11.303	14.702	−3.482	−6.906	−6.990
1.2.04	−5.374	9.187	−1.877	−6.084	−4.148
1.2.05	−21.881	35.228	−9.778	−9.873	−6.304
1.2.06	−7.158	11.109	−2.412	−6.157	−4.619
1.2.07	−9.087	14.091	−2.805	−7.481	−5.282
1.2.08	−39.331	66.179	−23.136	−15.356	−11.644
1.3.01	−6.108	9.254	−1.754	−5.179	−3.788
1.3.02	−11.032	18.270	−4.480	−7.363	−4.604
1.3.03	−5.068	7.773	−1.769	−4.924	−3.988
1.3.04	−5.788	8.241	−1.710	−5.176	−4.433
1.3.05	−7.520	11.264	−2.206	−6.019	−4.482
1.3.06	−6.459	7.817	−1.679	−4.601	−4.923
1.3.07	−8.669	11.528	−2.640	−5.609	−5.390
1.3.08	−18.352	22.869	−6.191	−8.769	−10.443
1.3.09	−14.630	18.485	−5.205	−7.259	−8.609
1.3.10	−26.879	38.652	−12.178	−11.682	−12.086
1.4.01	−7.012	11.807	−2.302	−6.623	−4.130
1.4.02	−8.229	11.575	−2.489	−6.816	−5.959
1.4.03	−31.128	55.392	−19.803	−13.400	−8.938
1.4.04	−5.644	9.355	−1.740	−5.391	−3.419
1.4.05	−28.509	53.497	−18.772	−12.827	−6.610
1.5.01	−1.684	3.322	−0.711	−2.544	−1.617
1.6.01	−3.413	7.014	−1.531	−4.613	−2.543
1.7.01	−4.308	8.894	−2.071	−5.331	−2.815
1.8.01	−6.308	13.292	−3.226	−6.679	−2.921

Table 14: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $Cl-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−3.212	8.111	−0.830	−7.122	−3.053
1.1.02	−5.540	13.035	−2.660	−8.377	−3.542
1.2.01	−5.861	9.163	−2.248	−4.162	−3.107
1.2.02	−5.997	10.992	−1.434	−6.559	−2.997
1.2.03	−2.382	6.428	−0.726	−5.164	−1.845
1.2.04	−3.451	8.462	−0.803	−6.848	−2.639
1.2.05	−18.663	33.379	−8.590	−10.501	−4.374
1.2.06	−3.837	8.402	−1.184	−5.552	−2.170
1.2.07	−5.530	12.392	−1.471	−7.921	−2.530
1.2.08	−35.164	60.980	−21.714	−13.890	−9.789
1.3.01	−4.591	8.041	−1.138	−5.020	−2.708
1.3.02	−9.339	16.523	−3.858	−6.635	−3.309
1.3.03	−3.611	7.556	−0.776	−5.741	−2.572
1.3.04	−2.695	6.551	−0.587	−5.555	−2.285
1.3.05	−5.000	10.517	−1.542	−6.892	−2.917
1.3.06	−2.732	5.267	−0.644	−4.148	−2.257
1.3.07	−5.352	9.977	−1.742	−5.620	−2.737
1.3.08	−7.039	14.108	−1.978	−9.382	−4.291
1.3.09	−2.298	7.130	−0.927	−5.966	−2.061
1.3.10	−16.016	30.165	−8.324	−11.939	−6.114
1.4.01	−5.194	10.203	−1.525	−6.562	−3.078
1.4.02	−3.261	7.664	−0.964	−6.078	−2.639
1.4.03	−27.795	52.223	−18.580	−13.113	−7.265
1.4.04	−5.271	10.278	−1.506	−6.516	−3.015
1.4.05	−27.281	52.769	−17.837	−13.526	−5.876
1.5.01	−1.206	3.482	−0.257	−2.992	−0.973
1.6.01	−2.702	7.349	−1.064	−5.183	−1.599
1.7.01	−3.334	8.751	−1.593	−5.411	−1.587
1.8.01	−5.357	12.978	−2.692	−6.682	−1.753

Table 15: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $Br-H_2O$)t systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
2.1.01	−9.073	16.107	−4.445	−8.499	−5.910
2.1.02	−5.925	9.749	−1.941	−7.198	−5.315
2.1.03	−5.177	8.526	−1.772	−6.186	−4.609
2.2.01	−25.049	39.220	−11.660	−11.185	−8.673
2.2.02	−9.164	13.808	−2.938	−7.026	−5.320
2.2.03	−11.400	16.797	−3.749	−7.413	−5.766
2.2.04	−10.014	13.915	−2.895	−7.219	−6.214
2.2.05	−41.347	64.090	−21.896	−15.567	−14.720
2.2.06	−11.223	16.906	−3.494	−8.005	−5.817
2.2.07	−12.048	18.285	−3.683	−8.491	−5.936
2.3.01	−14.391	22.676	−6.127	−8.419	−6.261
2.3.02	−5.411	9.561	−2.009	−6.240	−4.099
2.3.03	−7.707	11.807	−2.614	−6.061	−4.574
2.3.04	−6.193	9.842	−2.321	−5.587	−4.259
2.3.05	−8.458	12.997	−2.841	−6.619	−4.922
2.3.06	−5.737	8.719	−1.734	−5.632	−4.384
2.3.07	−5.698	9.276	−2.046	−6.227	−4.695
2.3.08	−6.042	8.429	−1.618	−5.401	−4.633
2.3.09	−10.350	13.892	−3.429	−6.291	−6.179
2.3.10	−7.409	8.961	−1.971	−5.011	−5.430
2.3.11	−21.572	36.033	−11.998	−12.332	−9.870
2.3.12	−17.979	22.524	−6.051	−8.182	−9.688
2.3.13	−11.605	16.667	−3.981	−8.311	−7.230
2.3.14	−7.355	11.461	−2.409	−7.197	−5.500
2.4.01	−33.927	55.695	−19.628	−13.720	−11.579
2.4.02	−10.739	17.761	−3.944	−8.433	−5.355
2.4.03	−33.753	57.063	−19.877	−14.187	−10.754
2.4.04	−7.113	12.958	−2.874	−6.358	−3.387
2.4.05	−6.759	11.078	−2.487	−6.208	−4.375
2.5.01	−1.952	3.781	−0.851	−2.718	−1.739
2.6.01	−3.823	7.971	−1.848	−4.955	−2.655
2.6.02	−1.237	3.471	−0.528	−3.195	−1.489
2.7.01	−5.493	11.183	−2.929	−6.013	−3.251
2.7.02	−3.684	6.769	−1.287	−4.942	−3.144
2.8.01	−8.515	16.999	−4.622	−7.583	−3.722
2.8.02	−10.363	13.385	−2.928	−6.099	−6.006

Table 16: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $Br-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
2.1.01	−7.180	15.975	−3.677	−9.647	−4.529
2.1.02	−3.901	9.047	−1.165	−7.512	−3.531
2.1.03	−2.422	7.552	−0.848	−6.698	−2.416
2.2.01	−23.118	38.264	−10.758	−11.597	−7.209
2.2.02	−5.454	11.002	−1.364	−6.921	−2.736
2.2.03	−9.107	15.245	−2.519	−7.660	−4.041
2.2.04	−4.403	9.884	−1.186	−6.647	−2.352
2.2.05	−37.353	59.333	−20.488	−14.442	−12.950
2.2.06	−9.124	17.174	−2.457	−9.356	−3.763
2.2.07	−7.333	13.452	−2.066	−7.072	−3.019
2.3.01	−13.212	22.310	−5.475	−9.056	−5.432
2.3.02	−4.275	9.213	−1.287	−6.448	−2.797
2.3.03	−7.048	11.411	−2.009	−6.595	−4.242
2.3.04	−3.329	7.314	−1.229	−4.908	−2.152
2.3.05	−6.677	13.114	−2.144	−7.870	−3.577
2.3.06	−3.413	8.352	−0.699	−6.762	−2.522
2.3.07	−3.815	8.435	−0.974	−6.519	−2.873
2.3.08	−2.788	6.690	−0.632	−5.573	−2.304
2.3.09	−6.747	11.725	−2.540	−5.821	−3.383
2.3.10	−4.496	7.372	−1.130	−5.121	−3.375
2.3.11	−19.876	34.462	−10.697	−12.647	−8.758
2.3.12	−6.050	12.370	−1.878	−8.095	−3.651
2.3.13	−9.944	17.577	−3.315	−9.988	−5.671
2.3.14	−4.608	11.347	−1.432	−8.755	−3.447
2.4.01	−30.830	53.005	−18.501	−13.400	−9.726
2.4.02	−8.635	15.225	−2.950	−7.889	−4.248
2.4.03	−30.644	54.009	−18.725	−13.182	−8.542
2.4.04	−7.615	13.464	−2.442	−6.861	−3.454
2.4.05	−4.545	10.243	−1.399	−6.930	−2.631
2.5.01	−1.306	3.644	−0.331	−2.977	−0.970
2.6.01	−3.298	8.125	−1.488	−5.209	−1.870
2.6.02	−1.967	5.260	−0.584	−4.371	−1.662
2.7.01	−4.850	11.596	−2.343	−6.698	−2.295
2.7.02	−2.340	6.208	−0.736	−4.999	−1.867
2.8.01	−8.150	17.911	−4.034	−8.624	−2.897
2.8.02	−2.840	7.287	−0.776	−5.967	−2.295

Table 17: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $I-H_2O$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
3.1.01	−7.264	11.096	−1.535	−8.166	−5.869
3.1.02	−6.509	9.967	−1.310	−7.578	−5.430
3.1.03	−9.030	14.353	−2.244	−9.243	−6.164
3.2.01	−15.429	16.704	−4.612	−6.951	−10.288
3.2.02	−8.723	10.682	−2.551	−5.527	−6.119
3.2.03	−17.639	21.350	−5.243	−8.567	−10.098
3.2.04	−11.490	15.084	−3.181	−7.839	−7.426
3.2.05	−26.931	35.052	−10.118	−11.711	−13.707
3.2.06	−13.381	16.770	−3.727	−7.881	−8.218
3.2.07	−39.551	50.065	−14.024	−15.426	−18.936
3.2.08	−18.345	23.810	−4.890	−10.804	−10.228
3.3.01	−14.756	17.299	−4.406	−7.177	−9.039
3.3.02	−22.458	26.724	−7.875	−9.458	−13.067
3.3.03	−15.385	17.875	−4.570	−7.451	−9.531
3.3.04	−5.582	7.836	−1.501	−4.758	−4.006
3.3.05	−6.899	8.995	−1.831	−5.598	−5.333
3.3.06	−10.360	13.942	−3.161	−6.995	−6.574
3.3.07	−7.428	10.394	−2.120	−6.033	−5.186
3.3.08	−9.010	10.334	−2.528	−5.584	−6.787
3.3.09	−7.914	9.863	−2.222	−5.219	−5.492
3.3.10	−11.401	13.926	−3.772	−6.673	−7.920
3.3.11	−6.447	9.825	−2.198	−6.601	−5.421
3.3.12	−16.162	22.470	−5.126	−10.521	−9.338
3.3.13	−10.763	16.674	−3.382	−9.499	−6.970
3.3.14	−32.234	40.219	−13.332	−12.615	−17.962
3.3.15	−17.135	21.707	−6.345	−9.538	−11.311
3.4.01	−16.891	21.714	−2.912	−9.591	−7.681
3.4.02	−11.370	14.955	−1.930	−8.376	−6.721
3.4.03	−31.335	40.869	−7.030	−13.223	−10.719
3.4.04	−12.513	16.177	−2.019	−7.256	−5.611
3.4.05	−9.004	11.446	−1.453	−6.822	−5.833
3.4.06	−29.382	39.855	−7.197	−12.799	−9.522
3.5.01	−1.764	3.406	−0.694	−2.670	−1.722
3.6.01	−3.633	6.309	−1.109	−4.684	−3.117
3.6.02	−4.311	6.043	−1.125	−4.477	−3.870
3.6.03	−4.479	8.146	−1.675	−5.391	−3.399
3.7.01	−2.886	5.954	−0.947	−4.627	−2.506
3.7.02	−14.074	24.559	−7.092	−9.265	−5.871
3.7.03	−5.892	10.400	−2.092	−6.316	−3.899
3.8.01	−3.342	6.570	−0.955	−5.283	−3.011
3.8.02	−6.121	8.068	−1.177	−5.906	−5.136
3.8.03	−8.567	14.016	−2.241	−7.779	−4.571

Table 18: Total two-body SAPT(DFT) interaction energy contributions (kcal/mol) for $I-CH_4$) systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
3.1.01	−5.359	10.184	−0.856	−8.190	−4.220
3.1.02	−3.527	8.119	−0.433	−7.230	−3.071
3.1.03	−7.674	14.590	−1.445	−9.954	−4.484
3.2.01	−11.197	15.002	−3.997	−6.616	−6.809
3.2.02	−4.095	6.947	−1.362	−4.751	−3.261
3.2.03	−14.522	19.690	−4.131	−9.167	−8.131
3.2.04	−8.671	13.877	−2.131	−8.074	−4.998
3.2.05	−24.613	33.747	−9.123	−12.086	−12.075
3.2.06	−9.137	14.058	−2.280	−8.102	−5.460
3.2.07	−35.824	45.766	−12.770	−14.341	−17.170
3.2.08	−14.370	19.436	−3.472	−9.382	−7.789
3.3.01	−5.772	9.249	−1.854	−5.600	−3.977
3.3.02	−14.002	19.476	−5.541	−8.080	−8.146
3.3.03	−6.621	11.279	−1.976	−7.390	−4.708
3.3.04	−6.261	10.562	−1.544	−6.992	−4.234
3.3.05	−3.861	8.137	−0.899	−6.242	−2.864
3.3.06	−7.466	12.563	−2.111	−7.654	−4.669
3.3.07	−3.739	7.524	−0.827	−5.972	−3.014
3.3.08	−5.774	8.355	−1.675	−5.272	−4.365
3.3.09	−4.916	9.662	−1.280	−7.024	−3.558
3.3.10	−7.819	11.924	−2.893	−6.346	−5.133
3.3.11	−5.665	9.879	−1.395	−6.981	−4.162
3.3.12	−15.401	22.226	−4.092	−11.166	−8.433
3.3.13	−8.479	15.333	−2.081	−9.794	−5.021
3.3.14	−23.269	33.283	−9.854	−12.815	−12.655
3.3.15	−9.799	16.049	−3.161	−9.497	−6.408
3.4.01	−14.301	19.713	−1.937	−9.562	−6.087
3.4.02	−8.101	12.728	−0.833	−7.994	−4.201
3.4.03	−27.848	37.662	−5.801	−12.821	−8.808
3.4.04	−12.821	18.325	−1.963	−9.415	−5.873
3.4.05	−7.430	12.626	−0.879	−8.211	−3.894
3.4.06	−27.951	38.270	−6.399	−12.379	−8.458
3.5.01	−1.238	3.189	−0.253	−2.832	−1.134
3.6.01	−2.554	5.909	−0.561	−5.005	−2.212
3.6.02	−1.869	4.711	−0.352	−4.390	−1.900
3.6.03	−3.398	7.357	−1.045	−5.315	−2.401
3.7.01	−3.217	7.289	−0.706	−6.005	−2.639
3.7.02	−12.227	23.362	−6.208	−9.285	−4.357
3.7.03	−4.759	9.601	−1.477	−6.174	−2.809
3.8.01	−3.479	7.197	−0.711	−5.930	−2.922
3.8.02	−2.747	5.842	−0.168	−5.657	−2.729
3.8.03	−7.845	13.859	−1.596	−8.044	−3.626

Table 19: Two and Three-body XSAPT interaction energies (kcal/mol) for $Cl-H_2O$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−5.388	9.144	−1.789	−5.875	−3.907	−5.388	9.144	−1.760	−5.842	−3.846
1.1.02	−7.528	13.648	−3.525	−7.168	−4.573	−7.528	13.648	−3.502	−7.140	−4.522
1.2.01	−11.003	12.798	−3.258	−4.597	−6.060	−11.003	12.798	−3.316	−4.584	−6.105
1.2.02	−9.247	13.612	−2.863	−5.431	−3.929	−9.247	13.612	−2.658	−5.358	−3.652
1.2.03	−11.461	14.921	−3.724	−5.753	−6.018	−11.461	14.921	−3.528	−5.712	−5.781
1.2.04	−5.440	9.279	−1.997	−5.331	−3.489	−5.440	9.279	−2.324	−5.285	−3.770
1.2.05	−22.229	35.194	−9.481	−8.294	−4.809	−22.229	35.194	−10.022	−8.255	−5.311
1.2.06	−7.269	11.410	−2.492	−5.074	−3.425	−7.269	11.410	−2.215	−5.046	−3.119
1.2.07	−9.158	14.245	−2.803	−6.232	−3.948	−9.158	14.245	−2.695	−6.170	−3.779
1.2.08	−40.202	65.393	−22.994	−13.076	−10.880	−40.202	65.393	−24.300	−13.000	−12.109
1.3.01	−6.245	9.547	−1.703	−4.143	−2.543	−6.245	9.547	−2.139	−4.128	−2.965
1.3.02	−11.303	18.751	−4.478	−6.092	−3.121	−11.303	18.751	−5.161	−6.070	−3.783
1.3.03	−5.164	8.089	−1.910	−4.171	−3.155	−5.164	8.089	−1.845	−4.162	−3.081
1.3.04	−5.882	8.267	−1.925	−4.399	−3.938	−5.882	8.267	−1.789	−4.353	−3.757
1.3.05	−7.667	11.506	−2.300	−4.965	−3.427	−7.667	11.506	−2.254	−4.907	−3.323
1.3.06	−6.630	8.117	−1.828	−3.635	−3.976	−6.630	8.117	−1.693	−3.618	−3.824
1.3.07	−8.877	11.788	−2.802	−4.461	−4.351	−8.877	11.788	−2.635	−4.441	−4.165
1.3.08	−18.375	22.817	−6.771	−7.311	−9.640	−18.375	22.817	−6.723	−7.247	−9.528
1.3.09	−14.787	18.519	−5.866	−5.997	−8.130	−14.787	18.519	−5.709	−5.951	−7.928
1.3.10	−26.915	38.386	−12.598	−9.756	−10.884	−26.915	38.386	−12.483	−9.674	−10.687
1.4.01	−7.224	12.204	−2.229	−5.570	−2.819	−7.224	12.204	−2.506	−5.547	−3.072
1.4.02	−8.476	12.140	−2.592	−5.605	−4.533	−8.476	12.140	−2.395	−5.564	−4.295
1.4.03	−32.256	55.464	−18.529	−12.541	−7.863	−32.256	55.464	−18.905	−12.464	−8.160
1.4.04	−5.820	9.679	−1.643	−4.568	−2.352	−5.820	9.679	−1.910	−4.551	−2.603
1.4.05	−29.399	53.160	−17.676	−12.160	−6.074	−29.399	53.160	−19.136	−12.148	−7.522
1.5.01	−1.792	3.548	−0.801	−2.001	−1.047	−1.792	3.548	−0.840	−1.992	−1.076
1.6.01	−3.618	7.337	−1.646	−4.001	−1.928	−3.618	7.337	−1.694	−3.983	−1.958
1.7.01	−4.592	9.462	−2.169	−4.629	−1.927	−4.592	9.462	−2.215	−4.611	−1.956

Table 20: Two and Three-body XSAPT interaction energies (kcal/mol) for $Cl-CH_4$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
1.1.01	−3.073	7.700	−0.788	−6.020	−2.182	−3.073	7.700	−0.808	−5.986	−2.168
1.1.02	−5.473	12.763	−2.599	−7.216	−2.524	−5.473	12.763	−2.622	−7.189	−2.521
1.2.01	−5.961	9.369	−2.248	−3.629	−2.470	−5.961	9.369	−2.265	−3.636	−2.493
1.2.02	−5.913	10.690	−1.267	−5.347	−1.837	−5.913	10.690	−1.251	−5.329	−1.804
1.2.03	−2.366	6.234	−0.656	−4.068	−0.855	−2.366	6.234	−0.714	−4.021	−0.867
1.2.04	−3.351	8.087	−0.728	−5.824	−1.816	−3.351	8.087	−0.727	−5.795	−1.787
1.2.05	−18.819	32.767	−8.047	−8.702	−2.800	−18.819	32.767	−8.042	−8.668	−2.761
1.2.06	−3.853	8.294	−1.043	−4.401	−1.003	−3.853	8.294	−1.014	−4.379	−0.952
1.2.07	−5.411	12.011	−1.262	−6.672	−1.333	−5.411	12.011	−1.258	−6.663	−1.320
1.2.08	−35.861	59.606	−21.434	−11.913	−9.603	−35.861	59.606	−21.637	−11.888	−9.780
1.3.01	−4.608	8.079	−1.022	−3.978	−1.530	−4.608	8.079	−1.089	−3.962	−1.580
1.3.02	−9.461	16.632	−3.700	−5.327	−1.855	−9.461	16.632	−3.685	−5.292	−1.805
1.3.03	−3.543	7.299	−0.754	−4.530	−1.527	−3.543	7.299	−0.720	−4.492	−1.456
1.3.04	−2.601	6.167	−0.577	−4.494	−1.505	−2.601	6.167	−0.604	−4.461	−1.500
1.3.05	−4.953	10.257	−1.496	−5.495	−1.687	−4.953	10.257	−1.496	−5.443	−1.634
1.3.06	−2.713	5.248	−0.569	−3.175	−1.209	−2.713	5.248	−0.563	−3.156	−1.185
1.3.07	−5.371	9.884	−1.670	−4.362	−1.519	−5.371	9.884	−1.668	−4.344	−1.499
1.3.08	−6.813	13.572	−1.832	−7.649	−2.721	−6.813	13.572	−1.809	−7.610	−2.660
1.3.09	−2.195	6.714	−0.903	−4.837	−1.221	−2.195	6.714	−0.916	−4.824	−1.220
1.3.10	−15.858	29.409	−8.113	−9.835	−4.397	−15.858	29.409	−8.130	−9.792	−4.370
1.4.01	−5.300	10.325	−1.385	−5.569	−1.929	−5.300	10.325	−1.388	−5.535	−1.897
1.4.02	−3.314	7.623	−0.918	−4.919	−1.528	−3.314	7.623	−0.962	−4.877	−1.530
1.4.03	−28.751	51.756	−17.144	−12.196	−6.334	−28.751	51.756	−17.190	−12.147	−6.332
1.4.04	−5.303	10.206	−1.384	−5.416	−1.896	−5.303	10.206	−1.388	−5.355	−1.840
1.4.05	−27.966	51.844	−16.518	−12.683	−5.324	−27.966	51.844	−16.548	−12.607	−5.277
1.5.01	−1.190	3.361	−0.232	−2.335	−0.396	−1.190	3.361	−0.231	−2.331	−0.392
1.6.01	−2.744	7.208	−1.082	−4.327	−0.945	−2.744	7.208	−1.080	−4.289	−0.905
1.7.01	−3.454	8.855	−1.613	−4.540	−0.752	−3.454	8.855	−1.620	−4.535	−0.754

Table 21: Two and Three-body XSAPT interaction energies(kcal/mol) for $Br-H_2O$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
2.1.01	−9.140	16.441	−4.514	−7.867	−5.080	−9.140	16.441	−4.699	−7.812	−5.209
2.1.02	−5.979	10.069	−1.983	−6.268	−4.160	−5.979	10.069	−1.981	−6.230	−4.122
2.1.03	−5.247	8.804	−1.795	−5.189	−3.427	−5.247	8.804	−1.882	−5.150	−3.474
2.2.01	−25.350	38.638	−11.789	−9.664	−8.165	−25.350	38.638	−12.321	−9.557	−8.590
2.2.02	−9.270	13.917	−2.936	−5.880	−4.168	−9.270	13.917	−2.615	−5.789	−3.757
2.2.03	−11.529	16.946	−3.608	−6.199	−4.390	−11.529	16.946	−3.377	−6.113	−4.074
2.2.04	−10.201	14.335	−2.940	−5.979	−4.785	−10.201	14.335	−2.692	−5.912	−4.471
2.2.05	−42.112	63.065	−23.313	−13.930	−16.289	−42.112	63.065	−25.034	−13.851	−17.932
2.2.06	−11.307	17.028	−3.297	−6.913	−4.489	−11.307	17.028	−3.316	−6.867	−4.462
2.2.07	−12.266	18.759	−3.468	−6.899	−3.873	−12.266	18.759	−3.476	−6.844	−3.827
2.3.01	−14.645	23.079	−6.208	−7.219	−4.993	−14.645	23.079	−6.992	−7.193	−5.750
2.3.02	−5.591	10.006	−2.050	−5.322	−2.957	−5.591	10.006	−2.295	−5.321	−3.201
2.3.03	−7.887	12.349	−2.574	−4.993	−3.105	−7.887	12.349	−2.996	−4.967	−3.501
2.3.04	−6.454	10.595	−2.384	−4.471	−2.714	−6.454	10.595	−2.274	−4.445	−2.577
2.3.05	−8.621	13.237	−3.021	−5.664	−4.069	−8.621	13.237	−2.983	−5.596	−3.963
2.3.06	−5.869	8.907	−1.912	−4.781	−3.655	−5.869	8.907	−1.791	−4.728	−3.480
2.3.07	−5.840	9.690	−2.227	−5.271	−3.649	−5.840	9.690	−2.172	−5.249	−3.571
2.3.08	−6.194	8.731	−1.768	−4.488	−3.718	−6.194	8.731	−1.649	−4.445	−3.557
2.3.09	−10.572	14.174	−3.622	−5.161	−5.181	−10.572	14.174	−3.371	−5.138	−4.906
2.3.10	−7.607	9.353	−2.121	−3.975	−4.350	−7.607	9.353	−1.933	−3.956	−4.143
2.3.11	−21.464	35.682	−12.171	−10.616	−8.569	−21.464	35.682	−13.067	−10.560	−9.409
2.3.12	−18.131	22.539	−6.670	−6.945	−9.208	−18.131	22.539	−6.620	−6.944	−9.157
2.3.13	−11.568	16.863	−3.953	−6.978	−5.636	−11.568	16.863	−4.561	−6.950	−6.216
2.3.14	−7.543	11.949	−2.457	−5.803	−3.853	−7.543	11.949	−2.611	−5.769	−3.974
2.4.01	−35.183	55.560	−19.259	−13.393	−12.276	−35.183	55.560	−20.041	−13.319	−12.983
2.4.02	−11.170	18.699	−3.834	−7.229	−3.534	−11.170	18.699	−4.243	−7.176	−3.890
2.4.03	−34.716	56.409	−19.216	−13.946	−11.469	−34.716	56.409	−20.257	−13.836	−12.400
2.4.04	−7.391	13.435	−2.817	−5.644	−2.417	−7.391	13.435	−2.814	−5.636	−2.406
2.4.05	−7.002	11.770	−2.528	−5.187	−2.946	−7.002	11.770	−2.297	−5.161	−2.691
2.5.01	−2.071	4.123	−0.924	−2.184	−1.056	−2.070	4.123	−0.969	−2.174	−1.091
2.6.01	−4.088	8.510	−2.035	−4.427	−2.040	−4.088	8.510	−2.056	−4.421	−2.055
2.6.02	−1.321	3.621	−0.552	−2.773	−1.025	−1.321	3.621	−0.592	−2.770	−1.061
2.7.01	−5.834	11.912	−2.032	−5.390	−1.344	−5.834	11.912	−2.130	−5.372	−1.423
2.7.02	−3.896	7.290	−0.335	−4.116	−1.057	−3.896	7.290	−0.410	−4.098	−1.113

Table 22: Two and Three-body XSAPT interaction energies(kcal/mol) for $Br-CH_4$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding exchange effects, e.g., $E^{(2)}_{Disp} = E^{(2)}_{disp} + E^{(2)}_{exch-disp}$.

Name	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}	$E^{(1)}_{elst}$	$E^{(1)}_{exch}$	$E^{(2)}_{Ind}$	$E^{(2)}_{Disp}$	E_{int}
2.1.01	−7.125	15.812	−3.582	−8.581	−3.475	−7.125	15.812	−3.631	−8.547	−3.490
2.1.02	−3.779	8.789	−1.086	−6.350	−2.426	−3.779	8.789	−1.112	−6.315	−2.417
2.1.03	−2.363	7.333	−0.691	−5.445	−1.167	−2.363	7.333	−0.772	−5.416	−1.219
2.2.01	−23.220	37.079	−10.833	−10.017	−6.992	−23.220	37.079	−10.861	−9.971	−6.973
2.2.02	−5.375	10.620	−1.159	−5.731	−1.646	−5.375	10.620	−1.159	−5.700	−1.615
2.2.03	−9.034	14.868	−2.216	−6.346	−2.728	−9.034	14.868	−2.199	−6.318	−2.683
2.2.04	−4.386	9.662	−0.998	−5.452	−1.173	−4.386	9.662	−1.001	−5.440	−1.164
2.2.05	−37.961	57.755	−21.823	−13.054	−15.083	−37.961	57.755	−21.923	−12.996	−15.124
2.2.06	−9.018	16.717	−2.102	−8.032	−2.435	−9.018	16.717	−2.104	−8.027	−2.432
2.2.07	−7.373	13.316	−1.691	−5.811	−1.560	−7.373	13.316	−1.705	−5.777	−1.540
2.3.01	−13.284	22.143	−5.413	−7.733	−4.286	−13.284	22.143	−5.421	−7.669	−4.231
2.3.02	−4.271	9.106	−1.196	−5.212	−1.573	−4.271	9.106	−1.245	−5.146	−1.556
2.3.03	−7.055	11.391	−1.899	−5.383	−2.945	−7.055	11.391	−1.888	−5.327	−2.878
2.3.04	−3.418	7.430	−1.134	−3.724	−0.846	−3.418	7.430	−1.225	−3.691	−0.903
2.3.05	−6.625	12.801	−2.168	−6.471	−2.464	−6.625	12.801	−2.152	−6.406	−2.382
2.3.06	−3.350	8.130	−0.662	−5.738	−1.620	−3.350	8.130	−0.658	−5.726	−1.605
2.3.07	−3.754	8.242	−0.913	−5.251	−1.675	−3.754	8.242	−0.913	−5.220	−1.644
2.3.08	−2.760	6.500	−0.620	−4.440	−1.321	−2.760	6.500	−0.638	−4.415	−1.313
2.3.09	−6.806	11.728	−2.538	−4.676	−2.292	−6.806	11.728	−2.555	−4.666	−2.298
2.3.10	−4.485	7.319	−1.091	−4.067	−2.324	−4.485	7.319	−1.085	−4.048	−2.299
2.3.11	−19.564	33.330	−10.681	−10.781	−7.696	−19.564	33.330	−10.716	−10.743	−7.693
2.3.12	−5.987	11.997	−1.817	−6.733	−2.539	−5.987	11.997	−1.821	−6.693	−2.504
2.3.13	−9.677	16.974	−3.195	−8.209	−4.106	−9.677	16.974	−3.155	−8.163	−4.021
2.3.14	−4.562	11.052	−1.328	−7.024	−1.861	−4.562	11.052	−1.318	−7.003	−1.831
2.4.01	−31.958	52.415	−18.021	−12.997	−10.561	−31.958	52.415	−18.073	−12.945	−10.561
2.4.02	−8.910	15.630	−2.758	−6.780	−2.819	−8.910	15.630	−2.787	−6.739	−2.806
2.4.03	−31.448	52.777	−17.944	−13.065	−9.680	−31.448	52.777	−18.035	−13.048	−9.754
2.4.04	−7.744	13.605	−2.295	−5.897	−2.332	−7.744	13.605	−2.357	−5.879	−2.376
2.4.05	−4.623	10.283	−1.270	−5.652	−1.262	−4.623	10.283	−1.309	−5.599	−1.249
2.5.01	−1.328	3.682	−0.304	−2.277	−0.227	−1.328	3.682	−0.297	−2.268	−0.211
2.6.01	−3.381	8.146	−1.547	−4.508	−1.290	−3.381	8.146	−1.562	−4.493	−1.290
2.6.02	−1.975	5.152	−0.602	−3.568	−0.994	−1.975	5.152	−0.615	−3.551	−0.989
2.7.01	−5.035	11.837	−1.384	−5.794	−0.376	−5.035	11.837	−1.380	−5.779	−0.357
2.7.02	−2.367	6.211	0.246	−4.068	0.021	−2.367	6.211	0.221	−4.053	0.012

Table 23: Three-body SAPT(DFT) interaction energy contributions (kcal mol^{-1}) for Cl & $Br-H_2O$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding nonadditive exchange effects: specifically, $E^{(2)}_{Ind,3} = E^{(2)}_{ind,3} + E^{(2)}_{exch-ind,3}$ and $E^{(3)}_{Disp,3} = E^{(3)}_{disp,3} + E^{(2)}_{exch-disp,3}$.

Complex	SAPT(DFT)		
	$E^{(1)}_{exch,3}$	$E^{(2)}_{Ind,3}$	$E^{(3)}_{Disp,3}$
1.1.01	−0.062	0.019	0.144
1.1.02	−0.010	0.023	0.132
1.2.01	−0.132	0.152	0.065
1.2.02	−0.059	0.116	0.185
1.2.03	−0.207	0.380	0.177
1.2.04	−0.082	−0.296	0.080
1.2.05	−0.061	−1.140	0.138
1.2.06	0.048	0.020	0.090
1.2.07	−0.044	0.173	0.181
1.2.08	−0.195	−2.349	0.291
1.3.01	−0.061	−0.473	0.056
1.3.02	−0.166	−0.884	0.098
1.3.03	−0.024	0.114	0.023
1.3.04	−0.062	0.191	0.105
1.3.05	−0.066	0.005	0.139
1.3.06	−0.048	0.190	0.079
1.3.07	−0.060	0.230	0.106
1.3.08	−0.158	0.302	0.169
1.3.09	−0.491	0.664	0.272
1.3.10	−0.327	0.520	0.251
1.4.01	−0.126	−0.248	0.099
1.4.02	−0.128	0.285	0.167
1.4.03	0.013	−1.657	0.289
1.4.04	−0.027	−0.415	0.062
1.4.05	−0.439	−2.632	0.177
1.5.01	−0.010	−0.067	0.020
1.6.01	−0.004	−0.154	0.064
1.7.01	0.002	−0.182	0.061
2.1.01	−0.068	−0.583	0.141
2.1.02	−0.046	−0.051	0.149
2.1.03	−0.055	−0.140	0.117
2.2.01	0.062	−0.816	0.259
2.2.02	−0.103	0.222	0.229
2.2.03	−0.002	0.013	0.204
2.2.04	−0.176	0.318	0.214
2.2.05	−0.163	−2.309	0.273
2.2.06	−0.007	−0.080	0.136
2.2.07	−0.045	−0.088	0.183
2.3.01	−0.161	−1.096	0.108
2.3.02	−0.032	−0.186	0.005
2.3.03	−0.036	−0.588	0.086
2.3.04	0.027	−0.088	0.093
2.3.05	−0.047	0.038	0.152
2.3.06	−0.066	0.173	0.084
2.3.07	−0.011	0.094	0.084
2.3.08	−0.079	0.194	0.116
2.3.09	−0.045	0.298	0.114
2.3.10	−0.047	0.229	0.089
2.3.11	0.026	−2.042	0.170
2.3.12	−0.095	0.280	0.037
2.3.13	−0.043	−0.920	0.073
2.3.14	−0.065	−0.200	0.127
2.4.01	0.031	−1.830	0.262
2.4.02	−0.047	−0.740	0.165
2.4.03	0.197	−1.991	0.325
2.4.04	0.016	−0.156	0.054
2.4.05	0.050	0.023	0.099
2.5.01	−0.011	−0.072	0.017
2.6.01	0.000	−0.073	0.027
2.6.02	−0.016	−0.032	0.013
2.7.01	0.023	−0.336	0.056
2.7.02	0.027	−0.140	0.070

Table 24: Three-body SAPT(DFT) interaction energy contributions (kcal mol^{-1}) for Cl & $Br-CH_4$ systems in the 3BXB dataset. The capitalized component names indicate the inclusion of the corresponding nonadditive exchange effects: specifically, $E^{(2)}_{Ind,3} = E^{(2)}_{ind,3} + E^{(2)}_{exch-ind,3}$ and $E^{(3)}_{Disp,3} = E^{(3)}_{disp,3} + E^{(2)}_{exch-disp,3}$.

Complex	SAPT(DFT)		
	$E^{(1)}_{exch,3}$	$E^{(2)}_{Ind,3}$	$E^{(3)}_{Disp,3}$
1.1.01	−0.160	0.039	0.160
1.1.02	−0.112	0.086	0.149
1.2.01	−0.004	−0.025	−0.003
1.2.02	−0.017	0.053	0.049
1.2.03	−0.113	0.016	0.135
1.2.04	−0.008	−0.001	0.049
1.2.05	−0.094	0.026	0.117
1.2.06	−0.029	0.027	0.082
1.2.07	−0.008	−0.037	0.016
1.2.08	−0.133	−0.100	0.148
1.3.01	−0.047	−0.054	0.062
1.3.02	−0.155	0.141	0.119
1.3.03	−0.053	0.037	0.094
1.3.04	−0.041	0.006	0.076
1.3.05	−0.081	0.002	0.153
1.3.06	−0.067	0.033	0.068
1.3.07	−0.099	0.081	0.096
1.3.08	−0.048	0.069	0.167
1.3.09	−0.008	0.021	0.026
1.3.10	−0.060	0.103	0.206
1.4.01	−0.062	0.024	0.120
1.4.02	−0.136	0.029	0.158
1.4.03	−0.125	−0.113	0.224
1.4.04	−0.151	0.046	0.193
1.4.05	−0.311	−0.031	0.310
1.5.01	−0.006	0.004	0.014
1.6.01	−0.075	0.028	0.111
1.7.01	−0.042	0.012	0.043
2.1.01	−0.147	0.033	0.182
2.1.02	−0.151	0.021	0.159
2.1.03	−0.150	0.007	0.148
2.2.01	−0.130	0.035	0.155
2.2.02	−0.045	0.018	0.063
2.2.03	−0.043	−0.043	0.079
2.2.04	−0.022	0.026	0.035
2.2.05	−0.138	−0.017	0.221
2.2.06	−0.010	0.012	0.015
2.2.07	−0.063	0.023	0.124
2.3.01	−0.188	0.110	0.201
2.3.02	−0.174	0.130	0.197
2.3.03	−0.137	0.096	0.169
2.3.04	−0.124	0.048	0.105
2.3.05	−0.115	0.050	0.180
2.3.06	0.003	−0.004	0.018
2.3.07	−0.054	0.053	0.122
2.3.08	−0.035	0.012	0.063
2.3.09	−0.063	0.047	0.067
2.3.10	−0.014	0.005	0.055
2.3.11	−0.046	−0.016	0.187
2.3.12	−0.043	0.036	0.077
2.3.13	−0.025	0.136	0.183
2.3.14	−0.081	0.076	0.154
2.4.01	−0.109	−0.197	0.221
2.4.02	−0.055	0.012	0.145
2.4.03	−0.156	−0.245	0.146
2.4.04	−0.068	−0.052	0.086
2.4.05	−0.152	0.022	0.184
2.5.01	−0.010	0.012	0.027
2.6.01	−0.046	0.022	0.062
2.6.02	−0.066	−0.003	0.075
2.7.01	−0.051	0.032	0.077
2.7.02	−0.069	0.013	0.079