Theoretical investigations of small molecule adsorption on pristine, doped, and perforated graphene

by

Sicheng Li

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 6, 2016

Keywords: quantum chemistry, intermolecular interactions, nanotubes

Approved by

Konrad Patkowski, Chair, Assistant Professor of Chemistry and Biochemistry
Vincent Ortiz, Ruth W. Molette Professor of Chemistry and Biochemistry
Rik Blumenthal, Associate Professor of Chemistry and Biochemistry
Jimmy Mills, Professor of Chemistry and Biochemistry

Abstract

Computational chemistry is a standard tool to understand chemical phenomena at the electronic and molecular levels, and is a rapidly growing branch of chemistry. In this thesis, three distinct computational chemistry projects about dispersion interaction between small molecules and models of nanotubes are investigated: (1) An accurate benchmark description of the interactions between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen; (2) Evaluation of DFT-D variants suitable for nanotube adhesion forces; (3) Description of the interactions between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen via local methods.

Chapter 1 presents a brief introduction to these subjects. Chapter 2 explains the methodologies used in this work. Chapter 3 presents the performance of a large variety of modern density functional theory approaches for the adsorption of carbon dioxide on molecular models of pyridinic N-doped graphene. The benchmark interaction energies were established at the complete-basis-set limit MP2 level plus a CCSD(T) coupled-cluster correction in a moderate but carefully selected basis set. Using a set of 96 benchmark CCSD(T)-level interaction energies as a reference, the performance of various DFT+D variants was examined. It turns out that several schemes such as B2PLYP-D3 and M05-2X-D3 exhibit average errors on the entire benchmark data set in the 5–10% range. The top DFT+D variants were then used to investigate the energy profile for a carbon dioxide transition through model N-doped graphene pores. The results obtained from these methods indicated that the largest, N4H4 pore allows for a barrierless CO_2 transition to the other side of a graphene sheet. In Chapter 4, three sets of benchmark CCSD(T)/CBS data, all involving a coronene molecule (flat or curved away from the adsorbate), interacting with **1** methane, **2** carbon dioxide, and **3** ethylene are combined to determine the performance of various DFT approaches for

describing interactions of solubilizer molecules with nanotubes. The combined data show that the simple PBE-D3(BJ)$_{\text{refit}}$ variant emerges as an optimal combination of accuracy and efficiency for weakly interacting complexes of this kind. Finally, Chapter 5 reports the results for various local methods on investigation of the dispersion interaction between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen. Unfortunately, so far no single local method could yield benchmark-level accuracy for the models of interest. Therefore, more detailed study is needed to generate acceptable results for local treatments.

Acknowledgments

There are many people who have helped me during this journey. Firstly, I would like to thank Dr. Konrad Patkowski for guiding me through my research. You are the best adviser I could ever have. I will always be grateful for your guidance, knowledge, and passion. I very much look forward to seeing the continuation of your work. I would also like to thank my committee: Dr. Vincent Ortiz who is the truly encouraging professor in quantum chemistry, Dr. Rik Blumenthal for an abundance of thermodynamical knowledge, and Dr. Jimmy Mills who gives me countless help and valuable suggestions during my study at the Department of Chemistry and Biochemistry. I am also deeply indebted to Dr. Orlando Acevedo who taught so many useful knowledge about Computational Chemistry. Next, I would like to thank Dr. Daniel Smith for your help during the whole process who trained me in many aspects, especially about how to do information searches, writing papers, and also how to write scripts. Now I fully understand that Google is our best friend, if we know how to get along with it. Plus, I want to say thank you to the people in Dr. Patkowski's group: Dr. Habib Ur Rehman, Dr. Narendra Nath Dutta, Jonathan Waldrop, Monika Kodrycka, and Reza Hemmati. Finally, I also want to express my deepest gratitude to my family for their endless support and encouragement. It has certainly been an interesting journey, thank you all for being a part of it.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	two-dimensional
AOs	atomic orbitals
BJ	Becke-Johnson
BO	Born-Oppenheimer
BSSE	basis set superposition error
CBS	complete basis set
CC	coupled cluster
CI	configuration interaction
CNTs	carbon nanotubes
CP	counterpoise correction
DF	density fitting
DFT	density functional theory
DFTB	semiempirical density functional based tight-binding method
DLPNO	domain based local pair natural orbital
FCI	full configuration interaction
FNO	frozen natural orbitals
GGA	generalized gradient approximation

GTO Gaussian-type orbitals

HF Hartree-Fock

KS Kohn-Sham

LDA local density approximation

LSDA local spin density approximation

MBD many-body dispersion

MCSCF multi-configurational self-consistent field

MOs molecular orbitals

MP Møller-Plesset

MP2 second order Møller-Plesset perturbation theory

MP2-F12 explicitly correlated MP2

MUE mean unsigned error

MURE mean unsigned relative error

N-PHACs nitrogen-containing polyheterocyclic aromatic compounds

N-PHACs nitrogen-containing polyheterocyclic aromatic compounds

NCI non-covalent interactions

PAH polycyclic aromatic hydrocarbons

PES potential energy surface

QMC Quantum Monte Carlo

RPA random-phase approximation

SCF	self-consistent field
STO	Slater-type orbitals
TS	Tkatchenko-Scheffler
WFT	wavefunction theory

Chapter 1

Introduction

Computational chemistry is a standard tool to understand chemical phenomena at the electronic and molecular levels, and is a rapidly growing branch of chemistry. It has very broad applications, from spectroscopy investigation, reaction pathways, drug design and delivery, hydrogen storage, to catalysis. In this present work, we will utilize quantum chemistry tools to investigate the properties of small molecule adsorption on pristine, doped, and perforated graphene. To begin with, we need to have an idea about adsorption. There are two types of adsorption: chemisorption and physisorption. Chemisorption, a.k.a chemical adsorption, is adsorption in which the forces involved are valence forces of the same kind as those functioning in the formation of chemical compounds, defined by IUPAC.[1] Thus, the distinction between chemisorption and physisorption and is the same as between chemical and physical interaction in general. Different from chemisorption, physisorption (or physical adsorption) is adsorption in which the forces involved are intermolecular forces, such as van der Waals forces, which do not involve a significant change in the electronic orbital patterns of the species involved. In this work, we are only interested in the harder to study, the physisorption phenomenon.

In this thesis, we are particularly interested in non-covalent interactions (NCI) between small molecules and carbon nanostructures. The question here is why are people interested in NCI? The reason for that is these interactions are involved in many aspects of chemistry such as drug binding, protein folding, supramolecular assembly, and reaction pathways. Particularly, we are interested in the NCI between small molecules and carbon nanostructures; these interactions are a hot topic of many theoretical and experimental investigations.

Specifically, we have studied carbon dioxide and ethylene physisorption onto pristine, doped, and perforated graphene.

Carbon-based nanomaterials such as carbon nanotubes and graphene have a broad range of applications in nanotechnology, electronics, optics and making high strength composite materials. Nanomaterials often have unusual mechanical strength and unique electrical properties, optical and chemical properties that can be modified by chemical doping and by noncovalent interactions with other adsorbed molecules.[2] More importantly, nanostructures can be used as a medium for separation of different gases through adsorption. This application is particularly significant for capturing carbon dioxide from exhaust gases, a vital molecule in this research as well. This separation is a crucial task on the way towards greener energy.

An *ab initio* description of the physical adsorption[3] of various molecules on carbon nanostructures is highly desirable since the accuracy of empirical adsorption potentials is limited. However, this physisorption is significantly difficult for *ab initio* electronic structure methods, especial for density functional theory (DFT), since many noncovalent interactions are dominated by London dispersion forces,[4] which are not well described by many popular theoretical methods. Thus, numerous approaches have been developed to describe the noncovalent interactions. Researches have also proposed various techniques to speed up computations so that larger systems may be studied. New DFT methods have been extensively benchmarked[5] against databases of accurate wavefunction-based interaction energies. The accuracy obtained by the best approaches (on average, about 0.3-0.5 kcal/mol at the minima), while much better than the original DFT, varies dramatically between different systems and can not guarantee an accurate description of adsorption energies that amount to a few kilocalories per mole. Therefore, it is necessary to validate the accuracy of different DFT functionals against wavefunction-based interaction energies for carefully chosen model systems before picking any DFT method to compute the noncovalent interaction energies. Only approaches that adequately pass this validation can be expected to yield accurate adsorption properties for extended nanostructures. Such a validation, and the subsequent use of

the optimal DFT method(s) to achieve accurate graphene and nanotube adsorption energies, are the foundations of this research. **The central purpose of this research is to obtain precise ab initio interaction potentials for several small molecules (CO_2, C_2H_4, CH_4) adsorbed onto models of pristine and N-doped graphene and single-walled carbon nanotubes.** To this end, we will first carefully choose a large number of relevant weakly interacting dimers as representatives. Then we are going to apply high-level *ab initio* methods on these models to generate benchmark interaction energies that are accurate to 0.1 kcal/mol or even better. These benchmark data will enable us to evaluate the accuracy of various DFT functionals in the next step and to pick the best approach that will then be applied to compute interaction energies between the adsorbate molecules and large nanotube fragments as well as infinite periodic nanotubes. The interaction energies obtained in this manner for different distances and orientations of the adsorbate molecules can be utilized to construct highly accurate potential energy surfaces. Finally, these potentials can be combined with ab initio adsorbate-adsorbate interaction potentials to build reliable adsorption isotherms for pure gases and binary mixtures as well as to explore the vibrational dynamics of adsorbed molecules.

The specific objectives of this study are:

- Firstly, to obtain accurate coupled-cluster [CCSD(T)]-level benchmarks for the lowest-energy structures of dimers of CO_2, C_2H_4, CH_4 with polycyclic aromatic hydrocarbons (PAH) and nitrogen-containing polyheterocyclic aromatic compounds (N-PHACs) containing one or more pyridinic nitrogen atoms.

- Secondly, to pick a DFT variant that recovers benchmark interaction energies accurately and consistently throughout the entire range of intermolecular distances by using the DFT-D3 scheme and then to apply this approach to compute accurate interaction energies between the chosen adsorbates and finite and infinite fragments of carbon nanotubes and N-doped nanotubes.

- Finally, in order to obtain benchmark CCSD(T)-level results for larger systems, the local coupled cluster methods will be tested.

The importance of the research in this thesis will be in several aspects. To begin with, it will provide new precise benchmarks for approximate and efficient *ab initio* approaches. Furthermore, it will contribute to the development of DFT with a proper account of dispersion. Additionally, it will lead to more reliable potential energy surfaces for molecular simulations involving carbon nanostructures and gas-phase small molecules. Finally, it will strengthen our knowledge of the physisorption phenomenon in general. As far as the energy-related research is concerned, the developments proposed here are likely to help in the design process of carbon-based nanomaterials for hydrogen and hydrocarbon storage, molecular sensing, and selective CO_2 capture for a cleaner and more efficient utilization of fossil fuels.

Chapter 2

Methods and theoretical background

Computational chemistry has developed into a standard tool to understand chemical phenomena by applying mathematical approximations and equations as well as computational software. It has broad applications. One can use it to optimize the structures and calculate properties for molecules. It can be used to simulate reactions in order to interpret and predict chemical phenomena. The fundamental basis of computational chemistry is quantum mechanics. Among many different methods, the ones that do not include any empirical or semi-empirical factors in their equations and also do not incorporate any experimental data are called *ab initio methods*. In other words, the solutions of these methods are directly derived from theoretical principles.

There are many properties that can be obtained by applying computational chemistry. Here we only name a few important ones:

- Equilibrium and transition structures
- Reaction rates and pathways
- Thermochemical properties such as bond dissociation energies and enthalpies of formation
- Polarizabilities and hyperpolarizabilities
- NMR, IR, Raman, and UV spectra

2.1 Schrödinger equation

The Schrödinger equation is the centerpiece of quantum mechanics.[6,7] This second-order partial differential equation describes how a physical system evolves with time. Its importance in quantum mechanics is an analogy of Newton's law in classical mechanics.

The Schrödinger equation can be divided into two subcategories: time-dependent and time-independent.

The time-dependent Schrödinger equation is:

$$\hat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi \tag{2.1}$$

where i is the square root of -1, $\hbar$ is Planck's constant divided by 2π, t is the time coordinate, Ψ is the wavefunction of the system and characterizes the position and momenta of the particles involved in the system,[8] and finally $\hat{H}$ is the Hamiltonian operator.

The time-independent Schrödinger equation has the form:

$$\hat{H}\Psi = E\Psi \tag{2.2}$$

For this thesis, all the solutions are within the time-independent subcategory. The exact solutions of the many-electron Schrödinger equation cannot be obtained, even for a two-electron hydrogen molecule (H_2). Therefore, we have to use several approximations to derive the solutions for the Schrödinger equation.

2.2 Born-Oppenheimer Approximation

The Born-Oppenheimer (BO) approximation is indispensable in quantum chemistry. The many electron wavefunction is a function of nuclear and electronic coordinates: $\Psi(R, r)$. Here R is nuclear coordinates while r is electronic coordinates. The motions of nuclei and electrons are coupled with each other. The BO approximation originates from the significant mass difference between nuclei and electrons: the mass of a proton is approximately two

thousand times larger than the mass of an electron. Thus, as the nuclei are much heavier, their velocities are much smaller. In this way, nuclei are almost stationary compared to electrons. In other words, the nuclei appear fixed compared to the electrons. Therefore, intuitively, the motion of electrons can be considered as circling mass points around the fixed nuclei. This is called "the Born-Oppenheimer approximation".[9] Hence, we can separate the total wavefunction into a nuclear function Ψ_N and electronic function Ψ_{el} without losing the accuracy:

$$\Psi(r, R) = \Psi_N(R)\Psi_{el}(r; R) \tag{2.3}$$

where $\Psi_{el}(r; R)$ is an electronic wavefunction that depends parametrically on the nuclear positions and $\Psi_N(R)$ is a nuclear wavefunction. Now, we can solve the electronic part of the Schrödinger equation separately.

$$\hat{H}_{el}\Psi_{el} = E_{el}\Psi_{el} \tag{2.4}$$

E_{el} represents the potential energy surface which depends on the nuclear configuration. The BO approximation is ubiquitous in quantum chemistry.

2.3 Hartree-Fock Theory (HF theory)

As one of the basic fundamental theories of quantum chemistry, the solution of the HF equations is the building block for most of the advanced *ab initio* methods which provide a more accurate description of a many-electron system.[6–8] However, one of the crucial missing pieces of the HF theory is that it neglects the electron correlation in multi-electron systems. In the HF theory, electron-electron interaction is treated in an average way, in other words, each electron only interacts with a static electron cloud of all the other electrons, instead of an interaction with real electrons. In this way, the lowest energy obtained from the HF theory will be always greater than the true energy of the system in the sense of the variational principle.

The working form for the HF theory is the Roothaan-Hall equations.[8] They can be simply written as matrix equations:

$$\mathbb{FC} = \mathbb{SCE} \tag{2.5}$$

where $\mathbb{F}$ is the Fock matrix, which plays the role of the Hamiltonian in the Schrödinger equation, $\mathbb{E}$ is the diagonal matrix of orbital energies, $\mathbb{C}$ is the coefficient matrix for the expansion of the molecular orbitals in terms of the basis functions, and $\mathbb{S}$ is the overlap matrix.[9]

The Roothaan-Hall equations can be solved by the self-consistent field (SCF) method. The procedure for SCF is: First, all one- and two-electron integrals are calculated followed by generation of an initial starting guess for the MO coefficients. Second, the initial density matrix is formed and then the formation of the Fock matrix using the core integrals and the density matrix times two-electron integrals follows. Third, the Fock matrix is diagonalized, and the eigenvectors contain the new MO coefficients. Last, the new density matrix is formed and if it is sufficiently close to the previous density matrix, then the self-consistency is achieved. Otherwise, a new Fock matrix is formed and the same procedure is repeated until the density matrix reaches convergence.

2.4 Basis set approximation

Mathematically, any set of functions can be used as a basis set.[6–8,10,11] In theoretical and computational chemistry, basis set is defined as a set of functions (called basis functions) whose linear combinations are used to create molecular orbitals (ψ_j),

$$\psi_j = \sum_{i=1}^{n} c_{ij}\chi_i \tag{2.6}$$

where the coefficients c_{ij} are also called the molecular orbital expansion coefficients, which can be determined numerically by using the variational principle. χ_i denotes an arbitrary function. Larger basis sets provide more accurate approximations to the orbitals by imposing

fewer restrictions on the locations of the electrons in space. At the same time, the energy converges towards the HF limit of the method more closely.

Slater-type orbitals (STO) and Gaussian-type orbitals (GTO)[7,10,12] are the types of orbitals used in quantum chemistry. The STOs have the functional form:

$$\psi_{\zeta,n,l,m}(r,\theta,\phi) = NY_{l,m}(\theta,\psi)r^{n-1}e^{-\zeta r} \tag{2.7}$$

where N is the normalization constant and $Y_{l,m}$ are the typical spherical harmonic functions. The parameter ζ refers to the orbital exponent determining the size of orbitals. STO orbitals are important historically. However, their application is now limited for two main reasons. A major deficiency of STOs is that they are not appropriate for numerical computations of multi-centered integrals while solving the Schrödinger equation due to the high cost in computer time. Most quantum chemistry codes use GTOs as basis functions, and they have the general form:

$$\psi_{\alpha,n,l,m}(r,\theta,\phi) = Nx^l y^m z^n e^{-\alpha r^2} \tag{2.8}$$

where the sum of l, m, and n determines the type of orbital. In this function, the dependence on r^2 in the exponential makes GTOs inferior to STOs in two ways: First of all, a GTO has zero slope at the nucleus while an STO has a cusp. As the consequence, GTOs have difficulty in describing the proper behavior near the nucleus. Secondly, GTOs fall off too rapidly far from the nucleus compared to STOs, and the whole wavefunction is therefore represented poorly, which indicates that more GTOs are needed for obtaining a certain accuracy compared to STOs. Nevertheless, GTOs have the important advantage over STOs that a GTO integral evaluation is computationally cheaper than that for STO especially for two-electron integrals. Thus, GTOs are preferred and generally used in calculations. Fortunately, a linear combination of GTOs could be used to better mimic the behavior of

an STO. This combination has the following form:

$$\Phi_\mu = \sum_{i=1}^{L} d_{i\mu} \Phi_i(\alpha_{i\mu}) \tag{2.9}$$

$d_{i\mu}$ is the expansion coefficient of the primitive Gaussian function Φ_i, which has an exponent $\alpha_{i\mu}$. L is the number of functions in the expansion. Actually the GTOs are not orbitals. They are just simpler functions and are called primitive Gaussians. For molecular calculations, these gaussian primitives have to be contracted. In other words, certain linear combinations of them will be used as basis functions. In quantum chemistry contraction means "a linear combination of gaussian primitives to be used as basis function." [12] Such a basis function will have fixed coefficients and exponents. The contractions are sometimes called Contracted Gaussian Type Orbitals. Contracted Gaussians are used most often in quantum mechanics calculations.

There are several terms associated with basis sets. The minimal basis set [7,10,12] means that one basis function is applied per occupied atomic orbital. In other words, a minimal basis set is a representation that only consists of these functions required to accommodate all of the electrons in each atom. Quantum chemists found out that as least three Gaussians are needed to properly mimic one STO. This combination is named as STO-3G. Therefore, in quantum chemistry, the minimal basis set for a calculation is STO-3G. Minimal basis sets cannot describe the wavefunction sufficiently since the number and size of the orbitals are fixed for all systems. To improve the performance of the minimal basis set, the number of basis functions for each orbital with different orbital exponents has been doubled and tripled, called Double-Zeta and Triple-Zeta, respectively. In chemical bonding, the core orbitals only weakly affect bonding properties while valence orbitals are very important. Furthermore, split valence bases are used to add additional flexibility for describing valence orbitals. On the other hand, diffuse basis functions are added to deal with systems that allow electrons to flow away from the nucleus, for instance, for excited states and anions.

Another modification of basis sets is to make the atomic orbitals polarized by the influence of their surroundings. The polarization functions in basis sets are usually denoted as * or (d), for example, 6-31G(d). Finally, one should mention Dunning's correlation-consistent basis sets.[13] Thom Dunning pointed out that basis sets optimized at the HF level might not be sufficient for correlated computations. Dunning's basis sets are used dominantly in this thesis. The first group of them is the cc-pVXZ family, which means a Dunning correlation-consistent, polarized valence, X-zeta basis; where X=D,T,Q,5,6,7. When a prefix "aug" is added, it means one set of diffuse functions is added for every angular momentum present in the basis. The Dunning basis sets are designed to converge smoothly toward the complete basis set (CBS) limit.

2.5 Electron correlation methods

The deficiency of the HF method is the lack of instantaneous correlation of the movements of electrons, which in turn lowers the total electron-electron potential energy. In this sense, the exact energy of the molecule will always be lower than that calculated by the HF method.

$$E_{\text{exact}} = E_{\text{HF}} + E_{\text{corr}} \tag{2.10}$$

Since the HF method does not include electron correlation, many theoretical methods have been proposed in order to capture electron correlation properly. There are two categories of electron correlation. The first category is dynamical correlation, which is mainly caused by the instantaneous repulsion of the electrons. In the HF method, electrons often get too close to each other, because the electrostatic interaction is treated in only an average manner. As a result, the electron-electron repulsion term is too large resulting in higher energy. It is called dynamical electron correlation because it is related to the actual movements of the individual electrons. The dynamical correlation can be treated by perturbation theory,[7,10,12] coupled cluster (CC) theory,[7,10,12,14] and density functional theory (DFT).[15] The

second one is non-dynamical or static correlation. It is related to the fact that in certain circumstances a single Slater determinant is not a good approximation to the true ground state, because there are other Slater determinants with comparable energies. On the other hand, non-dynamical correlation can be treated by the multi-configurational self-consistent field (MCSCF) method[16] via a linear combination of determinants. Electron correlation methods are referred to as post-HF methods since the correlation correction terms are added to the basic HF energy.

2.5.1 Full configuration interaction

The full configuration interaction (FCI) method is the most complete treatment of the molecular system possible within the finite set of basis functions. In CI, the exact wavefunction is a linear combination of determinantal wavefunctions, each of which corresponds to an electronic state of the molecule. In this way, the ground electronic state is represented as a mixture of interacting electronic configurations.

$$\Psi_{\text{exact}} = c_0\Psi_0 + c_1\Psi_1 + c_2\Psi_2 + ... \tag{2.11}$$

where Ψ_0 is the ground electronic configuration of the Fock operator, Ψ_1 the first electronically excited configuration, Ψ_2 the second, etc., and the coefficients c refer to the contribution of that particular electronic configuration to the full consideration of the wavefunction. These coefficients are variationally optimized.

2.5.2 Møller-Plesset (MP) perturbation theory

Møller-Plesset (MP) perturbation theory is one of the most economical electron correlation methods.[7,10,12] In the MP scheme, the exact Hamiltonian operator is represented as

$$H = H_0 + \lambda V \tag{2.12}$$

where H_0 is the Fock operator and λV is a perturbation operator applied to H_0. The parameter λ may vary from 0 to 1. The Rayleigh-Schrödinger perturbation theory offers a description of the perturbed system as:

$$H\Psi = (H_0 + \lambda V)\Psi = E\Psi. \tag{2.13}$$

Here, the wave function and energy can be expanded in a power series of λ.

$$\Psi_\lambda = \lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i \Psi^{(i)} \tag{2.14}$$

$$E_\lambda = \lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i E^{(i)} \tag{2.15}$$

The value of i determines the order of correction. Due to the formalism of the MP method, the electron correlation energy is not included until $n = 2$, which is corresponding to second-order Møller-Plesset perturbation theory (MP2). The wavefunction and energy are therefore represented as a power series.

$$\Psi_{\mathrm{MP2}} = \Psi^{(0)} + \lambda\Psi^{(1)} \tag{2.16}$$

$$E_{\mathrm{MP2}} = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} \tag{2.17}$$

Substitution of these terms into the Schrödinger equation gives:

$$H_0\Psi^{(2)} + V\Psi^{(1)} = E^{(0)}\Psi^{(2)} + E^{(1)}\Psi^{(1)} + E^{(2)}\Psi^{(0)} \tag{2.18}$$

Premultiplying by $\Psi^{(0)}$ and integrating over all space yields the following forms for the energy.

$$E^{(0)} = \left\langle \Psi^{(0)} | H_0 | \Psi^{(0)} \right\rangle \tag{2.19}$$

$$E^{(1)} = \left\langle \Psi^{(0)} | V | \Psi^{(0)} \right\rangle \tag{2.20}$$

$$E^{(2)} = \left\langle \Psi^{(0)} | V | \Psi^{(1)} \right\rangle \tag{2.21}$$

$E^{(0)} + E^{(1)}$ is the Hartree-Fock energy. $E^{(2)}$ is the first approximation to the electron correlation energy, expanded into the following form:

$$E_0^{(2)} = -\frac{1}{4} \sum_{ab}^{virt} \sum_{ij}^{occ} \frac{|\langle ab||ij\rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{2.22}$$

where i and j are the occupied orbitals, a and b are the unoccupied virtual orbitals, ϵ denotes the orbital energy, and $\langle ab||ij\rangle$ are the antisymmetrized electron repulsion integral in physicist's notation, defined as:

$$\langle ab||ij\rangle = \int\int \psi_a^*(1)\psi_b^*(2)\frac{1}{r_{12}}[\psi_i(1)\psi_j(2) - \psi_j(1)\psi_i(2)]d\tau_1 d\tau_2 \tag{2.23}$$

2.5.3 Coupled cluster theory

The central equation of coupled cluster (CC) theory is

$$\Psi = e^{\hat{T}}\Phi_0 \tag{2.24}$$

where Ψ is the exact ground state wavefunction, Φ_0 is the ground state HF reference wavefunction. CC theory uses the exponentiated $\hat{T}$ excitation operator

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ... \tag{2.25}$$

where the subscript represents the number of electrons each operator will excite. As the complexity of the CC expansion is N^{2n+2}, where N is the number of total basis functions and n is the largest possible excitation level, it is intuitive to limit the overall number of the excitations. This truncation leads to a hierarchy of CC methods, for example, $\hat{T} \equiv \hat{T}_1 \rightarrow$

CCS and $\hat{T} \equiv \hat{T}_1 + \hat{T}_2 \rightarrow$ CCSD. Where "S", "D", "T", "Q", ... denotes the singles ($\hat{T}_1$), doubles ($\hat{T}_2$), triples ($\hat{T}_3$), and quadruples ($\hat{T}_4$), etc. When the number of excitations n is the number of electrons in the system, then the CC method is identical to FCI, for instance, for a four electron system CCSDTQ $\equiv$ FCI.

A set of non-linear equations produced from the CC method must be solved iteratively. If CCn is obtained iteratively, CC(n + 1) can be computed perturbatively. This leads to the famous CCSD(T) method where CCSD is solved iteratively and then (T) is achieved perturbatively.

2.5.4 Density functional theory

So far, all the methods we have been covering are called wavefunction methods. The energies and properties obtained are based on solving for the electron wavefunction. Another strategy for treating electron correlation is the popular density functional theory (DFT). DFT is based on the theorem by Hohenberg and Kohn,[15] which states that the ground state molecular properties can be obtained from the ground state electron probability density ρ. The electron density corresponding to a normalized N-electron wavefunction can be defined as

$$\rho(r) = N \sum_{s_1} ... \sum_{s_1} \int dr_2 ... \int dr_N |\Phi(r_1, s_1, r2, s2, ...r_N, s_N)|^2, \tag{2.26}$$

where r and s are the spatial and spin variables, respectively.

Because there is a one-to-one mapping between the ground state ρ and the ground state energy, the latter is obtained once the former is found. The primary molecular property of interest in these studies is the energy and nuclear geometry of stationary points on the potential energy surface (PES). The ground state energy E is a functional of ρ and the total energy can be described as:

$$E[\rho] = T_s[\rho] + V_{ext}[\rho] + J[\rho] + E_{xc}[\rho]. \tag{2.27}$$

Where $T_s[\rho]$ is the kinetic energy of a set of n independent electrons, moving in an effective electron potential which leads to the density $\rho(\vec{r})$, $V_{ext}[\rho]$ is the potential energy in the field of the nuclei plus any other external perturbation, and $J[\rho]$ is the total Coulomb interaction energy. The last term is the exchange-correlation energy ($E_{xc}[\rho]$). This term is the key problem in DFT,[17] since the exact form of the $E_{xc}[\rho]$ functional is unknown, thus approximations must be used. Various approximate functionals for $E_{xc}[\rho]$ exist in DFT calculations. However, the electron correlation in DFT is not systematically improvable.[10] Approximations to the $E_{xc}[\rho]$ are discussed below.

The simplest approximation is the local density approximation (LDA). LDA assumes that individual volume elements of the system have the density of a uniform electron gas and that the electron density varies smoothly throughout the entire volume. In the scheme of LDA, the $E_{xc}[\rho]$ term is represented as the interaction between the electron density $\rho(r)$ and $\epsilon_{xc}[\rho(r)]$, the per-electron exchange-correlation energy, integrated over the whole space.

$$E_{xc}^{LDA}[\rho(r)] = \int \rho(r)\epsilon_{xc}[\rho(r)]dr \tag{2.28}$$

Every point r is surrounded by a volume element dr of constant electron density. While this approximation is credited for the early success of DFT, it often gives unsatisfactory results in chemical applications. Thus, several corrections for the non-uniformity of atomic and molecular density have been constructed. For instance, LDA performs well for metallic systems, but overbinds organic compounds.[18]

The local spin density approximation (LSDA) divides the density into spin-up and spin-down and defines the net spin density as the difference between the two.[19]

$$\sigma(r) = \rho_{\uparrow}(r) - \rho_{\downarrow}(r) \tag{2.29}$$

LSDA is good for bond lengths and vibrations, however, it overestimates dipole moments and molecular binding energies while underestimating reaction barrier heights.[8,20,21]

The next generation of exchange-correlation approximations is the generalized gradient approximation (GGA). This approximation has been embraced due to its simplicity. In this approximation, not only the electron density at a certain point but also its gradient are taken into consideration. GGA functionals correct many of the drawbacks of the LSDA approximation and produce better thermochemical predictions but still underestimate reaction barriers.[17] In order to further improve the performance of GGA, the next step would be the inclusion of second derivative corrections, which are termed as meta-GGA functionals. In these methods, the exchange-correlation potential is dependent on the kinetic energy density, τ, formulated as

$$\tau(r) = \sum_{i=1}^{N} \frac{1}{2} |\nabla \psi_i(r)|^2 \tag{2.30}$$

where ψ_i is solved from a Kohn-Sham calculation.

A further improvement kicks in by hybridizing HF and DFT methods,[22] in which a HF exchange term and DFT exchange functionals are mixed in conjunction with a correlation functional. This model depends on a linear combination of HF exchange with density functional exchange-correlation contributions:

$$E_{xc}^{\text{hybrid}} = a_{x0} E_{xc}^{\text{LDA}} + (1 - a_{x0}) E_x^{\text{HF}} + a_{x1} \Delta E_x^{\text{GGA}} + E_c^{\text{LDA}} + a_c \Delta E_c^{\text{GGA}}. \tag{2.31}$$

Where the three semiempirical parameters a_{x0}, a_{x1}, and a_c can be determined by fitting the heats of formation of a standard set of molecules. A very famous example is the B3LYP[23,24] method, which uses the Becke 88 (B)[25] exchange functional together with the Lee-Yang-Parr (LYP) correlation functional.

$$E_{xc}^{\text{B3LYP}} = a_{x0} E_{xc}^{\text{LDA}} + (1 - a_{x0}) E_x^{\text{HF}} + a_{x1} \Delta E_x^{B} + E_c^{\text{LDA}} + a_c \Delta E_c^{\text{LYP}}. \tag{2.32}$$

Due to the incorporation of non-local Hartree-Fock exchange into a semi-local exchange-correlation of DFT, such functionals are referred to as hybrid functionals. These functionals

are usually more accurate than the purely local functionals for main group thermochemistry.[17]

Finally, we have double-hybrid functionals, which include a portion of the MP2 correlation energy as computed with Kohn-Sham (KS) orbitals. The example of a doubly-hybrid functional in this thesis is B2PLYP.[26] Among these DFT methods, only the double-hybrid functionals include long-range correlation; that is the reason any GGA or hybrid-GGA is lacking in dispersion.

2.6 Dispersion in DFT

DFT is popular and used in many quantum chemistry calculations, however, it has a very poor description of dispersion.[27,28] Therefore, a wide variety of new methods have been proposed to overcome the inherent inability of DFT to tackle the lack of long-range correlation[17,29–35] and these new methods provide enormous improvement over standard GGA or hybrid functionals when it comes to noncovalent interaction energies. These newly developed methods reach well into the 0.2–0.3 kcal/mol range of accuracy,[5] as averaged over popular databases of weak interaction energies,[36–38] in the van der Waals minimum region. This performance has long surpassed chemical accuracy (1 kcal/mol). Nevertheless, this accuracy (which corresponds to a relative accuracy of about 10% for the systems considered here) is by no means consistent across either different systems or different intermolecular separations. Therefore, careful benchmarking of different DFT functionals against accurate wavefunction-based interaction energies for relevant models to choose an optimal variant on a case-by-case basis is necessary.

There are four areas of approaches to treat dispersion in DFT, getting steadily more accurate (and hence expensive) as we progress:[27]

- ground level. The ground level for DFT-based dispersion correction schemes is these methods without long range asymptotics. In general, functionals without long-range asymptotics cannot describe the dispersion interaction of well separated molecules.

However, the "Minnesota functionals"[39] are an example of a new breed of functionals that are fitted to a dataset such that they can describe binding energies accurately at separations around minima, but they cannot provide an accurate picture for a long range separation.

- semi-empirical, pair-wise methods (excluding environment) employing the C_6 terms, DFT-D2[29] is among this group. For almost all these pair-wise methods, they introduce scaling factors. A short-range scaling is applied to turn off dispersion when atoms are close to each other. DFT-D2 is the most famous out of semi-empirical approaches. They fit the C_6 terms to *ab initio* or experimental data.

- approaches that go beyond semi-empirical and introduce environment-dependent C_6 coefficients and also some *ab initio* information. DFT-D3,[34] the Becke-Johnson (BJ) model (XDM),[40] and Tkatchenko-Scheffler (TS)[32] are among this group. Including the environment dependence and *ab initio* data makes these methods significantly more transferable. The TS method begins with C_6 coefficients found for interactions between free atoms, computed with self-interaction corrected time-dependent density functional theory.

- density functional approaches based on the vdW-DF-04[41] method of Langreth and Lundqvist (and co-workers), which is to take a double integral over a pair of points in space, and at each point calculate the product of the charge densities at the two points and a kernel which depends on both points. The recently developed vdW functional is the VV10[35] approach of Vydrov and Van Voorhis.

- and approaches which go beyond pair-wise additivity, such as many-body dispersion (MBD)[42] and the random-phase approximation (RPA).[43] The TS method is extended to include many-body dispersion by representing the interactions between atomic densities in terms of quantum harmonic oscillators (QHO).[44,45]

It should be noticed that for extended systems such as graphene and nanotubes, experiments may provide useful information about the strength of noncovalent interactions like dissociation energies (D_0) and adsorption enthalpies, however, direct information on their nature is often not available. Therefore, it is necessary to rely on theory and high accuracy calculations to understand the noncovalent interactions among them. One of the promising ways to generate a precise description of these interactions is the fixed-node diffusion Monte Carlo (DMC/FN-DMC) method,[46] which is a member of the quantum Monte Carlo (QMC) class.[47,48]

2.7 Statistics

In this thesis, two most used statistics are mean unsigned error (MUE),

$$\mathrm{MUE} = \frac{\sum_{i=1}^{N} |E^i - E^i_{\mathrm{ref}}|}{N} \tag{2.33}$$

and mean unsigned relative error (MURE)

$$\mathrm{MURE} = \frac{\sum_{i=1}^{N} (|\frac{E^i - E^i_{\mathrm{ref}}}{E^i_{\mathrm{ref}}}|)}{N} \times 100\%. \tag{2.34}$$

Chapter 3

An accurate benchmark description of the interactions between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen

3.1 Abstract

We assessed the performance of a large variety of modern density functional theory approaches for the adsorption of carbon dioxide on molecular models of pyridinic N-doped graphene. Specifically, we selected eight polyheterocyclic aromatic compounds ranging from pyridine and pyrazine to 1,6-diazacoronene and investigated their complexes with CO_2 for a large range of intermolecular distances and including both in-plane and stacked orientations. The benchmark interaction energies were computed at the complete-basis-set limit MP2 level plus a CCSD(T) coupled-cluster correction in a moderate but carefully selected basis set. Using a set of 96 benchmark CCSD(T)-level interaction energies as a reference, we investigated the accuracy of DFT-based approaches as a function of the density functional, the dispersion correction, the basis set, and the counterpoise correction or lack thereof. While virtually all DFT variants exhibit some deterioration of accuracy for distances slightly shorter than the van der Waals minima, we were able to identify several schemes such as B2PLYP-D3 and M05-2X-D3 whose average errors on the entire benchmark data set are in the 5–10% range. The top DFT performers were subsequently used to investigate the energy profile for a carbon dioxide transition through model N-doped graphene pores. All investigated methods confirmed that the largest, N4H4 pore allows for a barrierless CO_2 transition to the other side of a graphene sheet.

3.2 Introduction

Novel carbon-based nanomaterials such as graphene and carbon nanotubes exhibit a wide range of mechanical and electronic properties and have been proposed for a variety of applications. Even more diversity, in particular, both n-type and p-type semiconductor character, can be introduced by doping pristine nanostructures with heteroatoms such as nitrogen, boron, or oxygen. The nitrogen-doped structures, the subject of this work, are of particular importance. The microscopic structure of N-doped carbon nanotubes has been studied using both experiment and molecular simulations[49–54] and two local patterns around dopant atoms have emerged. The first one, the "graphitic" structure, involves a simple substitution of a nitrogen atom in place of one of the carbon atoms. This structure involves an unpaired electron that ends up in a delocalized $\pi^\star$ state.[52,55] The second, "pyridine-like" pattern, which will be the focus of the present study, involves 2–4 sp^2 nitrogen atoms in pyridinic rings surrounding a vacancy. The three-nitrogen vacancy is a particularly popular model and the presence of such vacancies has been confirmed by scanning tunneling microscopy.[49] Slightly larger vacancies are interesting as potential "holes" for small molecules to enter and exit the nanotubes or to pass through a porous graphene membrane.[56–68]

Noncovalent interactions of graphene and carbon nanotubes with adsorbed molecules are of broad significance[2] due to the proposed applications of nanotubes in chemical sensing[69] and gas storage and separation[70] as well as the possibilities of tuning nanotube properties via noncovalent functionalization.[71] In particular, the separation of carbon dioxide from flue and exhaust gases through selective adsorption is one of the most promising ways to reduce global carbon emissions, and carbon-nanotube based materials, while not as effective as the most recent generations of metal organic frameworks,[72] provide a viable medium for CO_2 sequestration.[73] It should be noted that the interaction of carbon nanotubes (pristine or doped) with CO_2 and other adsorbed small molecules is dominated by dispersion and thus provides a difficult target for *ab initio* computational chemistry, most notably for methods based on density functional theory (DFT). A wide variety of new methods have been devised

to overcome the inherent inability of DFT to account for long-range correlation[17,29–35] and these new methods provide enormous improvement over standard generalized gradient approximation (GGA) or hybrid functionals when it comes to noncovalent interaction energies. As averaged over popular databases of weak interaction energies,[36–38] the novel DFT variants have long surpassed chemical accuracy (1 kcal/mol) in the van der Waals minimum region, reaching well into the 0.2–0.3 kcal/mol range.[5] However, this accuracy, which corresponds to a relative accuracy of about 10% for the systems considered here, is by no means consistent across either different systems or different intermolecular separations, and one has to select an optimal DFT variant on a case-by-case basis, through careful benchmarking against accurate wavefunction-based interaction energies for relevant models. We have previously performed such benchmarking and functional selection for models of graphene and pristine carbon nanotubes interacting with methane[74,75] and carbon dioxide.[76]

For the interaction of a CO_2 molecule with N-doped graphene and carbon nanotubes, a natural class of models are dimers of CO_2 with nitrogen-containing polyheterocyclic aromatic compounds (N-PHACs) containing one or more pyridinic nitrogen atoms. Such dimers, along with similar 1- and 2-ring complexes (involving, e.g., the purine molecule) were the subject of a high-level (up to the supermolecular coupled-cluster method with single, double, and perturbative triple excitations, CCSD(T)) computational study of Vogiatzis *et al.*[77] These authors identified the global-minimum structures for 13 N-PHAC–CO_2 dimers and obtained benchmark interaction energies close to the CCSD(T) complete basis set (CBS) limit. The study of Ref. 77 was later employed by Mackie and DiLabio[78] as a benchmark for their DFT-based study of the CO_2 adsorption on N-doped carbon nanotubes. However, none of the models considered in Ref. 77 had more than two aromatic rings, while our methane adsorption study[74] indicated that the one- and two-ring aromatic fragments provide quite poor models of extended carbon nanostructures.

In this paper, we attempt to find the best-performing DFT variant for the CO_2 adsorption on graphene and carbon nanotubes N-doped into a pyridinic structure. To this

end, we first select a set of sixteen benchmark N-PHAC–CO_2 structures that cover both the global-minimum, planar geometries and the three-dimensional stacked structures. The N-PHACs present in the benchmark set range from one ring (pyridine, pyrazine) to seven rings (1,6-diazacoronene). The coronene-sized models are the largest ones for which reliable benchmark interaction energies, accurate to below 0.1 kcal/mol at the van der Waals minima, can be obtained using the composite MP2/CBS+ΔCCSD(T) (CBS-extrapolated second-order Møller-Plesset perturbation theory plus a CCSD(T) correction in a moderately sized basis set) approach.[74,75] Alternatively, one could use some approximate, local coupled-cluster methods[79–82] but we have not pursued this approach here as it is not clear how the associated approximations affect the uncertainty of the benchmark. We consider six different intermolecular separations along a one-dimensional cut through the potential energy surface (PES) for each structure, resulting in 96 data points for which benchmark MP2/CBS+ΔCCSD(T) interaction energies are computed. A variety of DFT functionals with different atom-pairwise dispersion corrections and basis sets are then compared to this benchmark in order to find a top performer. This top performer is then used to obtain CO_2 adsorption energies on larger N-PHAC models, in particular, to model the barrier to the CO_2 transition through three different vacancies in an N-doped graphene sheet.

3.3 Methods and Computational Details

The MOLPRO[83] program was used to obtain all conventional and explicitly correlated MP2 and CCSD(T) interaction energies. The MP2 calculations employed density fitting (DF-)[84] and used standard orbital and auxiliary bases aug-cc-pVXZ≡aXZ[13,85] and aug-cc-pVXZ/MP2FIT,[86,87] respectively. To keep the errors of the DF approximation under control, we needed to resort to conventional, non-density-fitted Hartree-Fock calculations in the aDZ and aTZ bases. Moreover, for the explicitly correlated MP2 (MP2-F12) calculations, the `df-basis-exch` and `ri-basis` auxiliary bases were chosen as the aug-cc-pVXZ/JKFIT sets instead of the MOLPRO default non-augmented sets. We found that this change in auxiliary

bases was critical for the accuracy of MP2-F12 but had a negligible effect on the ΔCCSD(T) term so we kept the default auxiliary bases in the ΔCCSD(T)-F12 calculations. As the cost of DF-MP2 is just a small fraction of that of conventional MP2, only DF-MP2 calculations are feasible in the aQZ and a5Z bases for larger N-PHACs. The "DF-" qualifier will be dropped from now on. Unless noted otherwise, all computations utilized the counterpoise (CP) correction for the basis set superposition error (BSSE).[88,89] The 1s carbon, nitrogen, and oxygen electrons were not correlated.

3.3.1 Geometries of the Model Complexes

To obtain the lowest-energy geometry for each dimer, first, a geometry optimization for all N-PHAC and CO_2 monomers was performed at the MP2/aTZ level. The dimer was then optimized at the MP2/aTZ level with only the intermolecular degrees of freedom allowed (in other words, the intramolecular degrees of freedom were frozen). In order to consider different modes of interaction between CO_2 and N-PHACs, two arrangements, one with CO_2 stacked over the N-PHAC surface and parallel to it, and the other with the CO_2 carbon in the same plane as the N-PHAC, are taken into consideration. The parallelity of the CO_2 molecule to the N-PHAC plane (stacked structures) as well as the maximum point-group symmetry were forced during the optimization, but the intermolecular distance and the angles that do not affect symmetry and parallelity were freely optimized. The global minimum position for all complexes occurs when the CO_2 carbon lies in the N-PHAC plane. In addition to the minimum geometries, we computed interaction energies along radial cross sections through the potential energy surfaces passing through the minimum and lowest-energy stacked geometries. In other words, we shifted the distance z between the monomers with all the angles fixed relative to the line connecting the monomers. The values of z given throughout the rest of the text are the distances between the CO_2 carbon and the closest nitrogen for the in-plane dimers (except phenanthroline-CO_2, where the z value is the distance from the CO_2 carbon to the midpoint between the two closest carbon atoms in the

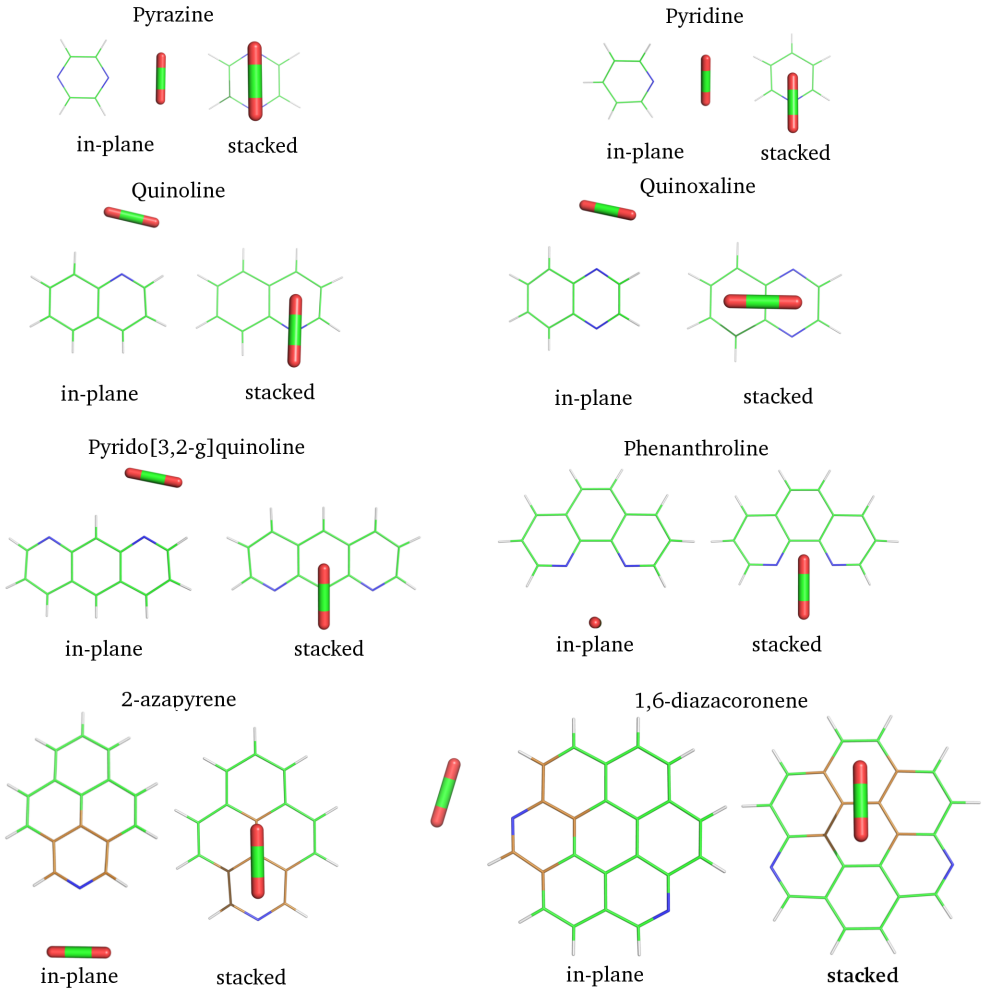

Figure 3.1: The structures of the model CO_2–N-PHAC complexes. The in-plane configurations represent the global minima, "stacked" are the related 3D stacked structures. The bronze-colored carbon atoms in the N-PHAC molecule as well as the closest nitrogen atom (blue) are those that have diffuse functions in the laDZ basis.

middle ring of phenanthroline). For all the stacked geometries, the z values are the distances between the CO_2 carbon and the N-PHAC plane. The lowest-energy in-plane and stacked geometries for each dimer are displayed in Fig. 3.1. The stacked configuration was picked to be parallel since all resulting geometries have at least C_s symmetry, lowering the overall computational cost compared to other possible stacked structures. The only exception is the quinoline-CO_2 complex for which, as shown in Fig. 3.1, the stacked configuration cannot

have C_s symmetry. However, we still forced the CO_2 molecule to be parallel to the quinoline plane to get a consistent set of sixteen geometries.

3.3.2 Benchmark Energies from Wave-Function Methods

Following the standard practice in the field,[90] the benchmark interaction energy is computed as

$$E_{\rm int}^{\rm benchmark} = E_{\rm int}^{\rm MP2}/({\rm a}(X{\rm –}1){\rm Z},\, {\rm a}X{\rm Z}) + \Delta E_{\rm int}^{\rm CCSD(T)}/({\rm a}(X'{\rm –}1){\rm Z},\, {\rm a}X'{\rm Z}), \tag{3.1}$$

where $E_{\rm int}^X = E_{\rm AB}^X - E_{\rm A}^X - E_{\rm B}^X$ is the supermolecular interaction energy at a given level of theory, $\Delta E_{\rm int}^{\rm CCSD(T)} = E_{\rm int}^{\rm CCSD(T)} - E_{\rm int}^{\rm MP2}$ is the CCSD(T) contribution missing at the MP2 level, and the notation (basis1, basis2) represents that the bases "basis1" and "basis2" have been utilized in the standard X^{-3} extrapolation for the correlation part of the interaction energy.[91] The X^{-3} extrapolation has been employed for both the conventional and explicitly correlated MP2 and CCSD(T) contributions.[92,93] The self-consistent field (SCF) part of the interaction energy was taken from the calculation employing the larger of the two bases and not extrapolated. The notation ΔCCSD(T)/aXZ indicates a correction that is calculated in the aXZ basis set without extrapolation. Moreover, MP2/$(X-1, X)$ and ΔCCSD(T)/$(X-1,\, X)$ will be the short-hand notations for $E_{\rm int}^{\rm MP2}$/(aug-cc-pV$(X-1)$Z, aug-cc-pVXZ) and $\Delta E_{\rm int}^{\rm CCSD(T)}$/(aug-cc-pV$(X-1)$Z, aug-cc-pV$X$Z), respectively.

To investigate the basis set convergence of the ΔCCSD(T) contribution, explicitly correlated CCSD(T)-F12 calculations were performed for seven 1-ring and 2-ring dimers (except for the C_1 stacked quinoline-CO_2 complex). The CCSD(T)-F12a and CCSD(T)-F12b approximations[94,95] use the default MOLPRO[83] values for the explicitly correlated *Ansätze*, geminal parameters, and auxiliary bases. Since the triples contributions to CCSD(T)-F12a and CCSD(T)-F12b do not include explicit correlation (an explicitly correlated (T) correction has been derived[96] but exhibits a steeper computational scaling), we employed the

popular estimate of the missing F12 contributions to $\Delta E^{(\mathrm{T})} = E^{\mathrm{CCSD(T)-F12}} - E^{\mathrm{CCSD-F12}}$ via scaling:

$$\Delta E^{(\mathrm{T}**)} = \Delta E^{(\mathrm{T})} \cdot \frac{E_{\mathrm{corr}}^{\mathrm{MP2-F12}}}{E_{\mathrm{corr}}^{\mathrm{MP2}}}, \tag{3.2}$$

where the subscript "corr" represents the correlation energy at a given level of theory. To ensure size consistency, the scaling factor calculated for the dimer was also employed in the counterpoise-corrected calculations for the monomers.[97] Throughout this work, the notations "(T)" and "(T**)" will refer to the unscaled and scaled triples corrections in CCSD(T)-F12, respectively. The CCSD(T)-F12 approach in its various approximate variants provides greatly improved weak interaction energies in double- and triple-ζ basis sets compared to conventional CCSD(T) results.[97–99]

3.3.3 DFT Calculations

Among the many new variants of DFT that include some form of dispersion, the three groups tested in this work are the functionals specifically optimized for benchmark weak interaction energies, the DFT+D approaches with an atom-pairwise dispersion correction added on top of a standard density-functional calculation, and the double hybrid DFT functionals. In this work, we examined a few representative members of each group. For the first group, the interaction-optimized functionals, we included M05-2X[100] and M06-2X.[17] Among the second group, DFT+D, we employed the widely popular B3LYP,[23,24] BLYP,[25] BP86,[25] PBE,[101] PBE0,[102,103] and LC-ωPBE[104] methods as well as Grimme's reparameterization[29] of Becke's B97 functional.[89] These five functionals were augmented by Grimme's atom-pairwise dispersion terms in the -D2,[29] -D3,[34] -D3(BJ),[105] -D3-E$^{(3)}$, and -D3(BJ)-E$^{(3)}$ variants. For the last group, we tested the B2PLYP[26] double hybrid functional.

All DFT interaction energies except for B2PLYP and LC-ωPBE were calculated using MOLPRO 2012.1[83] locally modified to include Grimme's reparameterization of B97. The requested energy convergence threshold for MOLPRO calculations was at least 10^{-7} hartree (10^{-8} for the M05-2X and M06-2X functionals which are known to exhibit particularly slow

convergence with respect to the integration grid[106,107]), and the corresponding autogenerated MOLPRO grids were used. All calculations employed density fitting with the default auxiliary basis sets[108] in MOLPRO 2012. The -D2, -D3, -D3(BJ), -D3-E$^{(3)}$, and -D3(BJ)-E$^{(3)}$ corrections were computed using Grimme's DFTD3 program V3 Rev. 2. The LC-ωPBE and B2PLYP interaction energies were calculated by the PSI4 code,[109] employing density fitting with the default PSI4 auxiliary basis sets. For the PSI4 calculations, the default 10^{-6} hartree energy convergence threshold and the Lebedev-Treutler (75,302) grid were utilized. As the CP correction is by no means guaranteed to improve DFT results, all DFT variants were tested both with and without this correction.

3.4 Results and Discussion

3.4.1 Benchmark Interaction Energies

Table 3.1: The MP2 and ΔCCSD(T) interaction energy contributions (in kcal/mol) for the lowest-energy structures of the pyrazine-CO_2 and quinoxaline–CO_2 complexes. The rows marked "ext." display the CBS-extrapolated results — the value in the aXZ column was obtained from the $(X-1, X)$ extrapolation.

	in-plane pyrazine-CO_2				stacked pyrazine-CO_2				in-plane quinoxaline-CO_2				stacked quinoxaline-CO_2			
method	aDZ	aTZ	aQZ	a5Z	aDZ	aTZ	aQZ	a5Z	aDZ	aTZ	aQZ	a5Z	aDZ	aTZ	aQZ	a5Z
MP2	-3.328	-3.667	-3.786	-3.838	-1.208	-1.544	-1.631	-1.672	-3.837	-4.221	-4.361	-4.420	-3.296	-3.765	-3.920	-3.987
ext.		-3.811	-3.871	-3.886		-1.686	-1.698	-1.711		-4.388	-4.459	-4.474		-3.969	-4.032	-4.051
MP2-F12	-3.851	-3.902	-3.893	-3.896	-1.635	-1.696	-1.705	-1.711	-4.439	-4.481	-4.481	-4.482	-3.985	-4.029	-4.043	-4.055
ext.		-3.925	-3.888	-3.898		-1.722	-1.712	-1.717		-4.501	-4.482	-4.483		-4.048	-4.053	-4.068
ΔCCSD(T)	0.082	0.043	0.037		0.468	0.503	0.518		0.113	0.071			1.097	1.139		
ext.		0.026	0.032			0.517	0.528			0.054				1.157		
ΔCCSD(T)-F12a	0.104	0.069	0.054		0.521	0.534	0.536		0.140	0.103			1.188	1.192		
ext.		0.054	0.044			0.540	0.537			0.087				1.194		
ΔCCSD(T**)-F12a	0.011	0.033	0.038		0.374	0.480	0.511		0.018	0.057			0.947	1.104		
ext.		0.043	0.041			0.525	0.533			0.073				1.170		
ΔCCSD(T)-F12b	0.198	0.107	0.075		0.644	0.580	0.557		0.254	0.148			1.400	1.270		
ext.		0.069	0.051			0.554	0.540			0.104				1.215		
ΔCCSD(T**)-F12b	0.105	0.071	0.058		0.496	0.526	0.532		0.131	0.102			1.159	1.182		
ext.		0.058	0.049			0.538	0.536			0.090				1.191		
ΔCCSD(T)-F12avg	0.104	0.070	0.056		0.509	0.530	0.534		0.135	0.102			1.173	1.187		
ext.		0.056	0.046			0.539	0.537			0.088				1.193		
CCSD(T)/aXZ	-3.245	-3.623	-3.755		-0.735	-1.040	-1.118		-3.723	-4.149			-2.192	-2.624		
CCSD(T)/$(X-1, X)$		-3.782	-3.850			-1.168	-1.174			-4.328				-2.806		
CCSD(T)-F12avg/aXZ	-3.747	-3.832	-3.837		-1.126	-1.166	-1.172		-4.303	-4.379			-2.811	-2.842		
CCSD(T)-F12avg/$(X-1, X)$		-3.868	-3.841			-1.183	-1.175			-4.411				-2.855		
MP2/(Q,5)+ ΔCCSD(T)/aXZ	-3.804	-3.843	-3.850		-1.243	-1.209	-1.194		-4.361	-4.403			-2.955	-2.911		
MP2/(Q,5)+ ΔCCSD(T)/$(X-1, X)$		-3.860	-3.854			-1.194	-1.183			-4.420				-2.894		
MP2/(Q,5)+ ΔCCSD(T)-F12avg/aXZ	-3.782	-3.816	-3.830		-1.203	-1.181	-1.178		-4.339	-4.372			-2.878	-2.864		
MP2/(Q,5)+ ΔCCSD(T)-F12avg/$(X-1, X)$		-3.830	-3.840			-1.172	-1.175			-4.386				-2.859		
MP2-F12/(Q,5)+ ΔCCSD(T)-F12avg/aXZ	-3.794	-3.828	-3.842		-1.208	-1.187	-1.183		-4.347	-4.381			-2.894	-2.881		
MP2-F12/(Q,5)+ ΔCCSD(T)-F12avg/$(X-1, X)$		-3.842	-3.852			-1.177	-1.180			-4.394				-2.876		

In this section we present how the benchmark wave-function-based N-PHAC–CO_2 interaction energies were obtained employing large-basis MP2 and CCSD(T) calculations. In order to understand the effects of basis set size on the MP2 interaction energy and on the ΔCCSD(T) contribution, we first examined the seven 1- and 2-ring systems that have at least C_s symmetry and obtained an extended set of conventional and explicitly correlated MP2 and CCSD(T) interaction energies including results in basis sets up to a5Z and aTZ, respectively. For the minimum geometry of the in-plane and stacked pyrazine-CO_2 systems, which are both small and highly symmetric, we computed the CCSD(T) results in an even larger, aQZ basis. The results for four representative systems are presented in Table 3.1. This table includes the estimates of the total interaction energy obtained by a straightforward CBS extrapolation of the CCSD(T) results or an augmentation of the MP2/CBS value with the ΔCCSD(T) correction that is either computed or CBS-extrapolated. Table 3.1 illustrates the accuracy to which the CBS limit can be determined when the system size limits the CCSD(T) basis set choice to aTZ (as in the general case of the 1- and 2-ring systems) or aDZ (for all N-PHACs larger than 2-ring, except for the planar phenanthroline-CO_2 where we ran aTZ since it has C_{2v} symmetry). We examined the CCSD(T)-F12 approach with and without the scaling of triples, and noted that the scaling is harmful for the CCSD(T)-F12a variant but beneficial for CCSD(T)-F12b (the first observation indicates that the CCSD(T)-F12a approach, formally more approximate than CCSD(T)-F12b,[94] highly benefits from a cancellation of errors between the CCSD part and the triples part[93]). Therefore, we use the average of the ΔCCSD(T)-F12a and ΔCCSD(T**)-F12b results, further denoted as ΔCCSD(T)-F12avg, as the benchmark value for the ΔCCSD(T) contribution. These values are listed in Table 3.1 along with the values of the ΔCCSD(T)-F12 and ΔCCSD(T**)-F12 interaction energy terms.

The nonextrapolated MP2 and MP2-F12 results in Table 3.1 all converge smoothly to the CBS limit. The (T,Q) and (Q,5) extrapolated values agree to within 0.04 kcal/mol

for each system. Furthermore, the non-extrapolated MP2-F12 results are better than non-extrapolated MP2; however, after extrapolation both methods produce similar results.

The ΔCCSD(T) corrections shown in Table 3.1 exhibit a moderately fast convergence with the basis set size. It is obvious that extrapolation does help in the convergence of ΔCCSD(T). The explicitly correlated coupled cluster methods, especially CCSD(T)-F12a, exhibit faster convergence than traditional CCSD(T). The convergence of ΔCCSD(T)-F12a and ΔCCSD(T**)-F12b is smooth and the extrapolations work well. For the stacked pyrazine–CO_2 complex, the results gathered in Table 3.1 provide benchmark values of -1.717 ± 0.006 kcal/mol for MP2 and 0.537±0.003 kcal/mol for ΔCCSD(T). This leads to the total MP2/CBS+ΔCCSD(T) interaction energy of −1.180±0.007 kcal/mol, where the uncertainties of the two contributions have been added quadratically.

It is not feasible to run CCSD(T)/aQZ for systems that contain more than one ring. It is also preferable to avoid doing CCSD(T)/aTZ for systems that contain more than two rings since these calculations are very demanding. If the CCSD(T)/aTZ calculations are feasible, as in the case of the seven symmetric 1- and 2-ring systems as well as the in-plane C_{2v} phenanthroline-CO_2 dimer, there are four sensible approaches to estimate the benchmark CCSD(T)/CBS limit from either conventional or explicitly correlated calculations: CCSD(T)/(D,T), MP2/(Q,5) + ΔCCSD(T)/(D,T), MP2/(Q,5) + ΔCCSD(T)/aTZ, and MP2/(Q,5) + ΔCCSD(T)/aDZ (note that the nonextrapolated CCSD(T)/aTZ values are inferior to the extrapolated and/or composite results, cf. Table 3.1). For the in-plane and stacked pyrazine-CO_2 complexes, these combinations lead to absolute errors in the range 0.003–0.070 kcal/mol compared to the total interaction energy given by MP2-F12/(Q,5) + ΔCCSD(T)-F12avg/(T,Q). It is noted that the explicitly correlated CCSD(T)-F12 approach clearly improves the basis set convergence, and the CCSD(T)-F12a and CCSD(T**)-F12b values become virtually identical upon extrapolation. Because of this, we will use the average of ΔCCSD(T)-F12a and ΔCCSD(T**)-F12b as the final ΔCCSD(T) contribution to the benchmark: that is, the benchmark interaction energy for the 1- and 2-ring

systems will be MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T). Since the CCSD(T)/aQZ calculations are very time-consuming, we will use the theory level defined above to produce the benchmark potential energy curve also for pyrazine-CO_2. For this system, the errors resulting from the restriction of coupled-cluster calculations to the aTZ basis set are 0.003–0.010 and 0.008–0.014 kcal/mol for MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) and MP2/(Q,5)+ΔCCSD(T)/(D,T), respectively. A further restriction to aDZ leads to errors of 0.028–0.058 and 0.048–0.063 kcal/mol for explicitly correlated and conventional CCSD(T), respectively. Even this last error, corresponding to less than 6% of the total interaction energy, is remarkably low.

The quinoxaline–CO_2 results in Table 3.1 exhibit similar convergence patterns. The conventional ΔCCSD(T) part shows smooth convergence and the aDZ and aTZ bases are sufficient to narrow this term down to about 0.02 kcal/mol. The highest-level conventional estimates of the CBS limit, the MP2/(Q,5) + ΔCCSD(T)/(D,T) results, are too low by 0.018–0.026 kcal/mol compared to the benchmark MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) value. The observed accuracy of the conventional MP2/(Q,5)+ΔCCSD(T)/(D,T) estimate is similar to that found for the pyrazine-CO_2 dimer. The same trend is also true for the MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/aDZ (errors of 0.02–0.05 kcal/mol) and MP2/(Q,5) +ΔCCSD(T)/aDZ (errors of 0.03–0.08 kcal/mol) estimates. Consequently, the satisfactory, better than 0.1 kcal/mol accuracy of even the simplest approach, MP2/(Q,5) +ΔCCSD(T)/aDZ, is likely transferable to dimers that involve larger N-PHACs. Therefore, all benchmarks for systems larger than two rings, and for the nonsymmetric stacked quinoline-CO_2 complex, will utilize the conventional MP2/(Q,5)+ΔCCSD(T)/aDZ level except for the C_{2v} phenanthroline-CO_2 system and the largest dimers, which will employ MP2/(Q,5)+ ΔCCSD(T)-F12avg/(D,T) and MP2/(Q,5)+ΔCCSD(T)/laDZ, respectively—see below. As seen from Table 3.1, the F12 approach does greatly improve the calculated MP2 results. The conventional MP2 results are nearly as accurate as long as the CBS extrapolation is performed,

but the uncertainties of conventional MP2 (computed as differences between the extrapolated MP2/(Q,5) energy and the calculated MP2/a5Z energy) are larger than those of the MP2-F12 interaction energies. In the case of quinoxaline-CO_2, the MP2 and MP2-F12 uncertainties amount to 0.054 and 0.001 kcal/mol for the in-plane dimer and 0.064 and 0.013 kcal/mol for the stacked dimer, respectively. Nevertheless, all MP2 interaction energies for three-ring and larger systems, and for the stacked quinoline-CO_2 structure, will be obtained from a conventional extrapolation at the (Q,5) level as the cost of MP2-F12/a5Z starts to become prohibitive.

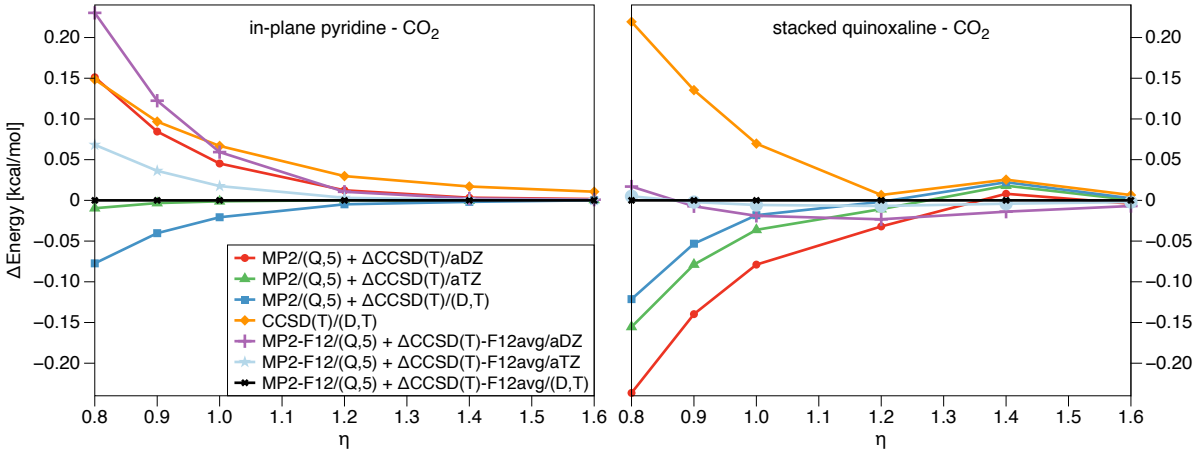

Figure 3.2: Differences between the benchmark MP2-F12/(Q,5) + ΔCCSD(T)-F12avg/(D,T) interaction energy and other CCSD(T)/CBS schemes as functions of η for the in-plane pyridine-CO_2 (left panel) and stacked quinoxaline-CO_2 (right panel) complexes.

Figure 3.2 displays the differences between the benchmark MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) result and various other CCSD(T)/CBS estimates for in-plane pyridine-CO_2 (left panel) and stacked quinoxaline-CO_2 (right panel) as functions of η (η is defined throughout the text as $\frac{z}{z_{min}}$, where z_{min} represents the minimum-energy z distance for each dimer). At the minimum separations for both complexes, all extrapolations shown agree with the benchmark to within 0.1 kcal/mol. For larger η, all considered variants are nearly as accurate as the benchmark MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) value. It is the short range, $\eta < 1.0$, where different benchmark variants start deviating more from each other. If only the aDZ basis set is available for the ΔCCSD(T) contribution, MP2/(Q,5)+ΔCCSD(T)/aDZ is

superior to MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/aDZ for in-plane pyridine-CO_2 but inferior for stacked quinoxaline-CO_2.

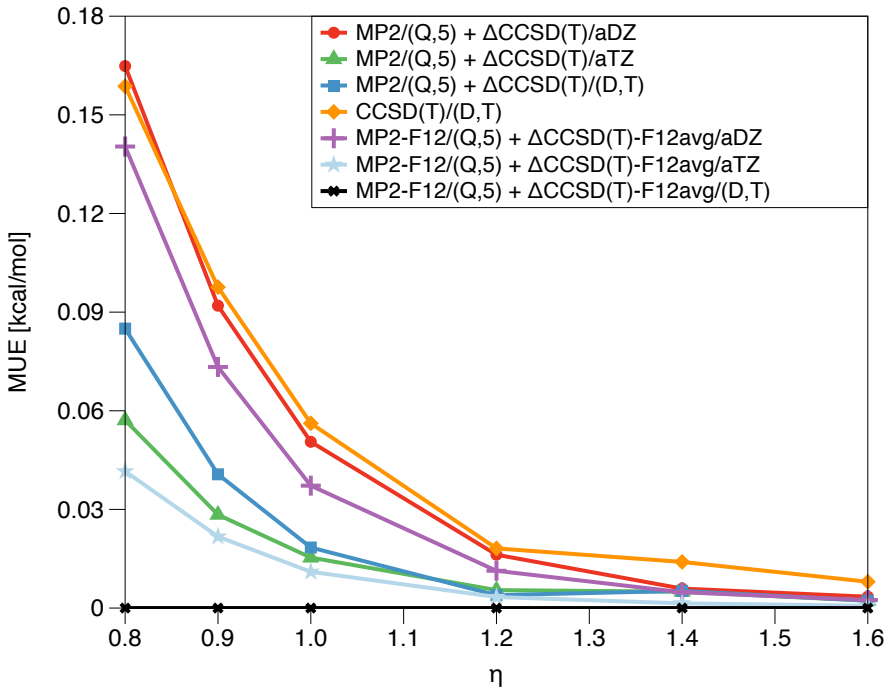

Figure 3.3: The mean unsigned error (MUE) for different CCSD(T)/CBS estimates as a function of η, using the MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) as the benchmark interaction energy for the seven symmetric 1- and 2-ring systems.

Since the performance of various CCSD(T)/CBS estimates for the in-plane and stacked dimers is quite different, we examine the mean unsigned errors (MUE) of different CCSD(T)/CBS schemes for all the symmetric 1- and 2-ring systems (7 dimers altogether) compared to the benchmark MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/(D,T) results in Fig. 3.3. At the minimum distance, $\eta = 1.0$, the values for all different schemes agree to within 0.06 kcal/mol. However, different CCSD(T)/CBS estimates start deviating from each other in the repulsive region. From a statistical point of view, the MP2-F12/(Q,5)+ΔCCSD(T)-F12avg/aDZ results are slightly superior to the MP2/(Q,5)+ΔCCSD(T)/aDZ ones, but the differences are small and consistent across the whole distance range. Thus, the limited increase in the

short-range accuracy does not warrant the additional computational effort of a CCSD(T)-F12 calculation and the benchmark interaction energies for larger systems will include the ΔCCSD(T) contribution from conventional CCSD(T).

The CCSD(T)/aDZ interaction energies cannot be computed for the 1,6-diazacoronene–CO_2 system due to the presence of diffuse basis functions on multiple centers leading to near linear dependencies in the basis set. In order to overcome the linear dependency issue, at least some of the offending diffuse basis functions have to be removed. However, the most popular scheme of removing diffuse functions from aXZ, the "calendar" basis sets,[110] may lead to persisting linear dependencies, a significant drop in accuracy, or both, as shown for the methane–pyrene and methane–coronene complexes.[74] Instead, an alternative basis set truncation scheme is employed where only the carbon dioxide atoms and the six closest heavy atoms (relative to the carbon in carbon dioxide) of the N-PHAC molecule retain diffuse functions; this approach is labeled as local-aug-cc-pVDZ (laDZ).[74,75] Pictorial demonstrations of the augmentation schemes in the 2-azapyrene–CO_2 and 1,6-diazacoronene–CO_2 complexes are displayed in Fig. 3.1. The bronze-colored carbon atoms in the N-PHAC molecule as well as the closest nitrogen atom (blue) are those that have diffuse functions in the laDZ basis. A comparison of the MP2 interaction energies and ΔCCSD(T) corrections calculated using the full aDZ basis, its laDZ subset, and the nonaugmented cc-pVDZ set across the whole range of η is shown in Table 3.2. This table gathers the results for 2-azapyrene–CO_2, for which we can get the CCSD(T) results for both aDZ and laDZ, and 1,6-diazacoronene–CO_2, for which we can only obtain the ΔCCSD(T) values in the laDZ basis.

The 2-azapyrene–CO_2 results in Table 3.2 demonstrate that diffuse functions are significant for both the MP2 interaction energy and the ΔCCSD(T) correction. The energy differences between full aDZ and cc-pVDZ are up to 4.8 kcal/mol for the MP2 interaction energy and up to 0.6 kcal/mol for the ΔCCSD(T) contribution. At the minimum distances, the deviations between ΔCCSD(T)/cc-pVDZ and ΔCCSD(T)/aDZ are 0.01 and 0.28 kcal/mol for the in-plane and stacked 2-azapyrene–CO_2 complexes, respectively. These

Table 3.2: The MP2 and ΔCCSD(T) contributions to the in-plane and stacked 2-azapyrene-CO_2 and 1,6-diazacoronene-CO_2 interaction energies (in kcal/mol) as functions of $\eta = z/z_{\min}$. No density fitting was used in this table.

		2-azapyrene–CO_2				1,6-diazacoronene–CO_2			
		MP2		ΔCCSD(T)		MP2		ΔCCSD(T)	
basis	η	in-plane	stacked	in-plane	stacked	in-plane	stacked	in-plane	stacked
cc-pVDZ	0.8	7.093	12.359	0.300	3.484	7.992	12.241	0.483	3.670
	0.9	-0.251	1.555	0.090	2.028	-0.363	1.183	0.179	2.236
	1.0	-2.322	-1.405	0.005	1.218	-2.587	-1.929	0.053	1.396
	1.2	-2.087	-1.525	-0.013	0.498	-2.215	-2.001	0.007	0.600
	1.4	-1.222	-0.854	-0.001	0.233	-1.258	-1.180	0.010	0.287
	1.6	-0.691	-0.462	0.002	0.118	-0.695	-0.675	0.009	0.147
laDZ	0.8	3.986	8.005	0.386	4.093	4.962	7.440	0.583	4.380
	0.9	-2.365	-1.533	0.139	2.447	-2.442	-2.321	0.230	2.759
	1.0	-3.855	-3.581	0.026	1.499	-4.089	-4.479	0.080	1.768
	1.2	-2.878	-2.541	-0.012	0.621	-3.039	-3.264	0.014	0.781
	1.4	-1.657	-1.320	0.001	0.292	-1.714	-1.793	0.015	0.382
	1.6	-0.956	-0.692	0.008	0.151	-0.957	-0.991	0.016	0.201
aDZ	0.8	4.005	7.595	0.366	4.085	4.811	6.955		
	0.9	-2.369	-1.777	0.119	2.441	-2.540	-2.644		
	1.0	-3.826	-3.723	0.015	1.493	-4.151	-4.695		
	1.2	-2.906	-2.584	-0.018	0.615	-3.109	-3.337		
	1.4	-1.678	-1.331	-0.004	0.288	-1.721	-1.812		
	1.6	-0.943	-0.691	0.005	0.147	-0.975	-0.993		

values should be contrasted with the corresponding differences between ΔCCSD(T)/laDZ and ΔCCSD(T)/aDZ, amounting to 0.01 kcal/mol in both cases. Thus, while for the in-plane structure the ΔCCSD(T) correction is quite unimportant, for the stacked geometry the partial augmentation present in the laDZ set is both necessary and sufficient to obtain an accurate value of this term. For the two dimers presented, the differences between the ΔCCSD(T) terms in two basis sets are roughly an order of magnitude smaller than the differences in the MP2 interaction energy. Thus, the results in Table 3.2 suggest that the ΔCCSD(T)/laDZ values for 1,6-diazacoronene-CO_2 should be within about 0.02 kcal/mol from the full aDZ results. Consequently, the benchmark 1,6-diazacoronene–CO_2 interaction energies will be obtained at the MP2/(Q,5)+ΔCCSD(T)/laDZ level.

The results in Tables 3.1–3.2 show that the ΔCCSD(T) correction is quite small, typically less than 0.1 kcal/mol at the minimum distance, for the in-plane dimers. For the stacked configurations, the ΔCCSD(T) correction is significantly larger and MP2 overbinds by up to 1.8 kcal/mol at the minimum distance. This is the case across all the in-plane and stacked structures (eight planar and eight stacked configurations, as shown in Fig. 3.1). In order to investigate this phenomenon for all systems, the mean unsigned relative errors (MURE) of MP2 and spin component scaled MP2 (SCS-MP2)[111] compared to the benchmark values are displayed in Fig. 3.4. As explained in more detail below, the averaging includes 95 geometries (the η=0.9 point of stacked pyridine–CO_2 is discarded because it is accidentally very close to zero). As the ΔCCSD(T) correction is small for in-plane complexes, the MP2/(aQZ, a5Z) result is very good with an overall MURE of 4.4% while the SCS-MP2/(aQZ, a5Z) results are quite poor (a MURE of 23.8%). The opposite is true for stacked complexes, with a MURE of 12.7% for SCS-MP2/(aQZ, a5Z) and 44.6% for MP2/(aQZ, a5Z). The MURE of the CP-uncorrected MP2 and SCS-MP2 interaction energies is also shown in Fig. S1. One can conclude from this figure that the lack of a CP correction significantly worsens the accuracy of MP2 but provides some limited improvement at the SCS-MP2 level.

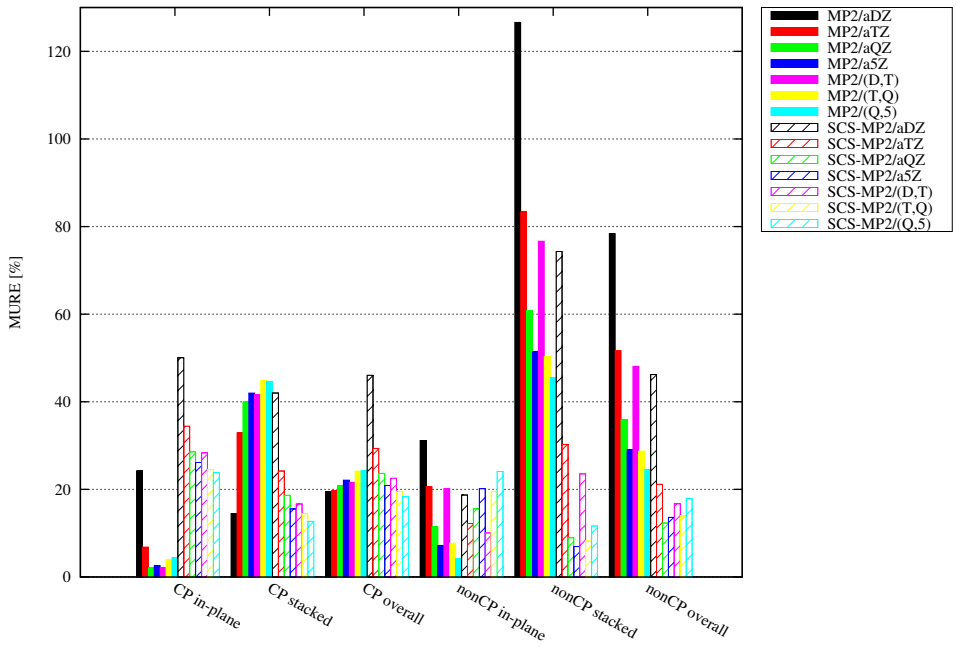

Figure 3.4: The mean unsigned relative errors (MURE) for MP2 and SCS-MP2 compared, both CP-corrected and nonCP-corrected, to the CCSD(T)/CBS-level benchmark values as defined in the text. The "CP" means CP-corrected, and "nonCP" represents the nonCP-corrected values. The "overall" label signifies the results of all stacked and in-plane structures.

The benchmark interaction energies computed so far do not include any monomer flexibility effects. These effects can be examined by comparing the van der Waals well depth obtained with the monomers frozen at their own optimized geometries (as is the case throughout the rest of this work) to the well depth calculated by a minimization of the CP-corrected interaction energy between fully flexible monomers. In the latter case, the quantity that needs to be minimized is

$$\begin{aligned} E_{\rm int}^{\rm flexible} &= [E^{\rm AB}(AB) - E^{\rm AB}(A) - E^{\rm AB}(B)] \\ &+[E^{\rm A}(A) - E_0^{\rm A}(A)] + [E^{\rm B}(B) - E_0^{\rm B}(B)] \end{aligned} \tag{3.3}$$

where the subscript 0 represents the nonrelaxed minimum geometry of the monomer, the superscripts signify the basis set (dimer-centered or monomer-centered), and the symbols in

parentheses denote the subsystems. On the example of the in-plane pyridine-CO_2 dimer, we first examined the effects of the CO_2 flexibility, with the pyridine monomer remaining rigid. The resulting flexible well depth, obtained by minimizing Eq. (3.3) with the system constrained to the C_{2v} symmetry, is larger by 0.166 kcal/mol (at the MP2/aTZ level) than the conventional rigid well depth. Furthermore, the changes in the C-O bond lengths do not exceed 0.0003 Å and the O-C-O angle change is 3.15°. The flexible-CO_2 energy is further lowered by 0.031 kcal/mol in a completely unrestricted MP2/aTZ dimer optimization. Thus, the total flexibility contribution to interaction energy is 0.197 kcal/mol, which is in good agreement with the previous theoretical result of 0.20 kcal/mol obtained for this system at the MP2/cc-pVTZ level.[77] The C-O bond length and the O-C-O angle keep almost the same values in the CO_2-only and fully flexible optimizations. Therefore, the flexibility effect comes mainly from the relaxation of the CO_2 molecule and the flexibility of the N-PHAC monomer will be neglected.

In the case of the stacked pyridine-CO_2 complex, we first optimized the tilt angle of CO_2, resulting in an interaction energy (still rigid and symmetric but not restricted to parallel configurations anymore) lower by 0.002 kcal/mol. Then, based on the geometry of this tilted configuration, the CO_2-only flexible minimization was performed and the relaxation energy was 0.011 kcal/mol. Finally, a totally unrestricted MP2/aTZ dimer optimization on stacked pyridine-CO_2 dimer was performed, lowering the interaction energy by a further 0.023 kcal/mol. Thus, the total flexibility contribution to interaction energy for stacked pyridine-CO_2 dimer is 0.034 kcal/mol. The CO_2-only relaxation energies for the N-PHAC-CO_2 dimers are gathered in Table 3.3 to demonstrate the flexibility effects. The flexibility effects for stacked complexes turn out to be very small except for the phenanthroline–CO_2 system.

We conclude that while the flexibility effects are significant for the in-plane orientations, they are negligible for the stacked ones. One may note that the stacked configurations are more representative of extended nanotubes than the in-plane ones, and that the precise

Table 3.3: The optimized minimum distance $z_{\min}$ (Å), the O-C-O angle deformation $\Delta\phi_{\mathrm{OCO}}$ (°), the CO_2-only flexible energy change ΔE_{flex} (kcal/mol), and the C-O bond length change Δr_{CO} (Å), calculated at the MP2/aTZ level for the lowest-energy N-PHAC–CO_2 structures. The N-PHAC monomer was kept rigid.

Complex	Orientation	z_{min}	$\Delta\phi_{\mathrm{OCO}}$	ΔE_{flex}	Δr_{CO}
pyridine-CO_2	in-plane	2.826	3.15	0.166	0.0003
	stacked	3.144	0.75	0.011	0.0011
pyrazine-CO_2	in-plane	2.840	2.58	0.112	0.0002
	stacked	3.370	0.13	0.001	0.0004
quinoline-CO_2	in-plane	2.846	2.97	0.148	0.0003
	stacked	3.138	0.73	0.011	0.0013
quinoxaline-CO_2	in-plane	2.873	2.61	0.114	0.0002
	stacked	3.080	0.31	0.003	0.0010
pyrido[3,2-g]quinoline-CO_2	in-plane	2.883	2.80	0.132	0.0008
	stacked	3.103	0.47	0.006	0.0015
phenanthroline-CO_2	in-plane	3.791	3.93	0.258	0.0004
	stacked	2.821	3.56	0.215	0.0014
2-azapyrene-CO_2	in-plane	2.813	3.30	0.183	0.0004
	stacked	3.139	0.41	0.005	0.0014

details of the geometry (flexible or rigid) are not relevant for the main purpose of this work, an assessment of the performance of different DFT-based approaches. Moreover, as we proceed to examine radial cross sections through the intermolecular potential energy surfaces, it would be cumbersome to reoptimize the monomer geometry at each intermolecular separation. Therefore, all calculations throughout the rest of this work utilize rigid monomers.

The accuracy of the benchmark energies obtained in this section does not only rely on the accuracy with which the CCSD(T)/CBS limit is estimated, but also on the smallness of the effects neglected in the frozen-core CCSD(T) calculation. In order to examine the effect of core-core and core-valence correlation on the interaction energy, we computed the all-electron CCSD(T)/aug-cc-pCVDZ and CCSD(T)/aug-cc-pCVTZ interaction energies for the four 1-ring systems at $\eta = 1.0$. In the aug-cc-pCVDZ basis, the all-electron and frozen-core interaction energies differ by 0.004–0.005 kcal/mol; for the aug-cc-pCVTZ set the corresponding differences range from 0.002 to 0.011 kcal/mol. Thus, the interaction

energy contributions from the core-core and core-valence correlation should not exceed a few hundredths of a kilocalorie per mole. The relativistic effects are likely even smaller as only light atoms are present. The effects of coupled-cluster excitations beyond CCSD(T) are the hardest to estimate, especially since their basis set convergence is often slow.[112,113] The most similar system for which these effects have been estimated is the benzene dimer, for which Pitoňák *et al.*[114] performed small-basis CCSD(TQ_f)[115] calculations, taking into account approximate quadruple excitations. The post-CCSD(T) interaction energy corrections obtained in this way ranged from 0.02 to 0.04 kcal/mol so we expect the post-CCSD(T) effects to be of similar magnitude for the systems considered here. Overall, our benchmark interaction energies are likely accurate to 0.1 kcal/mol or better at the minimum separations ($\eta = 1.0$).

3.4.2 DFT calculations

In this section, we examine how well different DFT functionals recover the MP2/CBS +ΔCCSD(T) benchmark interaction energies for the N-PHAC-CO_2 complexes. The 16 minima obtained in Sec. 3.3.1 were investigated at η times the minimum distance z_{min}, $\eta = 0.8$, 0.9, 1.0, 1.2, 1.4, and 1.6, giving a total of 96 CCSD(T) results. The DFT interaction energies were computed using the def2-SVP, TZVP, QZVP[116] and Dunning aDZ and aTZ basis sets combined with five possible variants of Grimme's dispersion correction: -D2,[29] -D3,[34] -D3(BJ),[105] -D3-$E^{(3)}$, and -D3(BJ)-$E^{(3)}$, both with and without the CP correction. As the M0x series are interaction-optimized functionals, we also examined their performance without an additional atom-pairwise dispersion term. This corresponds to a total of 450 different combinations of functionals, basis sets, dispersion corrections, and the CP correction or lack thereof (note that the dispersion corrections for the M0x series include only -D3 and -D3-$E^{(3)}$, and there is no -D2 correction for LC-ωPBE).

The accuracy of different DFT variants with respect to the CCSD(T)-level benchmark interaction energy, obtained as described above, will be investigated using mean unsigned

relative error (MURE). The results include all 96 points except for a point in the repulsive region (the η=0.9 geometry of stacked pyridine–CO_2) which is very close to zero and would accidentally dominate the MURE. The MURE values of all 450 different combinations of functionals, basis sets, dispersion corrections, and the CP correction or lack of it (nonCP) are collected in the Supporting Information of Ref. 117. Overall, the MUE/MURE values in the QZVP basis range from 0.17 kcal/mol / 6.4% for nonCP B2PLYP-D3 to 1.25 kcal/mol / 45.9% for CP B97-D2. This range of errors is very similar for smaller basis sets down to aDZ, but a further basis reduction to SVP increases the errors to the range between 0.28 kcal/mol (CP M06-2X-D3)/11.1% (CP LC-ωPBE-D3-$E^{(3)}$) and 1.33 kcal/mol (CP B97-D2)/52.2 % (nonCP BLYP-D3(BJ)). To examine which DFT variants provide the most consistent accuracy, we analyzed the MURE values for each η and each geometry type (in-plane/stacked) separately — the pertinent results are given in the Supporting Information of Ref. 117. The deterioration of accuracy at the short range is clear for most of the methods: at the shortest separation $0.8z_{min}$ in the QZVP basis, only the different variants of M05-2X, M06-2X, PBE0 (both CP and nonCP), LC-ωPBE, and B2PLYP-D3 (nonCP only) attain a MURE below 20%. As expected, the M05-2X and M06-2X approaches without additional dispersion perform poorly in the long range (MURE in the QZVP basis is over 20% for all $\eta \geq 1.2$). However, the M05-2X-D3 and M06-2X-D3 variants perform fairly well at all separations. The MURE values as functions of η for these functionals, along with the next best approaches B3LYP-D3(BJ), LC-ωPBE-D3 (with and without the CP correction), and B2PLYP-D3 (nonCP only) are presented in Fig. 3.5. Additionally, this figure shows the corresponding MURE separated into in-plane and stacked complexes. The PBE0-D3 functional was omitted from Fig. 3.5 as its reasonable (MURE 16–23% depending on the -D3 variant and CP/nonCP) accuracy at $\eta = 0.8$, mentioned above, deteriorates to 35–43% at $\eta = 0.9$.

The results in Fig. 3.5 indicate that the performance at different η varies significantly: in particular, the relative accuracy of virtually all DFT-based methods decreases at distances

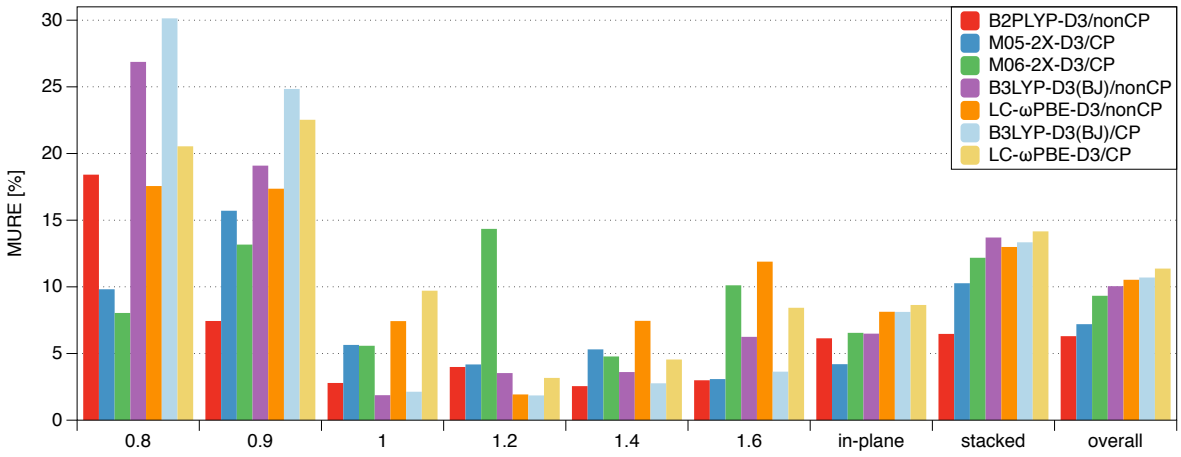

Figure 3.5: Mean unsigned relative errors (MURE) for the best-performing DFT-based methods in the QZVP basis set as functions of the relative intermolecular separation η (the overall value for all η is displayed as "Overall") compared against the MP2/CBS+ΔCCSD(T) benchmark interaction energies for the 95 model N-PHAC-CO_2 geometries. In addition, separate MURE values for the in-plane and stacked structures are displayed.

somewhat shorter than the minimum. Both the -D3 extensions of standard DFT variants such as B3LYP and PBE and the double hybrid functional B2PLYP-D3 perform very well at the van der Waals minima and at larger distances, with most MURE values below 5%. However, the errors increase several times in the mildly repulsive region of the interaction, with B2PLYP-D3 performing somewhat better than lower-rung functionals but still unsatisfactorily. The M06-2X-D3 approach presents a particularly interesting case: it provides, along with M05-2X-D3, by far the best accuracy at $\eta = 0.8$ but the errors vary irregularly with the separation, with large MURE values at $\eta = 1.2$ and 1.6 (but not 1.4). The behavior of M05-2X-D3 is much more stable although the errors at $\eta = 0.9$ are somewhat large. While Fig. 3.5 indicates that M05-2X-D3, along with B2PLYP-D3, should be the method of choice for studying N-PHAC–CO_2 potential energy surfaces, the oscillating accuracy of M06-2X-D3 shows, in our opinion, the dangers of strongly parameterized functionals.[118] Figure 3.5 also shows that all top-performing DFT variants except for B2PLYP-D3 provide somewhat better relative accuracy for the in-plane geometries than for the stacked ones, but the orderings

of functionals according to their MURE for in-plane and stacked complexes are remarkably similar.

The deterioration of the DFT+D accuracy in the mildly repulsive region of the interaction presents a serious problem as this region is extensively sampled in dynamics calculations and relevant for the computation of many interaction-dependent observables such as second virial coefficients and scattering cross sections. As argued in our recent work on interactions between CO_2 and pristine carbon nanotubes,[76] there are likely two reasons for this deterioration: overestimation of exchange by standard GGA functionals[119,120] and inadequacy of the damping functions used in the -D3 dispersion correction. The first of these issues can be alleviated by improving on the asymptotic behavior of the exchange functional by increasing the fraction of exact exchange at long range. This can be accomplished through the range separation (long-range correction) technique[121–123] as exemplified by the LC-ωPBE functional[104] (for which the -D3 parameters are available). This is the reason why we included the LC-ωPBE functional in the set of methods tested. However, Fig. 3.5 shows that the performance of LC-ωPBE at $\eta = 0.8$ and 0.9 is still not satisfactory. Therefore, we believe that at least a part of the problem lies in the damping forms of -D3 which have been optimized mostly for data at the van der Waals minimum separations.[34,105] Consequently, the authors of Ref. 76 proposed a refitting of the -D3, -D3(BJ), and Tang-Toennies[124] damped dispersion -D3(TT) to optimally reproduce the curved coronene–CO_2 benchmark interaction energies, also at intermolecular separations as short as 0.8 times the minimum. Therefore, *without any additional refitting,* we checked how the damping parameters of Ref. 76 perform relative to the original parameters from Refs. 34 and 105 on our N-PHAC–CO_2 benchmark dataset. The results are shown in Fig. 3.6, in the largest basis set QZVP with the CP correction. With the refitting, PBE-D3$_{\mathrm{refit}}$/CP is the top DFT performer (5.1%), followed by the B2PLYP-D3(TT)$_{\mathrm{refit}}$/CP (5.8%). Refitting improves results for almost all the variants except for the -D3(BJ)$_{\mathrm{refit}}$ approach for some functionals. While it was shown in Ref. 76 that

-$D3_{refit}$ (though not -$D3(BJ)_{refit}$ or -$D3(TT)_{refit}$) performs as well as original -D3 for the popular S22x5[125] and S66x8[37] databases, the transferability of the refitted damping parameters should not be taken for granted. Therefore, it is highly gratifying that the parameters from Ref. 76 improve the performance of standard DFT-D3 also for the N-PHAC–CO_2 complexes without any refitting required.

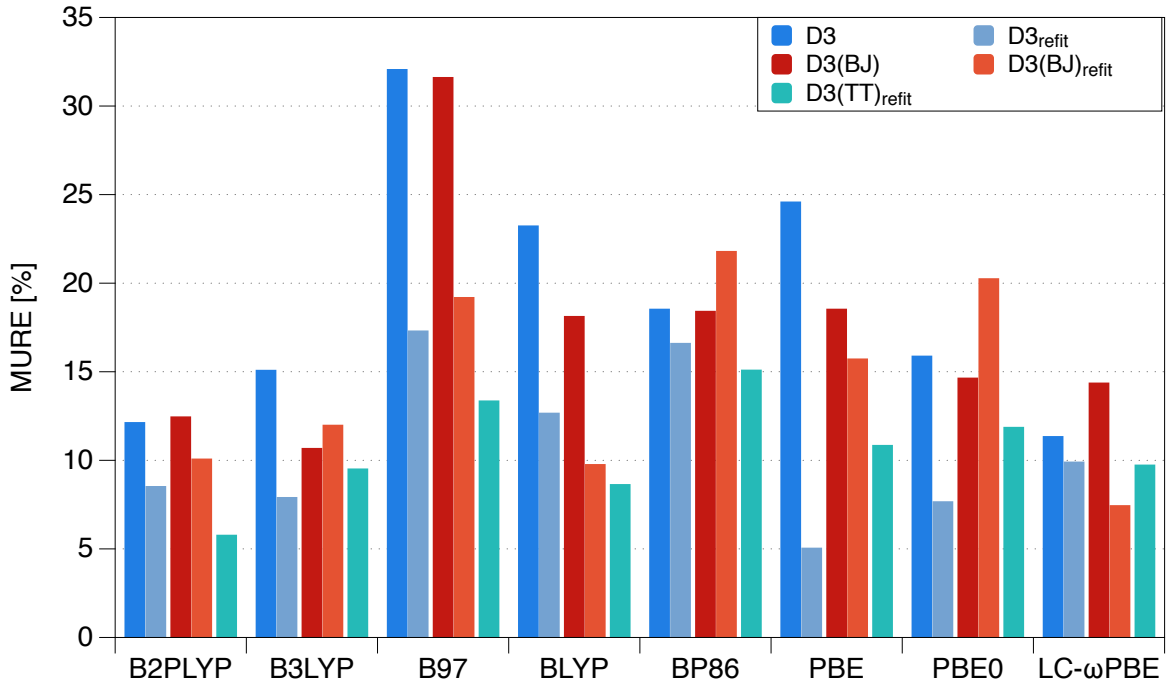

Figure 3.6: Mean unsigned relative errors (MURE) for CP-corrected DFT-D3 interaction energies using different damping functions (original and refitted in Ref. 76) in the largest basis set QZVP, against the MP2/CBS+ΔCCSD(T) benchmark interaction energies for the 95 model N-PHAC-CO_2 geometries.

Figure 3.7 illustrates the basis set dependence of the accuracy of the top-performing functionals. This figure indicates that larger basis sets improve the DFT+D performance, with the MURE generally decreasing in the order SVP → aDZ → TZVP → aTZ → QZVP. In the largest basis set, QZVP, the MURE values for CP and nonCP are almost identical to each other, except that B2PLYP/nonCP is superior to B2PLYP/CP. One should note that conventional functionals such as B3LYP and M05-2X exhibit faster basis set convergence

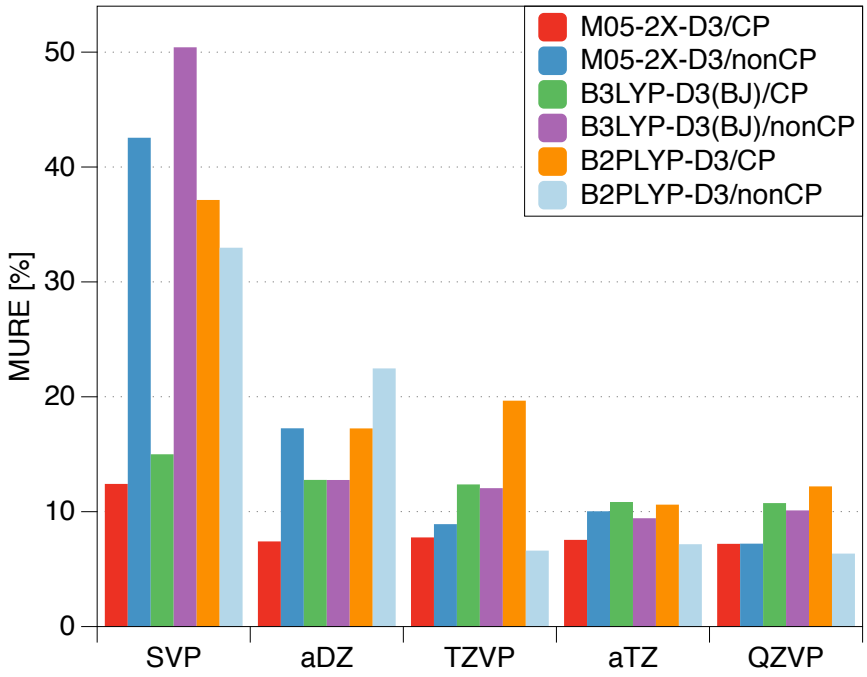

Figure 3.7: Mean unsigned relative errors (MURE) for the best performers: M05-2X-D3, B2PLYP-D3, and B3LYP-D3(BJ) in different basis sets, with and without the CP correction, against the MP2/CBS+ΔCCSD(T) benchmark interaction energies for the 95 model N-PHAC-CO_2 geometries.

than the B2PLYP double hybrid method, and their accuracy does not decay so dramatically in smaller basis sets (even SVP) as long as the CP correction is applied. This result is in agreement with the findings of Refs. 5 and 76 who found that the CP-corrected DFT results converge smoothly and different bases require similar damping parameters in the accompanying -D3 term.

The mean unsigned relative error for the best DFT functionals as a function of the N-PHAC size (the number of rings) is displayed in Fig. 3.8. In particular, we are interested to find out if there is any deterioration of the DFT-D3 accuracy with increasing system size due to the neglect of pairwise-nonadditive effects[126] and, if so, if the three-body -$E^{(3)}$ dispersion correction[34] helps alleviate this deterioration as suggested in a recent benchmark study of large weakly interacting complexes.[38] However, the results in Fig. 3.8 show that the relative

accuracy of top-performing DFT variants is highly uniform across systems of different sizes with an exception of M05-2X-D3 which strangely displays poor performance for the four-ring (2-azapyrene–CO_2) complexes. Moreover, the influence of the -$E^{(3)}$ term is minor in all cases. We conclude that the pairwise-nonadditive dispersion effects are relatively unimportant for the N-PHAC–CO_2 complexes considered in this work.

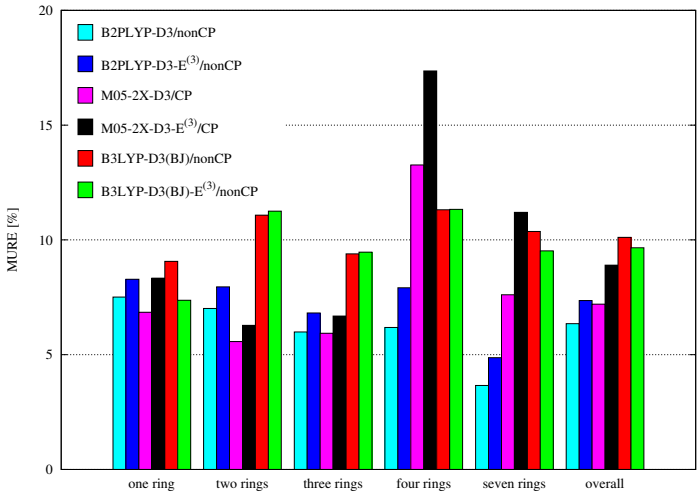

Figure 3.8: The mean unsigned relative errors (MURE) for the best performers: M05-2X-D3/CP, M05-2X-D3-$E^{(3)}$/CP, B2PLYP-D3/nonCP, B2PLYP-D3-$E^{(3)}$/nonCP, B3LYP-D3(BJ), B3LYP-D3(BJ)-$E^{(3)}$/nonCP, in the largest basis set QZVP computed, as functions of the number of rings in the N-PHAC molecule (the overall MURE for all 95 N-PHAC-CO_2 geometries is displayed in the last column) against the MP2/CBS+ΔCCSD(T) benchmark interaction energies.

To conclude the DFT analysis, the interaction energies at the MP2/CBS level, the MP2/CBS+ΔCCSD(T) benchmarks, and the top DFT performers: B2PLYP-D3/$^{\text{nonCP}}_{\text{QZVP}}$, B3LYP-D3(BJ)/$^{\text{nonCP}}_{\text{QZVP}}$, and M05-2X-D3/$^{\text{CP}}_{\text{QZVP}}$ for each N-PHAC are shown in Fig. 3.9. As mentioned above, the size of the ΔCCSD(T) correction is very small for in-plane geometries. As a result, the MP2 and CCSD(T) results for the eight in-plane configurations differ by only up to 0.07 kcal/mol. Conversely, the stacked dimers strongly benefit from the ΔCCSD(T) contribution: as expected, the MP2 values overestimate the interaction

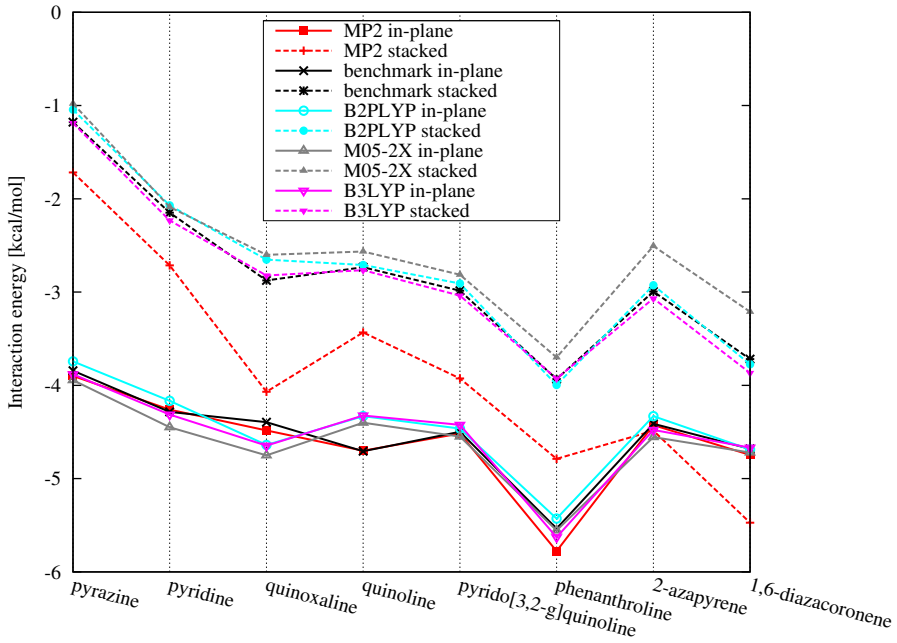

Figure 3.9: Comparison of the interaction energies calculated by different approaches for the in-plane and stacked minimum structures (obtained as described in the text) of all N-PHAC-CO_2 dimers considered here. The MP2 values are taken from MP2-F12/(Q,5) for the seven symmetric 1- and 2-ring systems and MP2/(Q,5) for larger systems. The benchmark values are calculated at the MP2+ΔCCSD(T) level as described in the text. The DFT results are: B2PLYP-D3/$^{\rm nonCP}_{\rm QZVP}$, M05-2X-D3/$^{\rm CP}_{\rm QZVP}$, and B3LYP-D3(BJ)/$^{\rm nonCP}_{\rm QZVP}$.

energies. For all of the investigated systems, the global minimum configurations are in-plane, even though MP2/CBS predicts an incorrect minimum structure for the 2-azapyrene and 1,6-diazacoronene complexes. For both in-plane and stacked structures, the best DFT performers do an excellent job of reproducing the CCSD(T)-level benchmark interaction energy. Overall, the best performing DFT/QZVP functional at the minimum separations ($\eta = 1.0$), B3LYP-D3(BJ)-E$^{(3)}$/nonCP, reproduces the minimum benchmark values to within 0.06 kcal/mol or 1.7% on the average. This level of accuracy is clearly fortuitous and does not carry on to other distances, but it illustrates that modern DFT variants are very capable of providing accurate van der Waals minimum energies as opposed to interaction energies in the mildly repulsive region.

3.4.3 Performance of Selected DFT Variants on Model N-Doped Graphene Holes

The DFT approaches that best reproduce the CCSD(T)-level benchmark results for CO_2 interacting with N-PHACs should be the methods of choice for studying interactions involving larger N-PHACs as well as with extended structures such as nitrogen-doped graphene sheets and carbon nanotubes. We performed the first step in this direction and computed, for the best DFT performers identified in Sec. 3.4.2, the CO_2 interaction energies with three larger N-PHACs representing the barriers to the CO_2 transition through model vacancies in the N-doped graphene surface. The geometries for these N3, N4, and N4H4 vacancy models interacting with CO_2 are presented in Fig. 3.10. The geometry of the vacancy was optimized at the MP2/aDZ, B3LYP/aTZ, and B3LYP/aDZ level for N3, N4, and N4H4, respectively, and the CO_2 molecule was confined to the perpendicular orientation along the symmetry axis of the vacancy. Figure 3.11 displays the interaction energies for the best DFT performers determined in Sec. 3.4.2: B3LYP-D3(BJ)/$^{\rm nonCP}_{\rm aTZ}$, M05-2X-D3/$^{\rm CP}_{\rm aTZ}$, and B2PLYP-D3/$^{\rm nonCP}_{\rm aTZ}$ as well as the CP-corrected MP2/aTZ interaction energies. From this figure, the energy barrier for CO_2 traveling through the (rigid) N3 vacancy is 507, 472, 485, and 450 kcal/mol for M05-2X-D3/$^{\rm CP}_{\rm aTZ}$, B2PLYP-D3/$^{\rm nonCP}_{\rm aTZ}$, B3LYP-D3(BJ)/$^{\rm nonCP}_{\rm aTZ}$, and MP2/aTZ, respectively. The highest energy barrier occurs when one of the oxygen atoms in carbon dioxide passes through the N3 plane. For the N4 vacancy-CO_2 complex, the respective energy barriers are 216, 209, 213, and 202 kcal/mol. Thus, the rigid N3 and N4 holes are impassable to CO_2, but we still need to consider the stretching of the vacancy by the passing CO_2 molecule that may significantly lower the barrier.

In the case of the N3 vacancy-CO_2 dimer, we will first examine the flexibility effects with the CO_2 carbon remaining in the center of the vacancy and allowing both the N-PHAC and the C–O bonds to stretch. The resulting flexible $z = 0$ barrier, obtained by minimizing Eq. (3.3) with the system constrained to the C_{2v} symmetry, is smaller by 97.7 kcal/mol (at the B3LYP/aDZ level) than the conventional rigid $z = 0$ barrier. Furthermore,

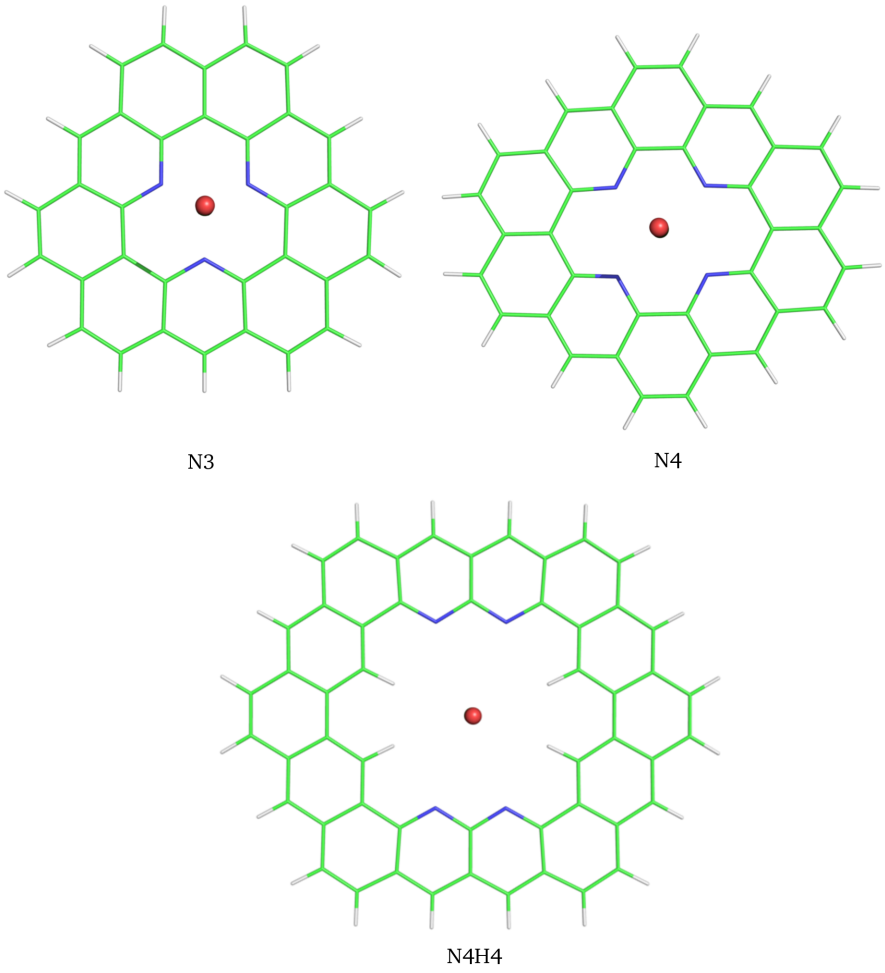

Figure 3.10: The structures for CO_2 interacting with the N-PHAC models of the N3, N4, and N4H4 vacancies in N-doped graphene.

the C–O bond length is larger by 0.237 Å and the distance from the center of the N3 vacancy to the nitrogen atoms is elongated by 0.148 Å. Based on the geometrical parameters from the B3LYP/aDZ flexible optimization, the single point calculations using MP2/aDZ, B3LYP-D3/aTZ, and M05-2X-D3/aTZ (all CP-corrected) were performed. The resulting interaction energies including the monomer deformation effects are 177.4, 174.0, and 186.4 kcal/mol, respectively.

When it comes to the N4 vacancy-CO_2 dimer, the flexibility effects, estimated by minimizing Eq. (3.3) with the system constrained to the C_{2v} symmetry, lower the rigid $z = 0$

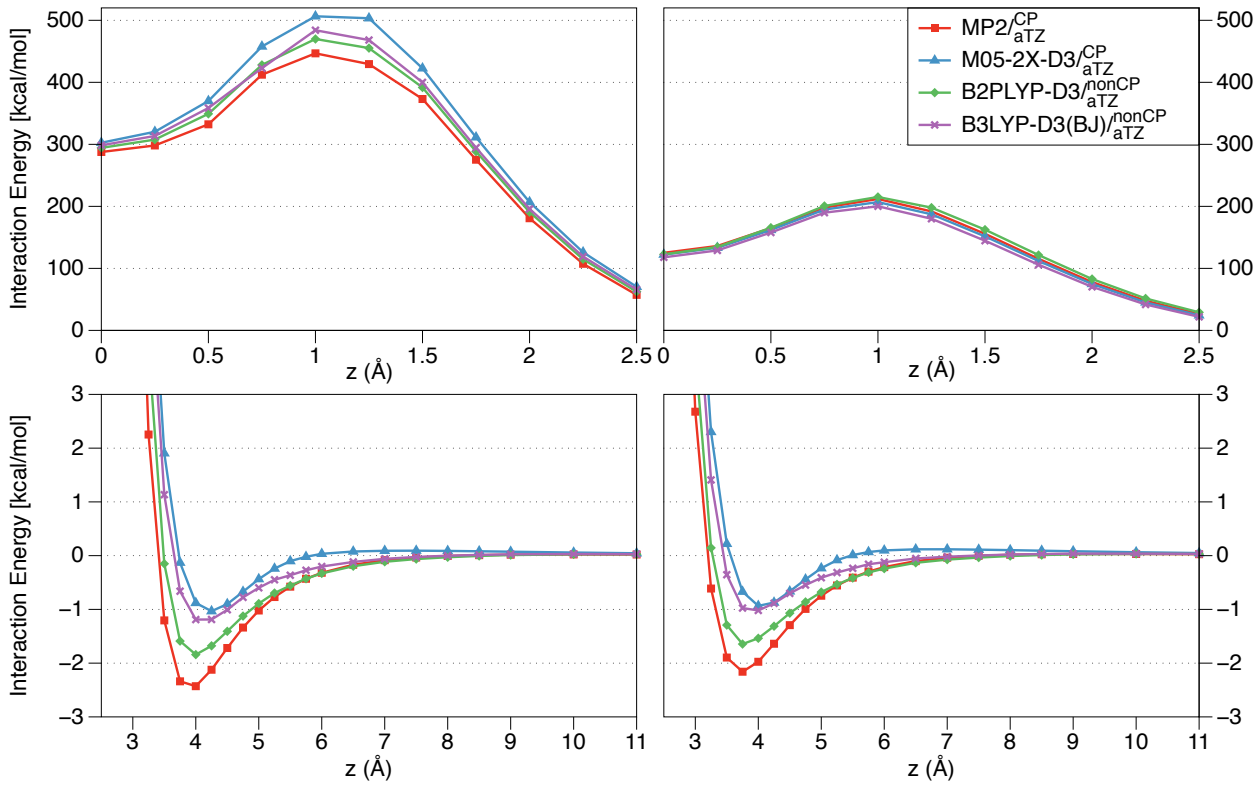

Figure 3.11: MP2 and DFT+D interaction energies (in kcal/mol) for the N3 vacancy–CO_2 (left panels) and N4 vacancy–CO_2 (right panels) complexes as functions of the distance z from the CO_2 carbon to the N-PHAC plane. The upper panels display interaction energies at the repulsive region while the lower panels show interaction energies at the minimum and long-range distances. The CO_2 molecule is located along the symmetry axis perpendicular to the N-PHAC plane as illustrated in Fig. 3.10. The interacting molecules are kept rigid.

barrier by 25.3 kcal/mol at the B3LYP/aDZ level. At the same time, the C–O bond length increases by 0.026 Å and the distance from the center of the N4 vacancy to the nitrogen atoms grows by 0.182 Å. Using the flexible geometry of the complex, we again performed single-point calculations at the B3LYP-D3/aTZ and M05-2X-D3/aTZ levels including the counterpoise and monomer deformation corrections, obtaining interaction energy values of 94.5 and 98.8 kcal/mol, respectively. Thus, the flexible energy barrier is still large enough to prevent the CO_2 molecule from passing through the N4 hole.

At $z = 4.0$ Å, close to the van der Waals minimum in Fig. 3.11, when CO_2 is oriented perpendicular to the N3 vacancy, the B2PLYP-D3/$^{\text{nonCP}}_{\text{aTZ}}$ interaction energy is -1.84 kcal/mol. However, an orientation of CO_2 parallel to the surface, at the same distance, is more favorable

with an interaction energy of -2.99 kcal/mol. At the same $z = 4.0$ Å, the B2PLYP-D3/$^{\text{nonCP}}_{\text{aTZ}}$ energies are -1.54 and -3.10 kcal/mol for the perpendicular and parallel N4 vacancy-CO_2 complexes, respectively.

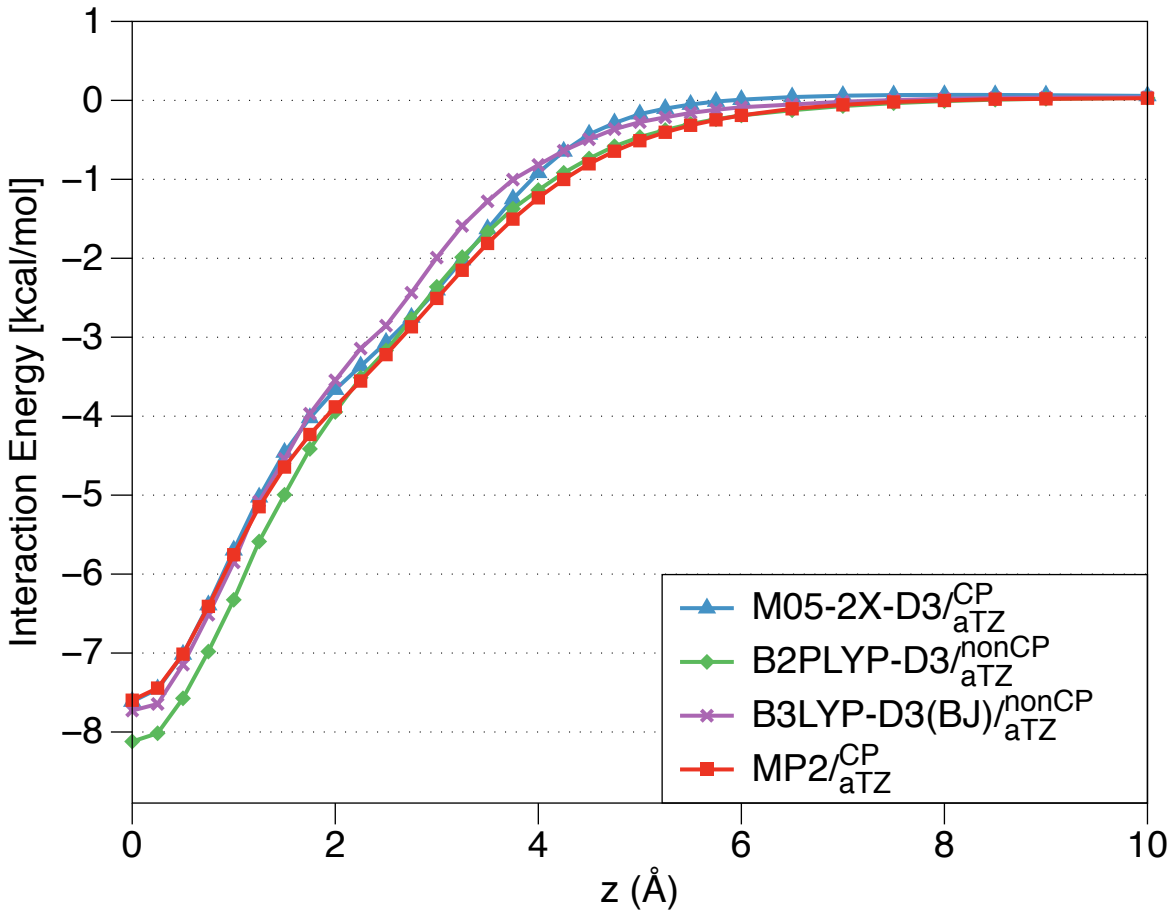

Figure 3.12: MP2 and DFT+D interaction energies (in kcal/mol) for the N4H4 vacancy–CO_2 complexes as functions of the distance z from the CO_2 carbon to the N-PHAC plane. The CO_2 molecule is located along the symmetry axis perpendicular to the N-PHAC plane as illustrated in Fig. 3.10.

It is interesting to explore a larger vacancy that allows the CO_2 molecule to move from one side of the N-PHAC to the other much easier than in the N3 and N4 cases. According to Ref. 58, the CO_2 molecule could pass the N4H4 vacancy almost freely. Therefore, we picked the N4H4 vacancy model, Fig. 3.10, as a representative of a larger hole to study the N-PHAC–CO_2 interaction energy. We kept the monomers rigid and restricted the symmetry to

D_{2h}. Figure 3.12 displays the interaction energies for the best DFT performers determined in Sec. 3.4.2: B3LYP-D3(BJ)/$^{\rm nonCP}_{\rm aTZ}$, M05-2X-D3/$^{\rm CP}_{\rm aTZ}$, and B2PLYP-D3/$^{\rm nonCP}_{\rm aTZ}$ as well as the CP-corrected MP2/aTZ interaction energies. From this figure, all methods predict a minimum when CO_2 is located at the center of the vacancy (in good agreement with the results of Ref. 58) with an interaction energy around -7.5 kcal/mol.

For comparison to the N3 and N4 vacancy models, we computed the B2PLYP-D3/$^{\rm nonCP}_{\rm aTZ}$ energies for the two parallel C_{2v} N4H4-hole-CO_2 complexes as well, at z = 4.0 Å. The interaction energies are -1.14 kcal/mol for the perpendicular orientation and -1.52, -1.72 kcal/mol for the two parallel C_{2v} N4H4 vacancy-CO_2 complexes (the oxygen atoms orient towards the nitrogens for the first one and towards the intra-vacancy hydrogens for the second one, respectively). Thus, the parallel orientations are energetically favorable at this distance. Overall, while the N3 and N4 vacancies are clearly too small for CO_2 (or, likely, any molecule) to pass through, the N4H4 hole is large enough, and provides a large enough dispersion interaction, to afford an (energetically) barrierless transition of CO_2 to the other side. This observation coincides, and the resulting well depth agrees quantitatively, with the findings of Ref. 58 which used the PBE-D2 level of theory and a wide (19,0) porous nanotube in place of graphene. The agreement between different approaches illustrated in Fig. 3.12 could not have been taken for granted, but it is highly gratifying, confirming that the DFT-based variants selected on the basis of their reproduction of benchmark data for smaller N-PHAC-CO_2 complexes are also appropriate, and consistent, for the study of carbon dioxide permeation through realistic models of porous N-doped graphene.

3.5 Summary

High-accuracy benchmark interaction energies were obtained for weakly interacting complexes of CO_2 with nitrogen-containing polyheterocyclic aromatic compounds, N-PHACs (pyrazine, pyridine, quinoline, quinoxaline, pyrido[3,2-g]quinoline, phenanthroline, 2-azapyrene, and 1,6-diazacoronene). The energies were computed by the supermolecular MP2 approach

extrapolated to the complete basis set limit plus a CCSD(T) correction calculated in a moderate basis set (up to aTZ for most 1- and 2-ring N-PHACs, aDZ for most 3- and 4-ring systems, and laDZ for 1,6-diazacoronene-CO_2). The calculations for 1- and 2-ring N-PHAC-CO_2 complexes (except for stacked quinoline-CO_2) utilized the explicitly correlated CCSD(T)-F12a/b approaches while all other systems were treated using conventional CCSD(T). An extensive basis set convergence analysis indicates that our benchmark interaction energies are accurate to a few hundredths of a kilocalorie per mole at the minimum separations. Our CCSD(T)-level results indicate that the global minimum structures for CO_2 interacting with N-PHACs are all in-plane. The ΔCCSD(T) correction is quite small (less than 0.08 kcal/mol at the minimum distance) for in-plane dimers. For the stacked configurations, the ΔCCSD(T) correction is significantly larger and the MP2 energies overbind by up to 1.8 kcal/mol at the minimum distance.

The newly developed CCSD(T)-level benchmarks were subsequently used to investigate the accuracy of several novel DFT approaches for the N-PHAC-CO_2 interaction energies. The comparisons included one-dimensional cuts through the N-PHAC-CO_2 potential energy surfaces passing through the lowest-energy structures for both the in-plane and stacked complexes, with distances ranging from 0.8 times the minimum to 1.6 times the minimum. Thus, the optimal DFT variant needs to provide a uniformly high accuracy for the entire potential energy curve, not just around the van der Waals minima. The tested approaches included M05-2X, M06-2X, B2PLYP, B3LYP, BLYP, PBE, PBE0, BP86, B97, and LC-ωPBE with the def2-SVP, TZVP, QZVP and Dunning aDZ and aTZ basis sets combined with five possible variants of Grimme's dispersion correction: -D2, -D3, -D3(BJ), -D3-$E^{(3)}$, and -D3(BJ)-$E^{(3)}$, both with and without the CP correction. In the largest, QZVP basis set, the three best approaches overall turned out to be B2PLYP-D3/nonCP, B2PLYP-D3(BJ)/nonCP, and M05-2X-D3/(both CP and nonCP), with mean unsigned relative errors on the 95 benchmark data points amounting to 6.4, 6.9, and 7.2%, respectively. Thus, a few DFT+D variants exhibit reasonable accuracy throughout the entire range of distances unlike the case of pristine

carbon nanotubes interacting with CO_2.[76] While a redesign of the atom-pairwise dispersion expression is not necessary for this work, the refitting of damping parameters performed for curved coronene-CO_2 complexes in Ref. 76 improved the performance of most DFT+D variants also for the N-PHAC-CO_2 models considered here.

The top performing DFT variants along with the MP2 approach were subsequently employed to study the barrier to a carbon dioxide transition through three model N-doped graphene pores. We found that only the largest of them, the N4H4 pore, is permeable to CO_2. For this pore, we obtained a quantitative agreement between all computed energy profiles and the results of Ref. 58. As the treatment of dispersion within the methods tested by us ranges from additive (DFT+D) to partially nonadditive (B2PLYP-D3) to fully non-additive (MP2) and no systematic discrepancies were observed as the model size increased, the pairwise-nonadditive effects on dispersion[44,126–128] are apparently not critical for the complexes considered in this work.

Acknowledgment

This research was supported by the Donors of the American Chemical Society Petroleum Research Fund, the NSF CAREER award CHE-1351978, and the startup funding from Auburn University.

Chapter 4

Evaluation of DFT-D variants suitable for nanotube adhesion forces

4.1 Introduction

Graphene and nanotubes exhibit a wide range of mechanical and electronic properties and have been proposed for a variety of applications. The construction of heterostructures by stacking different two-dimensional (2D) atomic crystals on top of each other is a new hot research topic in the past few years.[129] Graphene is the most important known 2D material. The 2D crystals are held together by weak adhesive forces which are responsible for the adsorption of all kinds of molecules to surfaces. Therefore, it is very intriguing to study the forces between small molecules and graphene. As we discussed in Chapter 1, the adsorption of small molecules on graphene or nanotubes is an example of physisorption. In this project, we are interested in the adhesive forces, which are the normal forces resisting the separation of adsorbate and adsorbent.[130] Understanding adhesive interactions is vital for interpreting the role that a special class of solubilizer molecules plays in the solubilization of solid carbon nanotubes (CNTs) in common organic solvents. Without suitable solubilizers,[131] solid CNTs cannot be dissolved in these solvents.[132] In order to dissolve a solid A into a solvent B, one has to rely on external forces, mainly stochastic forces coming from the solvent, to act on the solute molecules and pull them off the bulk against the resistance of adhesive forces that hold the solute molecules together. The efficiency of different solubilizers has strong dependence on their molecular structures, but this dependence can only be ascertained by detailed investigations of various model systems.

The quality of adhesive forces as derivatives of interaction energies with respect to intermolecular separation depends crucially on the quality of the energies. For the systems of interest, the total interaction energy is dominated by dispersion, as investigated in Chapter 3.

Accurate non-covalent interaction energies are obtained with high-level quantum theoretical methods, which are costly and can be applied to systems of only moderate size. Results from high-level quantum methods will be used as benchmark in this work. Therefore, one aim of this project is to find out an optimal DFT+D variant that can yield reliable forces for moderately sized systems, compared to the established high-accuracy benchmark values. Then, the next step is to modify computationally cheap low-level methods so that they can describe the forces accurately even for large systems that cannot be studied with high-level methods. Such low-level methods are necessary in molecular dynamics studies of such systems. The low-level method used in this project is the semiempirical density functional based tight-binding method (DFTB).[133]

4.2 The test set

In earlier studies[75,76] we have established the curved coronene molecule as a nanotube model that is both small enough to perform high-accuracy *ab initio* interaction energy calculations up to CCSD(T) and large enough to capture nearly all exterior adsorption energy as compared to larger models (the latter observation would not hold for interior adsorption). The curvature of the coronene molecule can be adjusted to model nanotubes of different types (zigzag/armchair/chiral) and different diameters. We have also investigated the dependence of benchmark interaction energies on the details of how the CCSD(T) complete basis set (CBS) limit is estimated, that is, the basis sets employed at different levels of theory and the presence/absence and details of the explicitly correlated (F12) treatment. Based on these investigations, we have generated[74–76,117] several sets of accurate benchmark interaction energies for models of nanotube adsorption that cover flat and curved aromatics, pristine and N-doped nanotubes, and CH_4 and CO_2 adsorbates, for a broad selection of curvatures, adsorbate orientations, and intermolecular distances (the latter aspect is especially crucial for estimating adhesion forces). In this work, we follow essentially the same approach.

4.3 Results and discussion

For the task of determining the performance of various DFT approaches for describing interactions of solubilizer molecules with nanotubes, we selected three sets of benchmark CCSD(T)/CBS data. These sets all involve a coronene molecule (flat or curved away from the adsorbate) interacting with **1** methane, **2** carbon dioxide, and **3** ethylene. Set **1** originates from Ref. 75 and has the same three angular orientations of the methane molecule, but the interior-adsorption geometries (those where the coronene molecule is curved towards the adsorbate) were removed. At the same time, the benchmark calculations were extended[134,135] to one shorter intermolecular distance ($R/R_{\mathrm{min}} = 0.8$) in line with the other datasets, for a total of 75 geometries. Set **2** originates from Ref. 76 (which already extends down to $R/R_{\mathrm{min}} = 0.8$) but the interior-adsorption geometries were removed for a total of 105 configurations. Set **3** has been constructed specifically for this project and involves the exterior-adsorption complexes of coronene-sized nanotube models (corresponding to the (5,5), (7,0), (9,0), and (12,0) nanotubes and flat graphene) interacting with ethylene at 5 intermolecular distances ($R/R_{\mathrm{min}} = 0.8, 0.9, 1.0, 1.2, 1.4$) and three symmetric angular orientations of the ethylene molecule — those in which one of the ethylene C_2 axes is perpendicular to the nanotube axis and passes through the center of the central coronene ring (Fig. 4.1). The geometry of the ethylene monomer is optimized at the DF-MP2/aTZ level. The curved coronene structures are obtained using the TUBEGEN program. Then, the desired fragment is cut from the TUBEGEN output so that its curvature corresponds to a particular zigzag (k,0) or armchair (n,n) nanotube. For each of the three C_2 axes, one has a choice of aligning one of the other two axes to be parallel with the nanotube axis (the two choices are equivalent for flat coronene and are related to each other by a 90-degree rotation around the intermolecular axis otherwise). We chose the possibility (depicted in Fig. 4.1) that led to a slightly stronger interaction (a lower interaction energy at the optimized intermolecular distance R_{min}). The value of R_{min} was optimized at the MP2/aug-cc-pVTZ level of theory

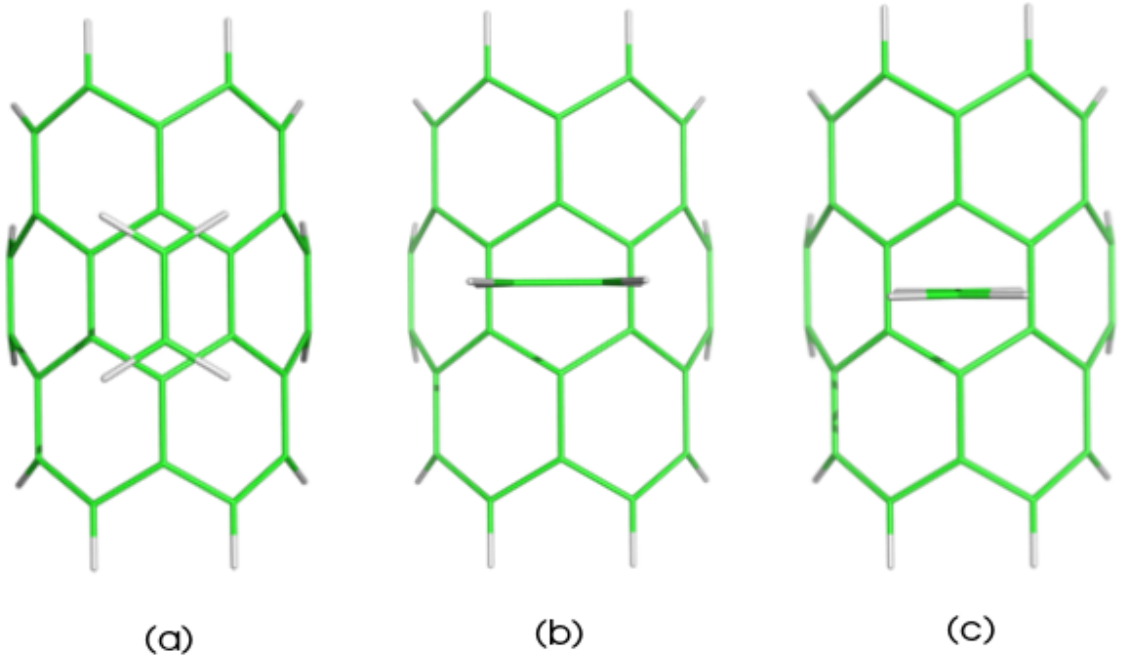

Figure 4.1: The ethylene-curved coronene configurations considered in this work (on the example of the (7,0) nanotube).

for each nanotube and each configuration. Thus, the overall set **3** consists of 75 configurations — all geometries and benchmark interaction energies are given in the Supporting Information to Ref. 135. The benchmark interaction energies have been calculated as

$$E_{\rm int} = E_{\rm int}^{\rm MP2}({\rm aTZ, aQZ}) + \Delta E_{\rm int}^{\rm CCSD(T)}({\rm laDZ}) \tag{4.1}$$

that is, the MP2 interaction energy was extrapolated to the complete-basis-set limit using the standard X^{-3} scheme,[91] and the post-MP2 interaction energy contribution $\Delta E_{\rm int}^{\rm CCSD(T)} = E_{\rm int}^{\rm CCSD(T)} - E_{\rm int}^{\rm MP2}$ was computed in the local-aug-cc-pVDZ≡laDZ basis set,[74] that is, the aDZ basis on all ethylene atoms and on the six coronene carbon atoms closest to the ethylene center of mass, and the nonaugmented cc-pVDZ basis on all remaining atoms of coronene. For the CO_2–curved coronene structures (Set **2**) the values of $\Delta E_{\rm int}^{\rm CCSD(T)}$ were obtained from explicitly correlated CCSD(T)-F12a and CCSD(T)-F12b calculations,[94,95] as explained in detail in Ref. 76; for the remaining two datasets, conventional CCSD(T) was employed. The 1*s* carbon electrons were not correlated. All wavefunction calculations were performed with the MOLPRO2012.1 code.[83]

The performance of different DFT variants on the combined 255-element dataset defined above will be evaluated using mean unsigned error (MUE) and mean unsigned relative error (MURE). It should be noted that this dataset does not feature any points close to where the potential energy curve crosses zero so no special weighting of relative errors is needed in contrast to Refs. 76 and 135. The DFT functionals considered include BLYP,[25] B3LYP,[23,24] B2PLYP,[26] BP86,[25] B97,[29,89] PBE,[101] PBE0,[102,103] and LC-ωPBE.[104] These functionals were augmented with Grimme's pairwise dispersion corrections -D3[34] and -D3(BJ)[105] using both original damping parameters and the parameters refitted in Ref. 135 to a broad database of CCSD(T)-level intermolecular interaction energies (the latter will be labeled by the "refit" subscript). Note that both refitted variants involve three damping parameters, similar to -D3(BJ) and unlike -D3 which involves only two parameters. All DFT calculations were

performed in the def2-QZVP basis set and include the counterpoise (CP) correction so they can be viewed as near CBS limit values for the respective functionals. Additionally, CP-corrected calculations in the aDZ basis were carried out to examine a more computationally efficient protocol. The resulting MUE and MURE values are presented in Tables 4.1 (QZVP basis) and 4.2 (aDZ basis).

Table 4.1: Mean unsigned errors (MUE, in kcal/mol) and mean unsigned relative errors (MURE, in percent) for different DFT-D approaches with respect to the CCSD(T)-level benchmark values for the full 255-element dataset presented here. All DFT computations utilized the def2-QZVP basis with the CP correction. The best performers are indicated in bold.

	MUE				MURE			
Functional	-D3	-D3(BJ)	-$D3_{3,\text{refit}}$	-$D3(BJ)_{\text{refit}}$	-D3	-D3(BJ)	-$D3_{3,\text{refit}}$	-$D3(BJ)_{\text{refit}}$
B2PLYP	0.172	0.118	**0.073**	0.124	6.61	5.65	**4.79**	6.87
B3LYP	0.404	0.215	0.265	**0.161**	16.11	10.80	13.14	**9.97**
BLYP	0.458	0.256	0.240	0.235	20.15	13.08	13.72	14.38
BP86	0.321	0.347	0.319	0.396	20.58	18.16	19.35	21.37
PBE0	0.330	0.242	0.243	**0.171**	16.90	14.42	14.21	**9.98**
PBE	0.449	0.280	0.296	**0.179**	22.98	15.96	16.28	**10.29**
LC-ωPBE	0.174	0.223	0.163	0.239	10.03	13.03	9.88	13.05
B97	0.483	0.461	0.297	0.343	22.55	27.56	14.84	18.90

Table 4.2: Mean unsigned errors (MUE, in kcal/mol) and mean unsigned relative errors (MURE, in percent) for different DFT-D approaches with respect to the CCSD(T)-level benchmark values for the full 255-element dataset presented here. All DFT computations utilized the aug-cc-pVDZ basis with the CP correction. The best performers are indicated in bold.

	MUE				MURE			
Functional	-D3	-D3(BJ)	-$D3_{3,\text{refit}}$	-$D3(BJ)_{\text{refit}}$	-D3	-D3(BJ)	-$D3_{3,\text{refit}}$	-$D3(BJ)_{\text{refit}}$
B2PLYP	0.275	0.201	**0.099**	**0.097**	10.58	8.92	**4.55**	**5.15**
B3LYP	0.414	0.213	0.275	**0.163**	16.49	10.68	13.32	**9.98**
BLYP	0.453	0.250	0.245	0.240	19.90	13.02	13.83	14.57
BP86	0.332	0.354	0.329	0.407	20.90	18.56	19.67	21.83
PBE0	0.329	0.248	0.233	**0.148**	16.44	14.26	13.24	**8.88**
PBE	0.451	0.283	0.284	**0.174**	22.95	15.99	15.64	**9.99**
LC-ωPBE	0.177	0.220	0.164	0.226	9.68	13.02	9.61	12.63
B97	0.470	0.472	0.297	0.356	22.30	28.65	15.38	19.93

The approaches that perform particularly well for a given degree of computational complexity (double hybrid and range-separated functionals are more demanding than conventional hybrid functionals which in turn are more demanding than GGAs) have been marked in bold. It is apparent that the performance of different DFT-D variants changes very little between the two bases which validates the use of the small aDZ set (with the CP correction) for reasonably accurate calculations on larger models. Moreover, while the double-hybrid B2PLYP-D3 approach performs significantly better than the lower-rung functionals, the top-performing GGA and hybrid variants exhibit very similar errors. In particular, the simple PBE-D3(BJ)$_{\text{refit}}$ variant emerges as an optimal combination of accuracy and efficiency for weakly interacting complexes of this kind.

4.4 Future work

The first task in the future will be to make a comparison of results obtained with high- and low-level methods. The benchmark interaction energies will be compared with interaction energies computed with DFTB plus an empirical dispersion correction. In this step, the hope is that reparametrizing the dispersion corrections, both Grimme's D3(BJ) and Hobza et al. original correction for DFTB,[136] will allow to bring low and high-level energies in good agreement and thus also yield good adhesive forces.

The following step is to consider using a QM/QM method with DFTB as the low-level method and the PBE-D3(BJ) with the reparametrized damping function as the high-level method. With this method we want to investigate systems with large adsorbates on CNTs of different curvature at a higher level than DFTB+D. The investigations with the low level method will be performed at the University of Graz, in the group of Dr. Sax.

Chapter 5

Description of the interactions between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen via local methods

5.1 Introduction

In general, a CCSD(T)/CBS computation,[137] the *gold standard*, is the preferred way to generate high-accuracy weak interaction energies for many systems. It is useful for testing more efficient computational tools, like a newly developed density functional. However, the requirement of a computationally demanding CCSD(T) calculation with a feasible basis set is the obstacle for studying intermolecular interactions between medium-sized molecules. Even worse, one cannot perform these calculations on somewhat larger systems. In our case, the systems involving molecules slightly larger than 2-azapyrene cannot be computed by CCSD(T) with the aug-cc-pVDZ basis. The application of traditional CCSD(T) calculations is typically limited to complexes smaller than 40 atoms. We can only offer CCSD(T)/laDZ for 1,6-diazacoronene to generate CCSD(T)-level results. Thus, it is impossible to utilize CCSD(T) for even larger systems such as 1,10-diazacircumcoronene. The geometry of this 1,10-diazacircumcoronene-CO_2 complex is displayed in Fig. 5.1. The motivation for going after 1,10-diazacircumcoronene-CO_2 and similar models in this project is checking if (1) the two-body nonadditivity effects become important and if (2) the many-body dispersion (MBD)[44] describes them well. In the work of Ref. 138 from Alexandre Tkatchenko and coworkers, the authors stated that the non-additive many-body dispersion (MBD) energy beyond the standard pairwise approach is critical for both the correct qualitative and quantitative description of polymorphism in molecular crystals. Recently, another study of Alexandre Tkatchenko and coworkers reported that the ubiquitous Van der Waals forces between polarizable nonmetallic nanostructures can be more completely understood in terms

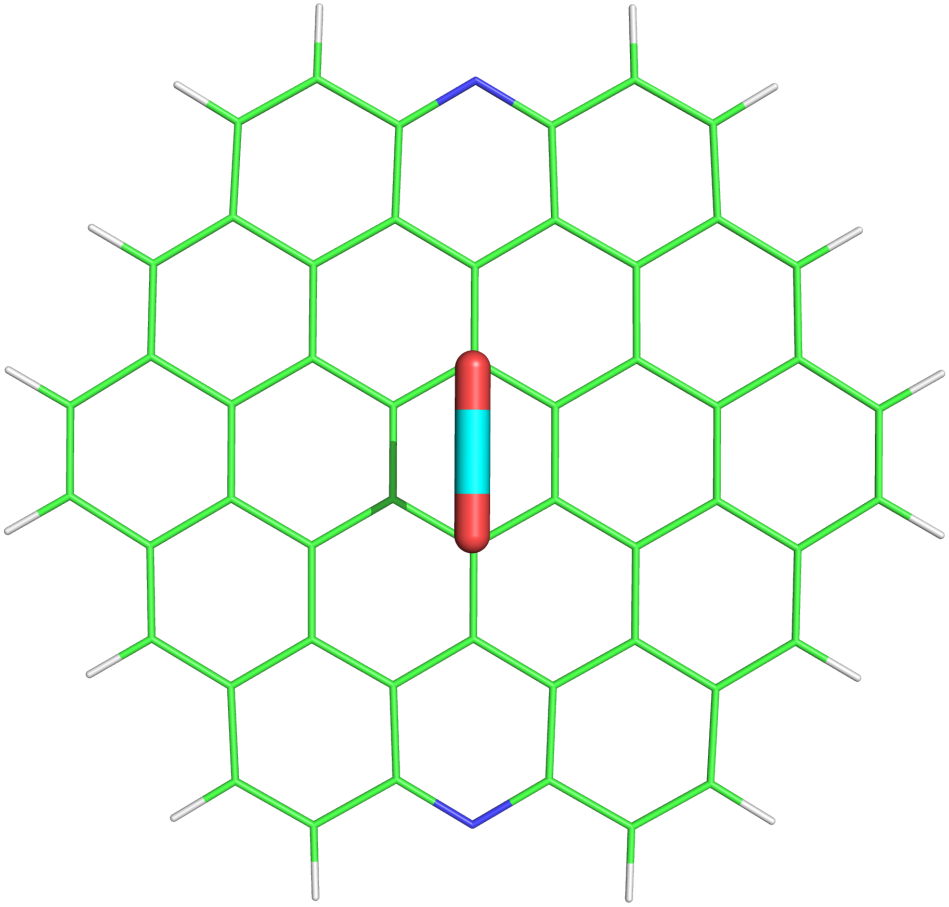

Figure 5.1: The 1,10-diazacircumcoronene-CO_2 configuration.

of collective interactions between wavelike charge density fluctuations, instead of simply a summation over pairwise interactions between instantaneous particle- or fragment-like dipolar fluctuations.[139] The methods we were using in our projects so far are often called wavefunction theory (WFT) based methods. If one goes beyond WFT, one of the promising alternatives is another family of methods that can describe noncovalent interactions for large systems: Quantum Monte Carlo (QMC). QMC is a family of stochastic methods for solving quantum many-body problems such as the stationary Schrödinger equation based on techniques such as use of stochastic processes and sampling of wave functions in the space of electron positions.[46,140] Even though we cannot obtain conventional CCSD(T) results for such large systems, we still need the same level of theory to assess the performance of DFT results and generate CCSD(T)-level benchmark values. If we stick to WFT methods in this thesis, it leaves us with only one way: local coupled cluster methods.

There are two basic categories for existing local correlation approaches.[141] The methods in the first category avoid the fragmentation of the system and solve the corresponding equations or compute perturbative corrections for the entire system at the same time. The main idea of these methods is proposed by the pioneering works of Pulay,[142,143] who adopted localized occupied molecular orbitals (MOs) and projected atomic orbitals (AOs) to describe the virtual space. For each occupied orbital pair a domain of spatially close virtual orbitals was built and employed to choose the allowed configurations in the wave function. The representative methods included in this category are Werner and co-workers'[144,145] and Schütz and co-workers'[146,146] various local correlation methods including local CC with single and double excitations (LCCSD) and CCSD with perturbative triples correction [LCCSD(T)], the pair natural orbital (PNO) CC methods of Neese and co-workers[82,147,148] and the orbital-specific virtuals (OSVs) approximation of Chan, Manby, and co-workers.[81,149]

The second category of the local correlation methods decomposes the system into fragments of manageable size, and thus the correlation energy is obtained as the sum of the

contributions of the individual fragments and of interactions thereof. Methods in this category are the local CC model based on the fragment MO theory,[150] cluster-in-molecule (CIM)[151] local correlation formalism of CCSD(T) (CIM-CCSD(T)), and divide-and-conquer (DC) methods.[152] Other methods also included in this subgroup are incremental methods.[153] In the incremental scheme, the adsorption energy is obtained as the sum of the HF energy, E^{HF}, of the system and the correlation contribution, E^{corr}.

$$E_{ads} = E^{HF} + E^{corr} \tag{5.1}$$

The localized orbitals are obtained from the HF canonical orbitals by using the Foster-Boys[154] procedure. Then the E^{corr} term is expanded in contributions from the correlation of electrons in these localized orbital groups by using the method of increments at a highly accurate quantum-chemical level in the following formula:

$$E^{corr} = \eta_{mol} + \sum_i \eta_i + \sum_i \eta_{mol,i} + \sum_{i>j} \eta_{ij} + \sum_{i>j} \eta_{mol,ij} + ... \tag{5.2}$$

where η_{mol} is the correlation energy change of molecule A due to adsorption, i, j, ... are the individual localized orbital groups of the surface or macromolecule B. The first two terms can be called one-body contributions because the correlation energy comes from one orbital group. $\sum_i \eta_{mol,i}$ and $\sum_{i>j} \eta_{ij}$ are the two-body contributions and so forth.

5.2 Methods and computational details

In order to assess the performance of these local methods, we adapted the geometries from Ref. 117, since we already established the highly accurate benchmark interaction energies for these configurations. Different from Ref. 117, we currently just tested the local methods on the minimum structures, instead of at different distances R. The geometries were presented in Fig. 3.1.

Table 5.1: The local CCSD(T) and conventional CCSD(T) interaction energies (in kcal/mol), and the corresponding timings (in hours), for the lowest-energy structure of the in-plane pyrazine-CO_2 complex in the aDZ basis.

method	interaction energy	time
NONCP CCSD(T)	-4.621	0.13
NONCP LCCSD(T)	-3.266	8.60
CCSD(T)	-3.245	0.13
LCCSD(T) with INTERACT=1	-3.118	29.53
DF-LCCSD(T)	-3.094	26.17
LCCSD(T)	-3.192	8.65
OSV-CCSD(T) with INTERACT=1	-2.990	1.62
DLPNO-CCSD(T)	-3.116	0.80

The local methods tested were OSV-CCSD(T),[155] LCCSD(T),[156] DF-LCCSD(T),[157] and DLPNO-CCSD(T).[148] DLPNO is the shorthand for domain based local pair natural orbital. For comparison, we also gathered the results from the frozen natural orbitals FNO-CCSD(T) method[158] in order to have a full picture for the local methods. One should notice that FNO-CCSD(T) is not a local method. All conventional CCSD(T) results were obtained from the previous project, i.e., Ref. 117. ORCA[159] was used to generate the data for DLPNO-CCSD(T) and FNO-CCSD(T) was computed with PSI4.[109] All other results were obtained using MOLPRO.[83] To have a better understanding of these calculations, their timings (in h) also will be included in Sec. 5.3.

5.3 Results and discussion

In order to find out a suitable local method to describe the interaction between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen, we first tested out the OSV-CCSD(T), LCCSD(T), DF-LCCSD(T), and DLPNO-CCSD(T) results for the in-plane pyrazine-CO_2 complex, computed in the aDZ basis. The interaction energies were gathered in Table 5.1. In this table, it is demonstrated that the LCCSD(T) method does reduce the basis set superposition error (BSSE), as claimed in the original literature (Ref.

156), as the difference between the NONCP and CP LCCSD(T) is only 0.074 kcal/mol. On the other hand, the conventional CCSD(T) is suffering from the BSSE, since the difference between NONCP CCSD(T) and CCSD(T) is larger than 1.3 kcal/mol. The *INTERACT=1* option in Table 5.1 means that individual molecules are identified automatically and all intermolecular pairs are automatically treated as strong pairs and included in the LCCSD treatment. However, this will make the LCCSD(T) calculations very expensive. On the other hand, the inclusion of this command does not help the results much, so we will not include it in later computations. The main problem for these local methods is the long running time. One can easily obtain the conventional high-accuracy CCSD(T) result within minutes, however, the calculations for local methods took hours to accomplish. Since the OSV-CCSD(T) result does not show acceptable accuracy, we will drop this method from now on. The conclusion that can be drawn from Table 5.1 is that DLPNO-CCSD(T) can produce reasonable results in an acceptable time. We want to obtain fairly accurate results ("silver standard") for systems with at least 50 atoms. In this manner, it is time to turn our attention to larger systems. The results for CCSD(T), DF-LCCSD(T), LCCSD(T), and DLPNO-CCSD(T) are presented in Table 5.2 for one- and two-ring systems of Ref. 117 in the aDZ basis.

It can be seen in Table 5.2 that the density fitting (DF) approximation is very helpful to reduce the running time, compared with the conventional scheme. The conventional LCCSD(T) calculations are very expensive, even when the systems contain only one ring. The running times for DF-LCCSD(T) are not so large for one-ring systems, however, they skyrocket for two-ring configurations. The running time is 472.5 h for stacked quinoxaline-CO_2, therefore, it is not feasible to investigate the interaction energies for larger systems via either LCCSD(T) or DF-LCCSD(T). Based on Table 5.2, DLPNO-CCSD(T) is the most promising local approach as the calculation can be done in less time than conventional CCSD(T). One can also notice that the differences between the results of DLPNO-CCSD(T)

Table 5.2: The local CCSD(T) and conventional CCSD(T) interaction energies (in kcal/mol), and timings (in hours), for the lowest-energy structures for one- and two-ring systems from Fig. 3.1, in the aDZ basis.

complex	method	interaction energy	time
in-plane pyrazine-CO_2	CCSD(T)	-3.245	0.13
	DLPNO-CCSD(T)	-3.116	0.80
	DF-LCCSD(T)	-3.094	26.17
	LCCSD(T)	-3.192	8.65
stacked pyrazine-CO_2	CCSD(T)	-0.735	0.13
	DLPNO-CCSD(T)	-0.732	0.9
	DF-LCCSD(T)	-0.931	5.0
	LCCSD(T)	-0.543	45.4
in-plane pyridine-CO_2	CCSD(T)	-3.636	0.16
	DLPNO-CCSD(T)	-3.464	0.2
	DF-LCCSD(T)	-3.336	2.2
	LCCSD(T)	-3.292	16.7
stacked pyridine-CO_2	CCSD(T)	-1.630	0.48
	DLPNO-CCSD(T)	-1.741	0.2
	DF-LCCSD(T)	-1.755	3.3
	LCCSD(T)	-1.373	22.1
in-plane quinoline-CO_2	CCSD(T)	-3.994	5.1
	DLPNO-CCSD(T)	-3.713	2.2
	DF-LCCSD(T)	-3.318	372.4
stacked quinoline-CO_2	CCSD(T)	-2.121	15.5
	DLPNO-CCSD(T)	-2.295	0.9
	DF-LCCSD(T)	-2.215	721.3
in-plane quinoxaline-CO_2	CCSD(T)	-3.723	4.9
	DLPNO-CCSD(T)	-4.634	0.8
	DF-LCCSD(T)	-3.539	278.1
stacked quinoxaline-CO_2	CCSD(T)	-2.192	3.6
	DLPNO-CCSD(T)	-2.459	0.9
	DF-LCCSD(T)	-2.329	472.5

and CCSD(T) are less than 0.3 kcal/mol, except in the case of in-plane quinoxaline-CO_2, where the difference is 0.911 kcal/mol. This result is out of the acceptable range.

Even though changing the default thresholds (T_{CutPNO}=3.33×10^{-7}, $T_{CutPairs}$=1×10^{-4}, T_{CutMKN}=1 × 10^{-3}) was not recommended in the ORCA documentation, it is logical to play with these thresholds trying to reduce the error. Here is the original statement about these thresholds from the ORCA manual: *(a) T_{CutPNO} controls the number of PNOs per electron pair. This is the most critical parameter and has a default value of 3.33e-7. (b) $T_{CutPairs}$ controls a perturbative selection of significant pairs and has a default value of 1e-4. (c) T_{CutMKN} is a technical parameter and controls the size of the fit set for each electron pair. It has a default value of 1e-3. All of these default values are conservative. Hence, no adjustment of these parameters is necessary. All DLPNO-CCSD truncations are bound to these three truncation parameters and should not be touched (Hence they are also not documented :-)).*

Apparently, we have to play with these thresholds in order to obtain accurate results for systems like quinoxaline-CO_2. The results of changing the first and second threshold are displayed in Table 5.3. In this table, the combination of the first two thresholds still could not reduce the error into the reasonable range. Even worse, the results are fluctuating instead of monotonously increasing or decreasing. One should notice that in this table, the other threshold values are the default ones. The conclusion from this table is that the combination of T_{CutPNO} amd $T_{CutPairs}$ is not enough to minimize the difference between the DLPNO-CCSD(T) and conventional CCSD(T) methods. In this manner, the third threshold, T_{CutMKN}, will kick in. From the results shown in Table 5.4, one can conclude that once T_{CutMKN} is tightened to at least 1×10^{-4}, then the DLPNO-CCSD(T) calculation could give the difference between DLPNO-CCSD(T) and CCSD(T) within 0.236 kcal/mol. When T_{CutMKN} is tighter than 1×10^{-7}, the DLPNO-CCSD(T) result is converged to a value of -3.487 kcal/mol.

Table 5.3: The DLPNO-CCSD(T) interaction energies with various combinations of thresholds, and conventional CCSD(T) interaction energies (in kcal/mol) for the lowest-energy structures of two N-PHAC-CO_2 complexes.

method	T_{CutPNO}	$T_{CutPairs}$	interaction energy
in-plane quinoxaline-CO_2			
DLPNO-CCSD(T)/aDZ	3.33×10^{-3}	1×10^{-4}	-4.905
	8.33×10^{-4}	1×10^{-4}	-4.961
	3.33×10^{-4}	1×10^{-4}	-4.657
	9.33×10^{-5}	1×10^{-4}	-4.563
	8.33×10^{-5}	1×10^{-4}	-4.260
	6.33×10^{-5}	1×10^{-4}	-4.989
	4.33×10^{-5}	1×10^{-4}	-4.571
	3.33×10^{-5}	1×10^{-4}	-4.430
	3.33×10^{-6}	1×10^{-4}	-4.623
	2.33×10^{-6}	1×10^{-4}	-4.561
	4×10^{-7}	1×10^{-4}	-4.668
	3.33×10^{-7}	1×10^{-4}	-4.634
	3.33×10^{-8}	1×10^{-4}	-4.744
	2.33×10^{-8}	1×10^{-4}	-4.747
	2.33×10^{-8}	1×10^{-5}	-4.676
	2.33×10^{-8}	1×10^{-6}	-4.630
	2.33×10^{-8}	1×10^{-7}	-4.630
	2.33×10^{-9}	1×10^{-4}	-4.764
	2.33×10^{-9}	1×10^{-7}	-4.706
	2.33×10^{-10}	1×10^{-4}	-4.770
	2.33×10^{-10}	1×10^{-6}	-4.738
CCSD(T)/aDZ			-3.723
in-plane 1,6-diazacoronene-CO_2			
DLPNO-CCSD(T)/laDZ	4×10^{-6}	1×10^{-4}	-3.076
	4×10^{-7}	1×10^{-4}	-2.341
	3.33×10^{-7}	1×10^{-4}	-3.174
	4×10^{-8}	1×10^{-4}	-3.234
DLPNO-CCSD(T)/aDZ	3.33×10^{-7}	1×10^{-4}	-3.835
CCSD(T)/laDZ			-4.009

Table 5.4: The local CCSD(T) and conventional CCSD(T) interaction energies (in kcal/mol) for the lowest-energy in-plane quinoxaline-CO_2 complex in the aDZ basis. The T_{CutPNO} and $T_{CutPairs}$ thresholds are fixed at 3.33×10^{-7} and 1×10^{-4}, respectively.

method	T_{CutMKN}	interaction energy
DLPNO-CCSD(T)	default(1×10^{-3})	-4.634
	1×10^{-2}	-3.398
	7.5×10^{-3}	-3.391
	5×10^{-3}	-3.515
	2.5×10^{-3}	-3.526
	7.5×10^{-4}	-4.637
	5×10^{-4}	-4.620
	2.5×10^{-4}	-4.589
	1×10^{-4}	-3.525
	7.5×10^{-5}	-3.528
	5×10^{-5}	-3.523
	2.5×10^{-5}	-3.516
	1×10^{-5}	-3.502
	7.5×10^{-6}	-3.498
	5×10^{-6}	-3.494
	2.5×10^{-6}	-3.491
	1×10^{-6}	-3.489
	1×10^{-7}	-3.487
	1×10^{-8}	-3.487
	1×10^{-9}	-3.487
	1×10^{-10}	-3.487
	1×10^{-11}	-3.487
	1×10^{-12}	-3.487
CCSD(T)		-3.723

Therefore, it seems that the appropriate thresholds to produce reasonable results for the in-plane quinoxaline-CO_2 have been found. Once we got this T_{CutMKN} parameter settled, we went back to all the other systems with the newly found value. In order to be on the safe side, we decided for the value $T_{CutMKN} = 1\times10^{-5}$.

The interaction energies with the third threshold $T_{CutMKN} = 1\times10^{-5}$ (the other two kept as default) are presented in Table 5.5. In particular, the calculation with this threshold for in-plane quinoxaline-CO_2 has a energy difference of 0.221 kcal/mol compared to the CCSD(T) result. This agreement is satisfactory. All the other results also are within the acceptable range except for stacked 2-azapyrene-CO_2 and larger systems. Therefore, the third threshold individually cannot lead to DLPNO-CCSD(T) interaction energies that are accurate enough. Based on all the results in Tables 5.1 - 5.5, another strategy to reduce the error for the case of stacked 2-azapyrene-CO_2 is to use a combination of the three thresholds together. The information from all difference combinations of the three thresholds is shown in Table 5.6. In this table, we can tell that the $T_{CutPairs}$ value has strong impact on the interaction energies. In order to have a converged result, it has to be tighter than 1×10^{-5}. Once this threshold is settled, all the interaction energies are less than -2.0 kcal/mol for in-plane quinoxaline-CO_2. It is shown again in this table that a single threshold of T_{CutMKN} cannot reduce the error. For instance, if we only change T_{CutMKN} to 1×10^{-9}, the interaction energy of in-plane quinoxaline-CO_2 is -2.726 kcal/mol, which gives a difference of 0.496 kcal/mol. This is not accurate enough. Therefore, a combination of T_{CutPNO}, $T_{CutPairs}$, and T_{CutMKN} with values of 3.33×10^{-8}, 1×10^{-6}, and 1×10^{-6}, respectively, is recommended based on the results of Table 5.6.

At this point, we are done with playing with the three documented thresholds, however, we are trying our best to obtain high-accuracy results out of ORCA. In this spirit, we tried one more threshold, the T_{CutTNO} one. This threshold is related to the accuracy for the (T) triple excitation correction. Its default value is 1×10^{-7}. Once again, we combined this threshold with the previous three to gather the values in Table 5.7.

Table 5.5: The DLPNO-CCSD(T) (with $T_{cutMKN} = 1\times10^{-5}$) and conventional CCSD(T) interaction energies (in kcal/mol) for the lowest-energy N-PHAC-CO_2 structures. Unless stated otherwise, the aDZ basis set was used.

complex	method	T_{CutMKN}	interaction energy
in-plane pyrazine-CO_2	CCSD(T)		-3.245
	DLPNO-CCSD(T)	1×10^{-5}	-3.079
stacked pyrazine-CO_2	CCSD(T)		-0.735
	DLPNO-CCSD(T)	1×10^{-5}	-0.728
in-plane pyridine-CO_2	CCSD(T)		-3.636
	DLPNO-CCSD(T)	1×10^{-5}	-3.430
stacked pyridine-CO_2	CCSD(T)		-1.630
	DLPNO-CCSD(T)	1×10^{-5}	-1.666
in-plane quinoline-CO_2	CCSD(T)		-3.994
	DLPNO-CCSD(T)	1×10^{-5}	-3.668
stacked quinoline-CO_2	CCSD(T)		-2.121
	DLPNO-CCSD(T)	1×10^{-5}	-2.208
in-plane quinoxaline-CO_2	CCSD(T)		-3.723
	DLPNO-CCSD(T)	1×10^{-5}	-3.502
stacked quinoxaline-CO_2	CCSD(T)		-2.192
	DLPNO-CCSD(T)	1×10^{-5}	-2.411
in-plane pyrido[3,2-g]quinoline-CO_2	CCSD(T)		-3.852
	DLPNO-CCSD(T)	1×10^{-5}	-3.586
stacked pyrido[3,2-g]quinoline-CO_2	CCSD(T)		-2.288
	DLPNO-CCSD(T)	1×10^{-5}	-2.461
in-plane phenanthroline-CO_2	CCSD(T)		-4.706
	DLPNO-CCSD(T)	1×10^{-5}	-4.660
stacked phenanthroline-CO_2	CCSD(T)		-3.298
	DLPNO-CCSD(T)	1×10^{-5}	-3.434
in-plane 2-azapyrene-CO_2	CCSD(T)		-3.811
	DLPNO-CCSD(T)	1×10^{-5}	-3.671
stacked 2-azapyrene-CO_2	CCSD(T)		-2.230
	DLPNO-CCSD(T)	1×10^{-3}	-2.776
		1×10^{-5}	-2.735
		1×10^{-6}	-2.726
in-plane 1,6-diazacoronene-CO_2	CCSD(T)/laDZ		-4.009
	DLPNO-CCSD(T)/laDZ	1×10^{-5}	-3.372
	DLPNO-CCSD(T)/aDZ	1×10^{-3}	-3.835

Table 5.6: The local CCSD(T) and conventional CCSD(T) interaction energies (in kcal/mol) for the lowest-energy structure of the stacked 2-azapyrene-CO_2 complex in the aDZ basis.

method	T_{CutPNO}	$T_{CutPairs}$	T_{CutMKN}	interaction energy
DLPNO-CCSD(T)	8.33×10^{-7}	1×10^{-5}	1×10^{-5}	-1.762
	3.33×10^{-7}	7.5×10^{-4}	1×10^{-5}	-1.732
	3.33×10^{-7}	5×10^{-4}	1×10^{-5}	-2.612
	3.33×10^{-7}	2.5×10^{-4}	1×10^{-5}	-1.738
	3.33×10^{-7}	1×10^{-4}	1×10^{-6}	-2.726
	3.33×10^{-7}	1×10^{-4}	1×10^{-9}	-2.726
	3.33×10^{-7}	1×10^{-5}	1×10^{-5}	-1.720
	8.33×10^{-8}	1×10^{-5}	1×10^{-5}	-1.894
	3.33×10^{-8}	7.5×10^{-4}	1×10^{-5}	-2.897
	3.33×10^{-8}	7.5×10^{-4}	1×10^{-6}	-2.889
	3.33×10^{-8}	5×10^{-4}	1×10^{-5}	-2.644
	3.33×10^{-8}	5×10^{-4}	1×10^{-6}	-2.637
	3.33×10^{-8}	2.5×10^{-4}	1×10^{-5}	-2.672
	3.33×10^{-8}	2.5×10^{-4}	1×10^{-6}	-2.665
	3.33×10^{-8}	1×10^{-4}	1×10^{-5}	-2.826
	3.33×10^{-8}	1×10^{-5}	1×10^{-6}	-1.958
	3.33×10^{-8}	1×10^{-6}	1×10^{-6}	-1.765
	3.33×10^{-8}	1×10^{-7}	1×10^{-7}	-1.763
CCSD(T)				-2.230

Table 5.7: The DLPNO-CCSD(T) (with four thresholds) and conventional CCSD(T) interaction energies (in kcal/mol) for the lowest-energy structure of the stacked 2-azapyrene-CO_2 complex in the aDZ basis.

method	T_{CutPNO}	$T_{CutPairs}$	T_{CutMKN}	T_{CutTNO}	interaction energy
DLPNO-CCSD(T)	8.33×10^{-7}	1×10^{-5}	1×10^{-5}	1×10^{-7}	-1.762
	3.33×10^{-7}	5×10^{-4}	1×10^{-5}	1×10^{-7}	-2.612
	3.33×10^{-7}	1×10^{-4}	1×10^{-3}	1×10^{-7}	-2.776
	3.33×10^{-7}	1×10^{-4}	1×10^{-5}	1×10^{-7}	-2.735
	3.33×10^{-7}	1×10^{-5}	7.5×10^{-4}	1×10^{-7}	-1.732
	3.33×10^{-7}	1×10^{-5}	2.5×10^{-4}	1×10^{-7}	-1.738
	3.33×10^{-7}	1×10^{-5}	1×10^{-5}	1×10^{-7}	-1.720
	8.33×10^{-8}	1×10^{-5}	1×10^{-5}	1×10^{-7}	-1.894
	3.33×10^{-8}	7.5×10^{-4}	1×10^{-5}	1×10^{-7}	-2.897
	3.33×10^{-8}	7.5×10^{-4}	1×10^{-6}	1×10^{-7}	-2.889
	3.33×10^{-8}	5×10^{-4}	1×10^{-5}	1×10^{-7}	-2.644
	3.33×10^{-8}	5×10^{-4}	1×10^{-6}	1×10^{-7}	-2.637
	3.33×10^{-8}	2.5×10^{-4}	1×10^{-5}	1×10^{-7}	-2.672
	3.33×10^{-8}	2.5×10^{-4}	1×10^{-6}	1×10^{-7}	-2.665
	3.33×10^{-8}	1×10^{-4}	1×10^{-5}	1×10^{-7}	-2.826
	3.33×10^{-8}	1×10^{-4}	1×10^{-5}	1×10^{-8}	-2.848
	3.33×10^{-8}	1×10^{-5}	1×10^{-4}	1×10^{-7}	-1.966
	3.33×10^{-8}	1×10^{-5}	1×10^{-4}	1×10^{-8}	-2.028
	3.33×10^{-8}	1×10^{-5}	1×10^{-5}	1×10^{-7}	-1.965
	3.33×10^{-8}	1×10^{-5}	1×10^{-5}	1×10^{-8}	-2.027
	3.33×10^{-8}	1×10^{-5}	1×10^{-6}	1×10^{-7}	-1.958
	3.33×10^{-8}	1×10^{-5}	1×10^{-6}	1×10^{-8}	-2.020
	3.33×10^{-8}	1×10^{-5}	1×10^{-6}	1×10^{-9}	-2.034
	3.33×10^{-8}	1×10^{-5}	1×10^{-6}	1×10^{-10}	-2.038
	3.33×10^{-8}	5×10^{-6}	5×10^{-6}	1×10^{-7}	-1.806
	3.33×10^{-8}	1×10^{-6}	1×10^{-6}	1×10^{-7}	-1.765
	3.33×10^{-8}	1×10^{-6}	1×10^{-6}	1×10^{-8}	-1.830
	3.33×10^{-8}	1×10^{-7}	1×10^{-7}	1×10^{-7}	-1.763
	3.33×10^{-8}	1×10^{-8}	1×10^{-8}	1×10^{-7}	-1.763
	3.33×10^{-8}	1×10^{-8}	1×10^{-8}	1×10^{-8}	-1.829
	1×10^{-8}	1×10^{-6}	1×10^{-6}	1×10^{-7}	-1.831
	1×10^{-8}	1×10^{-6}	1×10^{-6}	1×10^{-8}	-1.910
CCSD(T)					-2.230

Table 5.8: The conventional CCSD(T) and DLPNO-CCSD(T) (with default and optimal thresholds) interaction energies (in kcal/mol) and timings (in hours) for the lowest-energy N-PHAC-CO_2 structures in the aDZ basis. DLPNO-CCSD(T) refers to the default-threshold results, while DLPNO-CCSD(T)$_{Tight}$ denotes data for the optimal thresholds.

Complex	CCSD(T)	time	DLPNO-CCSD(T)	time	DLPNO-CCSD(T)$_{Tight}$	time
in-plane pyrazine-CO_2	-3.25	0.14	-3.12	0.8	-3.08	0.6
stacked pyrazine-CO_2	-0.74	0.13	-0.73	0.9	-0.54	0.6
in-plane pyridine-CO_2	-3.64	0.16	-3.46	0.2	-3.46	0.6
stacked pyridine-CO_2	-1.63	0.48	-1.74	0.2	-1.44	4.1
in-plane quinoxaline-CO_2	-3.72	4.9	-4.63	0.8	-3.50	3.4
stacked quinoxaline-CO_2	-2.19	3.6	-2.46	0.9	-1.86	3.7
in-plane quinoline-CO_2	-3.99	5.1	-3.71	2.2	-3.77	11.4
stacked quinoline-CO_2	-2.12	15.5	-2.30	0.9	-1.86	3.8
in-plane pyrido[3,2-g]quinoline-CO2	-3.85	26.6	-3.65	2.6	-3.62	14.2
stacked pyrido[3,2-g]quinoline-CO2	-2.29	23.0	-2.33	2.9	-1.96	14.0
in-plane phenanthroline-CO_2	-4.71	5.9	-4.70	2.4	-4.43	36.0
stacked phenanthroline-CO_2	-3.30	22.6	-3.15	2.6	-2.99	14.0
in-plane 2-azapyrene-CO_2	-3.81	12.8	-3.53	4.2	-3.64	22.0
stacked 2-azapyrene-CO_2	-2.23	39.9	-2.78	4.3	-1.83	44.0
in-plane 1,6-diazacoronene-CO_2	-4.01	96.6	-3.17	30.0	-3.35	76.0
stacked 1,6-diazacoronene-CO_2	-2.71	85.8	-2.55	35.0	-1.80	93.0

Changing this new threshold gives more accurate results but the improvement is not tremendous, as suggested in the ORCA manual, the previous three are the dominant thresholds. Based on the information in this table, we obtained an optimal combination of the four thresholds. The values are 3.33×10^{-8}, 1×10^{-6}, 1×10^{-6}, and 1×10^{-8} for T_{CutPNO}, $T_{CutPairs}$, T_{CutMKN}, and T_{CutTNO}, respectively.

Now, we collect all the default-threshold results, conventional CCSD(T), and optimal-threshold DLPNO-CCSD(T), as well as the timings in hours for these calculations in Table 5.8. From this table, it is obvious that the optimal-threshold results underestimate the interaction energies, however, the good news for them is that they are consistent. Different from the optimal-threshold values, the default-threshold ones are fluctuating. In the case of 1,6-diazacoronene-CO_2, regardless of the in-plane or stacked structure, neither the default nor the optimal thresholds provide an accurate description of the interaction. Another thing is that the running times of default-threshold calculations exhibit a truly near linear behavior, however, the running times for optimal thresholds are close to conventional CCSD(T). These factors make the DLPNO-CCSD(T) approach troublesome in terms of describing the

dispersion interaction between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen.

Even though we already concluded that the local methods could not generate accurate enough interaction energies for our target systems, we still want to try some methods to possibly obtain CCSD(T)-level silver standard benchmark results. That is the reason we did one further test by using the FNO-CCSD(T) method. There is one threshold to play with when one is running FNO-CCSD(T) in PSI4 – ***OCC_TOLERANCE***. This is a "*Cutoff for occupation of MP2 virtual NOs in FNO-QCISD/CCSD(T). Virtual NOs with occupations less than* ***OCC_TOLERANCE*** *will be discarded.*" The default value for ***OCC_TOLERANCE*** is 1.0×10^{-6}. When the default threshold is used, the differences between the FNO-CCSD(T) and CCSD(T) interaction energies are less than 0.02 kcal/mol, which is excellent. The results for FNO-CCSD(T) are collected in Table 5.9. While we are computing the interaction energies with ***OCC_TOLERANCE*** of 1×10^{-4}, the results are still acceptable. However, when 1×10^{-3} is used to obtain the results, the errors are out of control. Therefore, it is safe to conclude that FNO-CCSD(T) with the threshold ***OCC_TOLERANCE*** $= 1\times10^{-4}$ could produce fairly accurate results for the interaction between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen.

5.4 Summary

In this project of assessment of the performance of local methods for the description of the interaction between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen, there are three main conclusions. Firstly, LCCSD(T) and DF-LCCSD(T) are too expensive to obtain interaction energies for the complexes of interest, as can be seen from the LCCSD(T) and DF-LCCSD(T) results obtained from 1- and 2-ring systems in Tables 5.1 and 5.2. Secondly, the default-threshold DLPNO-CCSD(T) results are inconsistent across different dimers, especially when the larger systems are studied. Although it is claimed in the ORCA documentation that these thresholds should not be touched at all, it is obviously not

Table 5.9: The conventional CCSD(T) and FNO-CCSD(T) interaction energies (in kcal/mol) for the lowest-energy N-PHAC-CO_2 structures, in the aDZ basis. The FNO threshold is given in parentheses.

dimers	CCSD(T)	FNO-CCSD(T)
in-plane pyrazine-CO_2	-3.245	-3.247(1×10^{-6})
		-3.266(1×10^{-5})
		-3.258(1×10^{-4})
		-3.342(1×10^{-3})
stacked pyrazine-CO_2	-0.735	-0.738(1×10^{-6})
		-0.761(1×10^{-4})
		-0.885(1×10^{-3})
in-plane pyridine-CO_2	-3.636	-3.638(1×10^{-6})
		-3.665(1×10^{-4})
		-3.739(1×10^{-3})
stacked pyridine-CO_2	-1.630	-1.633(1×10^{-6})
		-1.656(1×10^{-4})
		-2.202(1×10^{-3})
in-plane quinoxaline-CO_2	-3.723	-3.726(1×10^{-6})
		-3.753(1×10^{-4})
		-3.838(1×10^{-3})
stacked quinoxaline-CO_2	-2.192	-2.196(1×10^{-6})
		-2.266(1×10^{-4})
		-2.942(1×10^{-3})
in-plane quinoline-CO_2	-3.994	-4.024(1×10^{-4})
		-4.097(1×10^{-3})
stacked quinoline-CO_2	-2.121	-2.159(1×10^{-4})
		-2.349(1×10^{-3})
in-plane pyrido[3,2-g]quinoline-CO_2	-3.852	-3.878(1×10^{-4})
		-3.952(1×10^{-3})
stacked pyrido[3,2-g]quinoline-CO_2	-2.288	-2.356(1×10^{-4})
		-2.565(1×10^{-3})
in-plane phenanthroline-CO_2	-4.706	-4.723(1×10^{-6})
		-4.782(1×10^{-4})
		-5.241(1×10^{-3})
stacked phenanthroline-CO_2	-3.298	-3.390(1×10^{-4})
		-3.892(1×10^{-3})
in-plane 2-azapyrene-CO_2	-3.811	-3.820(1×10^{-5})
		-3.825(1×10^{-4})
		-3.914(1×10^{-3})
		-3.819(1×10^{-4})(laDZ)
stacked 2-azapyrene-CO_2	-2.230	-2.235(1×10^{-6})
		-3.216(1×10^{-3})
		-2.073(1×10^{-4})(laDZ)
in-plane 1,6-diazacoronene-CO_2 (laDZ)	-4.009	-4.118(1×10^{-3})
stacked 1,6-diazacoronene-CO_2 (laDZ)	-2.710	-3.524(1×10^{-3})

the case as demonstrated in our research. On the other hand, it is very hard to manipulate these thresholds in order to yield decent benchmark values. The reason that we cannot get accurate interaction energies is that N-PHACs are highly nonlocal. Finally, the FNO-CCSD(T) results are the best out of all the test methods. One can even loosen the default threshold *OCC_TOLERANCE* to 1×10^{-4}.

As discussed very recently in Ref. 160, the new version of ORCA will be released in the near future. The truly linear version of DLPNO-CCSD(T)[160] will be included in this version and it is claimed to have improved accuracy in addition to improved efficiency. Our hope is that this new implementation could bring promising outcomes. Other than the truly linear DLPNO-CCSD(T), CIM-CCSD(T)[151] and methods of increments[153,161] might be tried as well.

Bibliography

[1] Everett, D. *Pure Appl. Chem.* **1972**, *31*, 577–638.

[2] Britz, D. A.; Khlobystov, A. N. *Chem. Soc. Rev.* **2006**, *35*, 637–659.

[3] Bruch, L. W.; Diehl, R. D.; Venables, J. A. *Rev. Mod. Phys.* **2007**, *79*, 1381–1454.

[4] Murdachaew, G.; de Gironcoli, S.; Scoles, G. *J. Phys. Chem. A* **2008**, *112*, 9993–10005.

[5] Burns, L. A.; Vazquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *134*, 084107.

[6] Jensen, F. *Introduction to computational Chemistry*; Wiley: New York, 1999.

[7] Levine, I. N. *Quantum Chemistry, Fifth Ed.*; Prentice-Hall, Inc.; New Jersey, 2000.

[8] Leach, A. R. *Molecular Modelling. Principles and Applications, 2nd ed.*; Prentice-Hall: Harlow, England, 2001.

[9] Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; Wavefunction Inc: Irvine, 2003.

[10] Cramer, C. J. *Essentials of Computational Chemistry*; Chichester: John Wiley Sons, Inc.; Hoboken, N. J., 2002.

[11] Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods, 2nd Ed.*; Gaussian Inc. : Pittsburgh, 1996.

[12] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; McGraw-Hill:New York, 1989.

[13] Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[14] Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697–1708.

[15] Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford Univ. Press: Oxford, 1989.

[16] Lawley, K. P. E. *Adv. Chem. Phys.: Ab Initio Methods in Quantum Chemistry-II*; Wiley: New York, 1987.

[17] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[18] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[19] Kohn, W.; Becke, A. D.; Parr, R. G. *J. Chem. Phys.* **1996**, *100*, 12974–12980.

[20] Becke, A. D. *Phys. Rev. A* **1986**, *33*, 2786–2788.

[21] Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612–5626.

[22] Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000.

[23] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[24] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[25] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[26] Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

[27] Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.

[28] Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289–320.

[29] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[30] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007**, *127*, 124108.

[31] Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[32] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

[33] Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. *Phys. Rev. Lett.* **2009**, *103*, 263201.

[34] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[35] Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

[36] Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

[37] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

[38] Risthaus, T.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 1580–1591.

[39] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289–300.

[40] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.

[41] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.

[42] DiStasio, R. A.; von Lilienfeld, O. A.; Tkatchenko, A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 14791–14795.

[43] Dobson, J. F.; Gould, T. *J. Phys.: Condens. Matter* **2012**, *24*, 073201.

[44] Tkatchenko, A.; DiStasio, Jr., R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.

[45] Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A.; Tkatchenko, A. *J. Chem. Phys.* **2014**, *140*, 18A508.

[46] Dubecký, M.; Jurečka, P.; Derian, R.; Hobza, P.; Otyepka, M.; Mitas, L. *J. Chem. Theory Comput.* **2013**, *9*, 4287–4292.

[47] Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. *Rev. Mod. Phys.* **2001**, *73*, 33–83.

[48] Austin, B. M.; Zubarev, D. Y.; Lester Jr., W. A. *Chem. Rev.* **2012**, *112*, 263–288.

[49] Czerw, R.; Terrones, M.; Charlier, J.-C.; Blase, X.; Foley, B.; Kamalakaran, R.; Grobert, N.; Terrones, H.; Tekleab, D.; Ajayan, P. M.; Blau, W.; Rühle, M.; Carroll, D. L. *Nano Lett.* **2001**, *1*, 457–460.

[50] Kang, H. S.; Jeong, S. *Phys. Rev. B* **2004**, *70*, 233411.

[51] Sun, C.; Wang, H.; Hayashi, M.; Chen, L.; Chen, K. *J. Am. Chem. Soc.* **2006**, *128*, 8368–8369.

[52] Usachov, D.; Vilkov, O.; Grüneis, A.; Haberer, D.; Fedorov, A.; Adamchuk, V. K.; Preobrajenski, A. B.; Dudin, P.; Barinov, A.; Oehzelt, M.; Laubschat, C.; Vyalikh, D. V. *Nano Lett.* **2011**, *11*, 5401–5407.

[53] Cho, Y. J.; Kim, H. S.; Baik, S. Y.; Myung, Y.; Jung, C. S.; Kim, C. H.; Park, J.; Kang, H. S. *J. Phys. Chem. C* **2011**, *115*, 3737–3744.

[54] Luo, Z.; Lim, S.; Tian, Z.; Shang, J.; Lai, L.; MacDonald, B.; Fu, C.; Shen, Z.; Yu, T.; Lin, J. *J. Mater. Chem.* **2011**, *21*, 8038–8044.

[55] Robertson, J.; Davis, C. A. *Diamond Relat. Mater.* **1995**, *4*, 441–444.

[56] Jiang, D.; Cooper, V. R.; Dai, S. *Nano Lett.* **2009**, *9*, 4019–4024.

[57] Du, H.; Li, J.; Zhang, J.; Su, G.; Li, X.; Zhao, Y. *J. Phys. Chem. C* **2011**, *115*, 23261–23266.

[58] Bucior, B. J.; Chen, D.-L.; Liu, J.; Johnson, J. K. *J. Phys. Chem. C* **2012**, *116*, 25904–25910.

[59] Hauser, A. W.; Schwerdtfeger, P. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13292–13298.

[60] Hauser, A. W.; Schwerdtfeger, P. *J. Phys. Chem. Lett.* **2012**, *3*, 209–213.

[61] Blankenburg, S.; Bieri, M.; Fasel, R.; Mllen, K.; Pignedoli, C. A.; Passerone, D. *Small* **2010**, *6*, 2266–2271.

[62] Schrier, J. *ACS Appl. Mater. Interfaces* **2011**, *3*, 4451–4458.

[63] Schrier, J. *ACS Appl. Mater. Interfaces* **2012**, *4*, 3745–3752.

[64] Liu, H.; Cooper, V. R.; Dai, S.; Jiang, D. *J. Phys. Chem. Lett.* **2012**, *3*, 3343–3347.

[65] Lu, R.; Rao, D.; Lu, Z.; Qian, J.; Li, F.; Wu, H.; Wang, Y.; Xiao, C.; Deng, K.; Kan, E.; Deng, W. *J. Phys. Chem. C* **2012**, *116*, 21291–21296.

[66] Lu, R.; Meng, Z.; Kan, E.; Li, F.; Rao, D.; Lu, Z.; Qian, J.; Xiao, C.; Wu, H.; Deng, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 666–670.

[67] Brockway, A. M.; Schrier, J. *J. Phys. Chem. C* **2013**, *117*, 393–402.

[68] Lu, R.; Meng, Z.; Rao, D.; Wang, Y.; Shi, Q.; Zhang, Y.; Kan, E.; Xiao, C.; Deng, K. *Nanoscale* **2014**, *6*, 9960–9964.

[69] Kauffman, D. R.; Star, A. *Angew. Chem. Int. Ed.* **2008**, *47*, 6550–6570.

[70] Cao, D.; Zhang, X.; Chen, J.; Wang, W.; Yun, J. *J. Phys. Chem. B* **2003**, *107*, 13286–13292.

[71] Hirsch, A. *Angew. Chem. Int. Ed.* **2002**, *41*, 1853–1859.

[72] Herm, Z. R.; Swisher, J. A.; Smit, B.; Krishna, R.; Long, J. R. *J. Am. Chem. Soc.* **2011**, *133*, 5664–5667.

[73] Kowalczyk, P. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2784–2790.

[74] Smith, D. G. A.; Patkowski, K. *J. Chem. Theory Comput.* **2013**, *9*, 370–389.

[75] Smith, D. G. A.; Patkowski, K. *J. Phys. Chem. C* **2014**, *118*, 544–550.

[76] Smith, D. G. A.; Patkowski, K. *J. Phys. Chem. C* **2015**, *119*, 4934–4948.

[77] Vogiatzis, K. D.; Mavrandonakis, A.; Klopper, W.; Froudakis, G. E. *ChemPhysChem* **2009**, *10*, 374–383.

[78] Mackie, I. D.; DiLabio, G. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2780–2787.

[79] Schütz, M.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358.

[80] Voloshina, E.; Usvyat, D.; Schütz, M.; Dedkov, Y.; Paulus, B. *Phys. Chem. Chem. Phys.* **2011**, *13*, 12041–12047.

[81] Schütz, M.; Yang, J.; Chan, G. K.; Manby, F. R.; Werner, H.-J. *J. Chem. Phys.* **2013**, *138*, 054109.

[82] Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013**, *139*, 134101.

[83] Werner, H.-J. et al. MOLPRO, version 2012.1, a package of ab initio programs. 2012; see http://www.molpro.net (accessed June 2, 2014).

[84] Werner, H.-J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149–8160.

[85] Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[86] Hättig, C. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

[87] Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

[88] Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.

[89] van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873–1885.

[90] Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200–10207.

[91] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

[92] Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2009**, *131*, 194105.

[93] Patkowski, K. *J. Chem. Phys.* **2012**, *137*, 034103.

[94] Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.

[95] Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.

[96] Köhn, A. *J. Chem. Phys.* **2009**, *130*, 131101.

[97] Marchetti, O.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3400–3409.

[98] Tew, D. P.; Klopper, W.; Hättig, C. *Chem. Phys. Lett.* **2008**, *452*, 326–332.

[99] Marchetti, O.; Werner, H.-J. *J. Phys. Chem. A* **2009**, *113*, 11580–11585.

[100] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 5121–5129.

[101] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[102] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[103] Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029–5036.

[104] Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.

[105] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[106] Johnson, E. R.; Becke, A. D.; Sherrill, C. D.; DiLabio, G. A. *J. Chem. Phys.* **2009**, *131*, 034111.

[107] Wheeler, S. E.; Houk, K. N. *J. Chem. Theory Comput.* **2010**, *6*, 395–404.

[108] Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[109] Turney, J. M. et al. *WIREs Comput Mol Sci* **2012**, *2*, 556–565.

[110] Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.

[111] Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[112] Smith, D. G. A.; Jankowski, P.; Slawik, M.; Witek, H. A.; Patkowski, K. *J. Chem. Theory Comput.* **2014**, *10*, 3140–3150.

[113] Demovičová, L.; Hobza, P.; Řezáč, J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19115–19121.

[114] Pitoňak, M.; Neogrády, P.; Řezáč, J.; Jurečka, P.; Urban, M.; Hobza, P. *J. Chem. Theory Comput.* **2008**, *4*, 1829–1834.

[115] Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1998**, *108*, 9221–9226.

[116] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[117] Li, S.; Smith, D. G. A.; Patkowski, K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 16560–16574.

[118] Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2014**, *140*, 18A527.

[119] Steinmann, S. N.; Wodrich, M. D.; Corminboeuf, C. *Theor. Chem. Acc.* **2010**, *127*, 429–442.

[120] Seth, M.; Ziegler, T.; Steinmetz, M.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 2286–2299.

[121] Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. *Chem. Phys. Lett.* **1997**, *275*, 151–160.

[122] Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425–8433.

[123] Gerber, I. C.; Ángyán, J. G. *Chem. Phys. Lett.* **2005**, *415*, 100–105.

[124] Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726–3741.

[125] Gráfová, L.; Pitoňák, M.; Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 2365–2376.

[126] Gobre, V. V.; Tkatchenko, A. *Nature Comm.* **2013**, *4*, 2341.

[127] Misquitta, A. J.; Spencer, J.; Stone, A. J.; Alavi, A. *Phys. Rev. B* **2010**, *82*, 075312.

[128] Dobson, J. F. *Surf. Sci.* **2011**, *605*, 1621–1632.

[129] Geim, A. K.; Grigorieva, I. V. *Nature* **2013**, *499*, 419–425.

[130] Lechner, C.; Sax, A. F. *J. Phys. Chem. C* **2014**, *118*, 20970–20981.

[131] Nakashima, N. *Sci. Technol. Adv. Mater.* **2006**, *7*, 609–616.

[132] Maurer, R. J.; Sax, A. F. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9893–9899.

[133] Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

[134] Smith, D. G. A. *Exploring intermolecular interactions through Coupled Cluster, Density Functional, and Multireference Symmetry-Adapted Perturbation Theories*. Ph.D. Dissertation, Auburn University, 2015.

[135] Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *J. Phys. Chem. Lett.* **2016**, *7*, 2197–2203.

[136] Koskinen, P.; Mkinen, V. *Comp. Mat. Sci.* **2009**, *47*, 237–253.

[137] Tsuzuki, S.; Honda, K.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104–112.

[138] Marom, N.; DiStasio, R. A.; Atalla, V.; Levchenko, S.; Reilly, A. M.; Chelikowsky, J. R.; Leiserowitz, L.; Tkatchenko, A. *Angew. Chem. Int. Ed.* **2013**, *52*, 6629–6632.

[139] Ambrosetti, A.; Ferri, N.; DiStasio, R. A.; Tkatchenko, A. *Science* **2016**, *351*, 1171–1176.

[140] Dubecký, M.; Mitas, L.; Jurečka, P. *Chem. Rev.* **2016**, *116*, 5188–5215.

[141] Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. *J. Chem. Phys.* **2013**, *139*, 094105.

[142] Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[143] Pulay, P.; Saebo, S. *Theoretica chimica acta* **1986**, *69*, 357–368.

[144] Werner, H.-J. *Mol. Phys.* **1996**, *89*, 645–661.

[145] Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

[146] Kats, D.; Korona, T.; Schütz, M. *J. Chem. Phys.* **2006**, *125*, 104106.

[147] Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, 114108.

[148] Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.

[149] Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K. L. *J. Chem. Phys.* **2011**, *134*, 044123.

[150] Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2005**, *123*, 134103.

[151] Li, W.; Piecuch, P.; Gour, J. R.; Li, S. *J. Chem. Phys.* **2009**, *131*, 114109.

[152] Li, W.; Li, S. *J. Chem. Phys.* **2004**, *121*, 6649–6657.

[153] Stoll, H. *Chem. Phys. Lett.* **1992**, *191*, 548–552.

[154] Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[155] Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2012**, *136*, 144105.

[156] Masur, O.; Usvyat, D.; Schütz, M. *J. Chem. Phys.* **2013**, *139*, 164116.

[157] Werner, H.-J.; Schütz, M. *J. Chem. Phys.* **2011**, *135*, 144116.

[158] DePrince III, A. E.; Sherrill, C. D. *J. Chem. Theory Comput.* **2013**, *9*, 2687–2696.

[159] Neese, F. *Wiley Interdiscip. Reviews: Comput. Mol. Sci.* **2012**, *2*, 73–78.

[160] Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016**, *144*, 024109.

[161] Lei, S.; Paulus, B. *Z. Phys. Chem.* **2016**, *230*, 651–666.