Theoretical investigations of weak interactions in fluorine containing compounds

inaugural-Dissertation
to obtain the academic degree
Doctor Rerum Naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by Matthias Berg
from Berlin
2017

This work was prepared under supervision of
Prof. Dr. Beate Paulus (Freie Universität Berlin)
from May 2012 until January of 2017.

1. Referee: Prof. Dr. Beate Paulus
2. Referee: PD Dr. Burkhard Schmidt

Date of the defence: 14.02.2017

Abstract

Subject of this thesis is the theoretical investigation of molecular systems, whose structure is governed by weak noncovalent interactions. Two studies are performed, where compounds of the element fluorine play a central role.

The first study regards the H_2-X_2 ($X = F, Cl, Br$) van der Waals dimers. In it, the question about the influence of quantum effects on the noncovalent bond between H_2 and X_2 molecules is pursued, where both molecules are approximated as rigid rotors. The four dimensional pair potential is calculated by quantum chemical methods and selectively characterized in terms of fundamental interaction components. The adiabatization of the potentials with respect to the rotational motion of H_2 allows the resolution of the differences in the interaction with respect to the nuclar spin isomers of dihydrogen, namely para- (pH_2) and orthohydrogen (oH_2). The two-dimensional rotationally adiabatic p/oH_2-X_2 pair potentials are compared the potentials of the Rg-X_2 (Rg = He, Ne, Ar) dimers. Finally, bound states of the pH_2-X_2 dimers are calculated, in order to determine the zero point energy. Quantum effects due to the nuclear motion dominate the binding energies of the pH_2-X_2 dimers. For pH_2-F_2 their relative contribution is with 85% largest. In the ground state, linear structures are predicted for pH_2-Cl_2 and pH_2-Br_2. In the case of pH_2-F_2 the calculations also yield a linear structure. However, due to the small difference between the binding energies of the linear and the T-shaped structure, no clear prediction can be made in view of the applied approximations.

The second study follows the question, whether folded (hairpin) conformers of n-octane molecules are more strongly stabilized by chain elongation on both ends with either alkyl groups (-C_mH_{2m+1}), perfluoroalkyl groups (-C_mF_{2m+1}) or one alkyl group and one perfluoroalkyl group. For that purpose the energy of the hairpin conformer is determined in relation to the energy of the respective linear conformer. The chain length (n), from which onwards the hairpin conformer is lower in energy than the linear conformer, is directly related to weak attractive intramolecular interactions. And thus to the question how intramolecular alkyl-alkyl, perfluoroalkyl-perfluoroalkyl and perfluoroalkyl-alkyl interactions compare in strength. To answer these questions, accurate quantum chemical calculations are performed. The method of increments at the local Møller-Plesset second order perturbation theory level and the local coupled cluster level is applied for the calculation of the large systems. Using the data from this method, maps of the intramolecular correlation interactions are analysed. Thermodynamic corrections are calculated, in order to make predictions at finite temperatures. The present results support the following ordering of interactions in terms of hairpin stabilization for groups of the same length: alkyl-alkyl > perfluoroalkyl-alkyl $\gtrsim$ perfluoroalkyl-perfluoroalkyl, where it has to be noted, that the differences are very small. At a temperature of 100 K hairpin formation of n-alkanes is predicted for $n \geq 18$. Hairpin formation of the investigated perfluoroalkylalkanes and 1,8-diperfluoroalkyloctanes is predicted for $n \geq 20$ and $n \geq 22$.

Kurzzusammenfassung

Gegenstand dieser Arbeit ist die theoretische Untersuchung molekularer Systeme, deren Struktur maßgeblich durch schwache nichtkovalente Wechselwirkungen bedingt ist. Dies geschieht anhand von zwei Studien, bei denen Verbindungen des Elements Fluor eine zentrale Rolle spielen.

In der ersten Studie werden die van-der-Waals Dimere des Typs H_2-X_2 ($X = F, Cl, Br$) betrachtet. Dabei wird der Fragestellung nachgegangen, wie sich unterschiedliche Quanteneffekte auf die nichtkovalente Bindung zwischen, als starren Rotatoren betrachteten, H_2 und X_2 Molekülen auswirken. Die vierdimensionalen Paarpotenziale werden quantenchemisch berechnet und punktuell über fundamentale Wechselwirkungen charakterisiert. Die Potentizale werden bezüglich der Rotationsbewegung des H_2 Moleküls adiabatisiert, was die Auflösung der intermolekularen Wechselwirkung nach Para- (pH_2) und Orthowasserstoff (oH_2) ermöglicht. Die zweidimensionalen rotationsadiabatisierten p/oH_2-X_2 Paarpotenziale werden mit den Potenzialen der Rg-X_2 (Rg = He, Ne, Ar) Dimere verglichen. Abschließend werden die gebundenen Zustände der pH_2-X_2 Dimere berechnet, um den Beitrag der Nullpunktsenergie zu bestimmen. Quanteneffekte der Kernbewegung dominieren die nichtkovalente Bindung der pH_2-X_2 Dimere. Ihr relativer Beitrag ist mit 85% für pH_2-F_2 besonders groß. Für die Grundzustände von pH_2-Cl_2 und pH_2-Br_2 können mit relativer Sicherheit lineare Strukturen vorhergesagt werden. Im Falle des pH_2-F_2 Dimers ergeben die Rechnungen zwar ebenfalls eine lineare Struktur, der geringe Unterschied zur Bindungsenergie der T-förmigen Struktur lässt in Anbetracht der verwendeten Näherungen aber keine eindeutige Vorhersage zu.

Die zweite Studie geht der Frage nach, ob sich gefaltete (hairpin) n-Octan Moleküle besser über eine beidseitige Kettenverlängerung mit Alkylgruppen (-C_mH_{2m+1}), Perfluoroalkylgruppen (-C_mF_{2m+1}), oder jeweils einer Alkylgruppe und einer Perfluoroalkylgruppe stabilisieren lassen. Dazu wird die Energie des jeweiligen hairpin Konformers relativ zu der Energie des linearen Konformers bestimmt. Die Kettenlänge (n), ab der das hairpin Konformer eine niedrigere Energie aufweist als das lineare Konformer, steht in direktem Zusammenhang mit stabilisierenden intramolekularen van-der-Waals-Wechselwirkungen. Und damit auch mit der Frage wie intramolekulare Alkyl-Alkyl, Perfluoroalkyl-Perfluoroalkyl und Alkyl-Perfluoroalkyl Wechselwirkungen der Stärke nach geordnet sind. Zur Beantwortung dieser Fragen werden genaue quantenchemische Rechnungen durchgeführt. Hierbei wird unter anderem die Inkrementenmethode auf Niveau der lokalen Møller-Plesset Störungstheorie zweiter Ordnung und der lokalen Coupled-Cluster-Theorie angewendet, um die genaue Berechnung der großen Systeme zu ermöglichen. Diese erlaubt es auch, die das hairpin Konformer stabilisierenden Wechselwirkungen innerhalb seiner Struktur zu kartieren. Um Vorhersagen bei endlicher Temperatur zu treffen, werden thermodynamische Korrekturen berechnet. Für die betrachteten Kettenmoleküle findet sich folgende Ordnung der intramolekularen Wechselwirkungen bei gleicher Gruppenlänge: Alkyl-Alkyl > Alkyl-Perfluoroalkyl $\gtrsim$ Perfluoroalkyl-Perfluoroalkyl, wobei die Unterschiede sehr klein sind. Bei einer Temperatur von 100 K ergeben sich folgende Kettenlängen ab der die Faltung bevorzugt ist: $n \geq 18$ für n-Alkane, $n \geq 20$ für Perfluoroalkylalkane und $n \geq 22$ für 1,8-Diperfluoroalkyloctane.

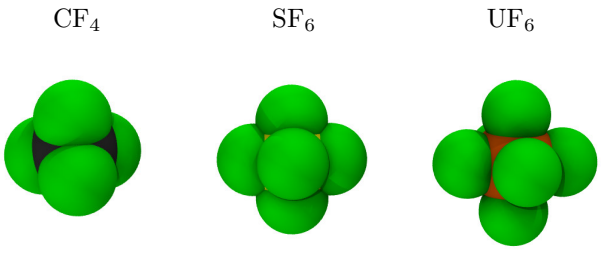
Contents

1 Introduction

The forces between any two or more molecules arise due to intermolecular interactions, also known as noncovalent interactions or sometimes called van der Waals interactions. These microscopic interactions also give rise to some of the macroscopic properties of molecular solids, liquids and gases. Intermolecular interactions can be partitioned into the following fundamental parts: exchange repulsion-, electrostatic-, induction- and dispersion interactions [1–3]. Another powerful concept is the identification of certain strong and directional interaction patterns, such as hydrogen bonds [4] or halogen bonds (σ-hole bonds) [5]. Closely related to the intermolecular interactions are the weak intramolecular interactions, which are important for the stability of conformations of macro molecules such as proteins and DNA. While intermolecular interactions are sometimes viewed separately from chemical bonds (strong intramolecular interactions), both concepts arise ultimately from the correlated quantum mechanical motion of electrons and nuclei. This also explains, why they cannot be clearly separated sometimes - the borders between both concepts are blurred.

The detailed study of intermolecular interactions is carried out the by the isolated examination of small van der Waals dimers, timers and larger clusters. Experimentally, this is possible by the combination of molecular beam and high-resolution spectroscopy methods [3, 6]. In this way empirical intermolecular interaction potentials are constructed which reproduce the observed spectra. Highly accurate *ab inito* studies of van der Waals complexes are possible, but restricted to small systems [7]. Limiting factors are the computationally demanding quantum chemical treatment of electron correlation and the fact that numerically accurate quantum mechanical treatments of the nuclear motion are limited to a few degrees of freedom. It is indeed the combination of experiment and theory that has led to the present level of understanding of intermolecular interactions. For example: it is now widely accepted, that is vital to account for dispersion interactions in computational modelling [8, 9].

Fluorine (F) is the most reactive of the elements, as it forms compounds with all other elements besides He, Ne and Ar [10]. Many elements reach their highest oxidation number in their fluorine compounds. Examples for such fluorine compounds are CF_4, SF_6 and UF_6. SF_6 has an important application as a dielectric medium in high-voltage circuit breakers. UF_6 is used to enrich uranium for the production of nuclear fuels and nuclear weapons. All three compounds have in common that they are gases at room temperature (CF_4 and SF_6) or slightly above (UF_6), see also figure 1.1a, despite their large molar masses. This property can be directly related to their particularly weak intermolecular interactions. Since all three compounds are apolar, dispersion interactions dominate the attractive part of their interaction potentials. Dispersion interactions can be derived from the molecular polarizabilities. The atomic polarizability of F is similar to that of H and clearly lower compared to those of the heavier halogen atoms Cl, Br and I, which qualitatively explains the weak intermolecular attraction of perfluorinated compounds, see figure 1.1b. The same reasoning helps to explain that F_2 and Cl_2 are gases at room temperature while Br_2 is a liquid and I_2 is a solid.

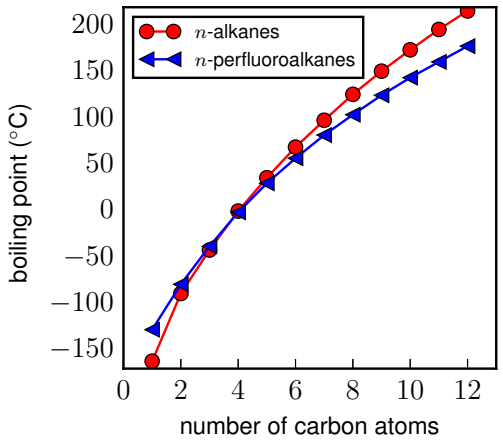


(a) Low boiling or sublimation points of perfluorinated compounds indicate weak intermolecular interactions [10].

$$U_{\text{disp}}^{AB} \sim -\frac{\alpha^A \alpha^B}{R_{AB}^6}$$

X	H	F	Cl	Br	I
α^X (Å^3)	0.67	0.56	2.2	3.1	4.7

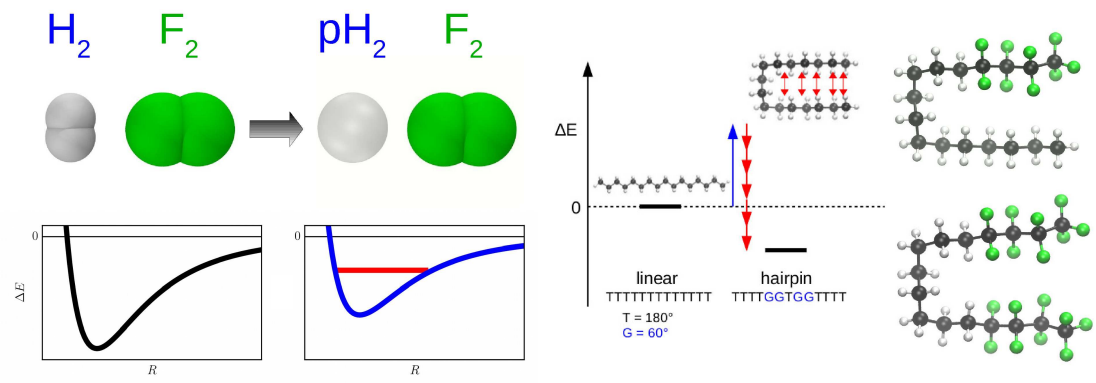
(b) Average dipole polarizabilities for ground state atoms [11, 12].



(c) Boiling points of n-alkanes and perfluorinated n-alkanes [12, 13].

Figure 1.1: The low polarizability of fluorine atoms (b) qualitatively explains the presence of weak dispersion interactions of perfluorinated compounds (a,c).

A particularly interesting question regards the origin of the low miscibility of liquid alkanes and perfluoroalkanes, despite their similar average molecular interactions, as estimated from their similar boiling points [12, 14–16], see figure 1.1c. The thermodynamic properties of liquids and their mixtures are studied by statistical methods such as Monte Carlo and molecular dynamics simulations, which ultimately rely on a potential for the intermolecular interaction [17, 18]. Accurate quantum chemical studies of small alkane and perfluoroalkane dimers are thus performed to provide a benchmark for more efficient methods such as force fields and dispersion corrected density functional theory [18–21]. These can then be used for the latter techniques, but also for other applications such as computational modelling of supramolecular complexes, which contain perfluoroalkyl groups. The tendency of compounds with large perfluoroalkyl-groups to rather mix with perfluoroalkane solvents than with water or lipophilic solvents, is exploited in the recovery of fluorophilic catalysts in liquid-liquid biphasic synthesis, which is an important part of "Green Chemistry" [12, 22]. Molecules with perfluoroalkyl groups aggregate in water, forming layers, micelles and other supramolecular architectures such as micro bubbles [16, 23, 24]. The unique properites of perfluoroalkanes and fluorinated polymers such as chemical inertness, high hydrophobicity, low friction, high small gas molecule carrying capacity,



(a) The interaction potential of H_2 with F_2 (black curve) changes due to the nuclear spin quantum effect (blue curve). The binding energy (red line) of the pH$_2$-F$_2$ dimer is dominated by nuclear quantum effects.

(b) Differences between weak intramolecular alkyl-alkyl, perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions may influence the stability of n-alkane, perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpin conformers with the same number of carbon atoms, with respect to their linear conformers.

Figure 1.2: Sketches for the two studies in this thesis.

lead to a wide range of further applications such as lubricants, medical fluids, liquid breathing, oil and water repellent fabrics, insulation materials, solvents and fire fighting foams [12, 16, 25].

This thesis contains two *ab initio* investigations, which are linked by the role of fluorine with respect to weak intermolecular interactions. Here only a brief introduction to the two studies is given, as each begins with a more detailed individual introduction.

Subject of the first investigation, see chapter 3, are the H_2-X_2 (X = F, Cl, Br) van der Waals dimers. Here, it is not only interesting to compare their intermolecular interaction potentials, but also to determine the individual interaction components, in order to explain the influence of electronic quantum effects on the shape of the potentials. However, the main motivation to study these van der Waals dimers is the importance of nuclear quantum effects, which are anticipated to be large due to the low mass of the H_2 molecule. The method of rotational adiabatization allows the determination of separate interaction potentials for the nuclear spin isomers ortho- (oH$_2$) and para-hydrogen (pH$_2$) [26, 27]. Thus, one of the goals of this study is to determine the changes in the interaction due to the nuclear-spin quantum effect, which is schematically shown in figure 1.2a. The rotationally adiabatic p/oH$_2$-X_2 potentials are then be compared to the Rg-X_2 (Rg = He, Ne, Ar) potentials, which are known for their low anisotropy [28]. Lastly, the bound states of the pH$_2$-X_2 potentials are calculated to determine the zero point energy with respect to the internal motions of the pH$_2$-X_2 dimers, which is indicated in figure 1.2a. The development of pH$_2$-X_2 pair potentials is further of importance, because they can be used in the construction of approximate many body potentials for X_2 impurities in pH$_2$ matrices [27]. The latter systems are used to study the fundamental dihydrogen dihalogen "detonating gas" reaction at cryogenic temperatures [29, 30]. The respective many-body potentials allow the study of interesting matrix quantum effects and allow to determine the initial state of the reaction [27].

In the second investigation, see chapter 4, weak intramolecular interactions between n-alkyl (-C_mH_{2m+1}) and perfluoro-n-alkyl groups (-C_mF_{2m+1}) are studied systematically in terms of their potential to stabilize folded (hairpin) conformers. The alkyl and perfluoroalkyl groups are thereby linked to a central octane unit (-C_8H_{16}-), which provides the necessary flexibility for folding. The stability of hairpin conformers depends on a fine balance between the introduction of four energetically unfavourable gauche carbon backbone dihedral angles and attractive intramolecular interactions, which depend on the overall chain length n, see also the scheme in figure 1.2b. The number of carbon atoms, *i.e.* the chain length, where the energy of the hairpin conformer is lower than the energy of the linear conformer is thus directly linked to the strength of the intramolecular interactions [31]. The present study is motivated by the recent experimental and theoretical investigation of n-alkane hairpin formation in via supersonic jet Raman spectroscopy by Lüttschwager *et al.* [32, 33] and related highly accurate theoretical investigations [34, 35]. This technique may allow to experimentally investigate the folding of the herein theoretically investigated partially fluorinated n-alkanes. Can we learn something about fluorophillic effects at the single molecule level? How do alkyl-alkyl, perfluoroalkyl-perfluoroalkyl and perfluoroalkyl-alkyl interactions compare in terms of hairpin stabilization? What are the hairpin critical chain lengths at low temperatures (100 K)? To answer these questions, we apply a combination of quantum chemical methods ranging from dispersion corrected density functional theory [9, 36–39] to explicitly correlated local electron correlation methods [40, 41] in conjunction with the method of increments [42–46]. The results of this investgation might aid the evaluation of force fields or density functional theory methods.

This thesis is organized as follows, the applied quantum chemistry methods are presented in chapter 2, including a brief recapitulation of the types of intermolecular interactions and their *ab initio* determination in section 2.3. Chapter 3 contains the investigation on the H_2-X_2 van der Waals dimers. The stability of hairpin conformers of partially fluorinated n-alkanes is investigated in chapter 4. Each study is closed by its own summary & conclusions and outlook sections.

2 Theory

2.1 The Born-Oppenheimer approximation

Many properties of molecules can be described by their state $|\Phi\rangle$, which is a solution of the non-relativistic time-independent Schrödinger equation

$$\hat{H}_{\mathrm{mol}}|\Phi\rangle = E_{\mathrm{mol}}|\Phi\rangle \tag{2.1}$$

where $\hat{H}_{\mathrm{mol}}$ is the molecular Hamilton operator and E_{mol} is the total energy of the molecule in state $|\Phi\rangle$. The M nuclei and N electrons of the molecule are described as point charges, at position vectors $\mathbf{R}_A$ and $\mathbf{r}_i$, respectively. The molecular Hamilton operator is set up by writing out the kinetic energy operators, $\hat{T}_A$, $\hat{T}_i$, for each particle and the operators for the Coulomb interactions $\hat{V}_{AB}$, $\hat{V}_{Ai}$, $\hat{V}_{ij}$ of each pair of particles. The Hamiltonian written in atomic units thus reads

$$\hat{H}_{\mathrm{mol}} = \sum_{i=1}^{N} \underbrace{-\frac{1}{2}\nabla_i^2}_{\hat{T}_i} + \sum_{A=1}^{M} \underbrace{-\frac{1}{M_A}\nabla_A^2}_{\hat{T}_A} + \sum_{i=1}^{N}\sum_{A=1}^{M} \underbrace{-\frac{Z_A}{r_{iA}}}_{\hat{V}_{Ai}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \underbrace{\frac{1}{r_{ij}}}_{\hat{V}_{ij}} + \sum_{A=1}^{M}\sum_{B=1}^{M} \underbrace{\frac{Z_A Z_B}{R_{AB}}}_{\hat{V}_{AB}} \tag{2.2}$$

where the charge of nucleus A is given by its atomic number Z_A, the Coulomb interactions further depend on the distances between electrons r_{ij}, nuclei and electrons r_{iA} and nuclei R_{AB}.[47]

The solution of equation 2.1, except for the smallest molecules like $H_2{}^+$ [48], is usually only possible by invoking the Born-Oppenheimer Approximation (BOA), where the motion of electrons is treated as instantaneous from the perspective of the nuclei and electrons move in the static potential of clamped nuclei. Adiabatisation of equation 2.1 according to the BOA gives rise to the electronic Schrödinger equation

$$\hat{H}_{\mathrm{el}}|\Psi\rangle = E|\Psi\rangle \tag{2.3}$$

where the electronic wave function is denoted $|\Psi\rangle$ and the total energy of the electrons in state $|\Psi\rangle$ in the field of the clamped nuclei is given by E. The electronic Hamilton operator $\hat{H}_{\mathrm{el}}$ is $\hat{H}_{\mathrm{mol}}$ (equation 2.2) for clamped nuclei, thus with $\hat{T}_A = 0$ and $\hat{V}_{AB} = \mathrm{const.}$. The electronic wave function $|\Psi\rangle$ depends fully on the electronic coordinates $\mathbf{r}_i$ and parametrically on the nuclear coordinates $\mathbf{R}_A$. Approximate solutions of equation 2.3 that have been used in this thesis are discussed in section 2.2.

As a consequence of the BOA, nuclei move on adiabatic potential energy surfaces (PES) $E(\mathbf{R}_A)$ found by solving equation 2.3 in the space spanned by the nuclear position vectors $\mathbf{R}_A$. The stationary nuclear states $|\Omega\rangle$ are solutions to the time-independent nuclear Schrödinger

equation

$$\hat{H}_{\mathrm{nuc}}|\Omega\rangle = E_{\mathrm{mol}}^{\mathrm{BOA}}|\Omega\rangle \tag{2.4}$$

where the molecular energy in the BOA is given by $E_{\mathrm{mol}}^{\mathrm{BOA}}$, an approximation to E_{mol} from equation 2.1, and $\hat{H}_{\mathrm{nuc}}$ denotes the nuclear Hamilton operator, which is given by

$$\hat{H}_{\mathrm{nuc}} = + \sum_{A=1}^{M} \underbrace{-\frac{1}{M_A}\nabla_A^2}_{\hat{T}_A} + E(\mathbf{R}_A). \tag{2.5}$$

The nuclear Schrödinger equation itself can be adiabatised by splitting the nuclear motion into translation, vibration and rotational degrees of freedom (DOF). Equation 2.4 is important for the development of rotationally adiabatic pair potentials as in chapter 3 and the calculation of thermodynamic corrections.

2.1.1 Limits of Born-Oppenheimer approximations

Whether the BOA or adiabatic separation breaks down or holds, depends on the size of non-adiabatic coupling terms (NACTs), which arise due to the adiabatic separation. In the following, we briefly show, where the NACTs appear and how their importance can be determined. Since, we will apply the BOA not only to electrons and nuclei but also to nuclear degrees of freedom, we consider a general case for light and heavy particles [27, 49].

A total Hamiltonian of a system of total mass m with light particles with mass scaled coordinates $\mathbf{l}$ and heavy particles with mass scaled coordinates $\mathbf{h}$ written in the adiabatic representation is,

$$\hat{H}(\mathbf{l},\mathbf{h}) = \hat{T}_h(\mathbf{h}) + \hat{H}_l(\mathbf{l};\mathbf{h}), \tag{2.6}$$

where $\hat{T}_h(\mathbf{h})$ denotes the kinetic energy operator of the heavy particles and $\hat{H}_h(\mathbf{l};\mathbf{h})$ denotes the light particle Hamiltonian. The latter Hamiltonian depends parametrically on the coordinates of the heavy particles and also contains their interactions. The Schrödinger Equation for the system is given by,

$$\left(\hat{H} - E\right)|\Psi(\mathbf{l},\mathbf{h})\rangle = 0, \tag{2.7}$$

where E is the total energy and $|\Psi(\mathbf{l},\mathbf{h})\rangle$ is the total wave function. In a Hilbert space of dimension N, the total wave-function can be expressed in terms of the Born-Oppenheimer expansion [50, 51],

$$|\Psi(\mathbf{l},\mathbf{h})\rangle = \sum_{j=1}^{N} \psi_j(\mathbf{h})|\phi(\mathbf{l};\mathbf{h})\rangle, \tag{2.8}$$

where the N heavy particle wave functions $\psi_j(\mathbf{h})$ take the role of expansion coefficients for the N light particle wave functions $|\phi_j(\mathbf{l};\mathbf{h})\rangle$. The light particle wave functions are the adiabatic eigenfunctions of the light particle Hamiltonian,

$$(\hat{H}_l(\mathbf{l};\mathbf{h}) - L_j)|\phi_j(\mathbf{l};\mathbf{h})\rangle = 0, \tag{2.9}$$

where L_j are the adiabatic energies of the light particles. By substitution of eq. 2.8 in eq. 2.7 and multiplication from the left by $\langle\phi_j(\mathbf{l};\mathbf{h})|$, differentiating with respect to the heavy particle kinetic energy operator $\hat{T} = -\frac{\hbar^2}{2m}\nabla^2$ and integration over the coordinates of the light particles

leads to as set N of coupled equations [49],

$$-\frac{\hbar^2}{2m}\nabla^2\psi_k + (L_j - E)\,\psi_k - \frac{\hbar^2}{2m}\sum_{j=1}^{N}\left(2\tau_{jk}\cdot\nabla + \tau_{jk}^{(2)}\right)\psi_j = 0. \tag{2.10}$$

The terms τ_{jk} and $\tau_{jk}^{(2)}$ are the so called non-adiabatic coupling terms (NACT). NACTs couple the motions of the light and heavy particles. They depend explicitly on the adiabatic light particle wave functions,

$$\tau_{jk} = \langle\phi_j(\mathbf{l};\mathbf{h})|\nabla\phi_k(\mathbf{l};\mathbf{h})\rangle \tag{2.11}$$

$$\tau_{jk}^{(2)} = \langle\phi_j(\mathbf{l};\mathbf{h})|\nabla^2\phi_k(\mathbf{l};\mathbf{h})\rangle. \tag{2.12}$$

The second order NACTs $\tau_{jk}^{(2)}$ can be derived from the first order term τ_{jk} [49]. Use of the Hellmann-Feynman theorem allows to express the first order NACTs via the following equation [49],

$$\tau_{jk} = \frac{\langle\phi_j(\mathbf{l};\mathbf{h})|\nabla\hat{H}_l|\phi_k(\mathbf{l};\mathbf{h})\rangle}{L_k - L_j}. \tag{2.13}$$

The BO approximation is the neglect of the coupling terms between the motions of the light and heavy particles, *i.e.* the right hand side of eq. 2.10. From eq. 2.11 follows, that the BOA holds, if light particle wave functions couple only weakly due to changes in the heavy particle coordinates. Additionally, NACTs are small when the heavy particles are slow with respect to the motions of the light particles. The herein often applied criterion, that the energy difference of two light particle states has to be large to for the application of the BOA originates from eq. 2.13. Whenever these criteria fail, the BOA breaks down and NACTs have to be in considered in some form, if the adiabatic light particle energies and wave functions should be employed [49].

Let us discuss some examples that are of importance to this work. In general, we only consider the nuclear motion (heavy particles) on the PES given by the adiabatic ground state energies of the electrons (light particles) and neglect the NACTs for the electronic-nuclear motion. This BOA is justified by the energetic separation of the closed shell electronic ground state and the first excited electronic state (in the relevant space of nuclear coordinates). The separation of rotational (heavy) and vibrational (light) degrees of freedom is justified for some cases, but not for others. For example the vibration of Cl_2 proceeds on a similar time scale as the rotation of H_2 [27]. Treating the dihalogen molecules as rigid, thus introduces a small error [27] into the RA potentials in chapter 3. For the same systems, the adiabatic separation of the rotational DOF of H_2 (light particle) from the rotational DOF of dihalogens (heavy particle) is justified by the large differences in their rotational constants. The bound states of the lowest RA PESs for the pH$_2$-X$_2$ dimers are calculated on the basis that the next highest pH$_2$-X$_2$ is energetically well separated from the latter. On the other hand, NACTs are assumed to be important for the first three RA PESs of oH$_2$-X$_2$, due to their small energetic differences. Because this complicates the bound state calculations for oH$_2$-X$_2$ considerably, they are not part of this work.

2.2 The electronic problem

One of the main goals of quantum chemistry is to find approximate solutions to the electronic Schrödinger equation, *i.e.* equation 2.3. Since we are dealing only with molecules in the electronic ground state, we will not cover excited electronic states. Two main theories exist to solve the electronic problem, wave function based methods and density functional theory (DFT). The wave function based methods are Hartree-Fock, described in section 2.2.1 and post-Hartree Fock methods, also called electron correlation methods, described in section 2.2.2. The latter section focuses on the single reference methods Møller-Plesset perturbation theory and coupled cluster theory, their local variants, the method of increments, and the inclusion of explicit correlation. Density functional theory and dispersion corrections in the frame work of DFT are discussed in section 2.2.3. The presentation of the theories mainly follows the textbooks on quantum chemistry by Szabo and Ostlund [47], Jensen [52] and Helgaker, Jørgensen and Olsen [53].

2.2.1 Hartree-Fock theory

Wave function electronic structure methods express many electron wave functions in terms of an orthonormal basis of spin orbitals χ, *i.e.* one electron wave functions. Spin orbitals

$$\chi = \psi(\mathbf{r})\sigma(\omega) \tag{2.14}$$

are products of a spatial orbital $\psi(\mathbf{r})$ and a spin function $\sigma(\omega)$ which depends on the spin coordinate ω. The spin functions σ for fermions, are either α (spin up) or β (spin down) functions. Adding the spin functions becomes necessary in non-relativistic theory in order to correctly describe the antisymmetry of fermionic wave functions with respect to particle exchange, namely

$$|\Psi(\mathbf{x}_i, \mathbf{x}_j)\rangle = -|\Psi(\mathbf{x}_j, \mathbf{x}_i)\rangle \tag{2.15}$$

where $\mathbf{x}_i = \{\mathbf{r}_i, \omega_i\}$.

The simplest anti-symmetrized product ansatz for $|\Psi\rangle$ in terms of orbitals is the Slater determinant

$$|\Psi_{\mathrm{SD}}\rangle = |\chi_i\chi_j\cdots\chi_N\rangle = (N!)^{-1/2}\begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix}. \tag{2.16}$$

In the Hartree-Fock (HF) approximation $|\Psi_{\mathrm{SD}}\rangle$ is used as an ansatz for $|\Psi\rangle$ to solve equation 2.3 for the electronic ground state $|\Psi_0\rangle$ and energy E_0 by application of the variational principle. It can be shown [47], that minimizing E for a Slater determinant by variation of the spin orbitals leads to a set of effective one electron equations

$$\hat{f}(i)\chi(\mathbf{x}_i) = \epsilon\chi(\mathbf{x}_i) \tag{2.17}$$

called Hartree-Fock equations, where $\hat{f}(i)$ denotes the Fock operator and ϵ denotes the orbital energy. The Fock operator has the form

$$\hat{f}(i) = -\frac{1}{2}\hat{\nabla}_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \hat{v}^{\mathrm{HF}}(i) \tag{2.18}$$

where the Hartree-Fock potential $\hat{v}^{\mathrm{HF}}(i)$ describes the interaction of electron i with the average field present due to all other electrons. Its form is

$$\hat{v}^{\mathrm{HF}}(i) = \sum_b \hat{\mathcal{J}}_b(1) - \hat{\mathcal{K}}_b(1) \tag{2.19}$$

where $\hat{\mathcal{J}}_b(1)$ is the Coulomb operator,

$$\hat{\mathcal{J}}_b(1) = \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1} \tag{2.20}$$

and $\hat{\mathcal{K}}_b(1)$ is the exchange operator,

$$\hat{\mathcal{K}}_b(1) = \int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \hat{\mathcal{P}}_{12} \chi_b(2) \tag{2.21}$$

where the operator $\hat{\mathcal{P}}_{12}$ exchanges the coordinates of two electrons. Here and in the subsequent text the short form for the coordinate of electrons, *i.e.* $\chi_b(\mathbf{x}_1) = \chi_b(1)$ is used. The Coulomb operator gives a local potential, *i.e.* at the position of one electron $\mathbf{x}_1$, due to the average field of a second electron in $\chi_b(2)$. The exchange operator reflects the Pauli principle and depends explicitly on the orbital it operates on, *i.e.* it is a nonlocal operator. The orbital dependence of both two electron operators implicates that $\hat{f}(i)$ depends on its own eigenfunctions. The therefore non-linear HF equations 2.17 are solved in an iterative procedure, namely the self consistent field (SCF) method. In practise one considers the dependence on the spin functions before hand and works in a finite basis of spatial orbitals. Here, only restricted HF theory (RHF) for closed shell molecules is considered, in which it is sufficient to work with spatial orbitals.

The molecular orbitals ψ_i (MO) are expressed as a linear combination,

$$\psi_i = \sum_{k=1}^{K} C_{ik}\phi_k, \tag{2.22}$$

of K atom centred basis functions ϕ_k (LCAO), that resemble atomic orbitals, where the MO coefficients are denoted by C_{ik}. This allows to rewrite the intergro-differential equations 2.17 as matrix equations, which are solved by finding the optimal MO coefficients via the SCF procedure. The Hartree-Fock equations in the finite spatial basis are the Roothaan-Hall equations

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.23}$$

where the Fock matrix is denoted by $\mathbf{F}$, the expansion coefficient matrix is $\mathbf{C}$, the overlap matrix is $\mathbf{S}$ and the diagonal orbital energy matrix is ϵ. The overlap matrix $\mathbf{S}$ becomes necessary as

the basis functions ϕ_k are normalized but in general not orthogonal. It's elements are given by

$$S_{kj} = \int d\mathbf{r}_1 \phi_k^*(1)\phi_j(1), \tag{2.24}$$

thus their value and sign depends on the type, orientation and distance of the two basis functions. The matrix $\mathbf{S}$ is Hermitian and has only positive eigenvalues. Linear dependencies in the basis functions lead to eigenvalues in $\mathbf{S}$ approaching zero. The Fock matrix depends on the (closed shell) density matrix $\mathbf{P}$, with elements

$$P_{kj} = 2\sum_a^{N/2} C_{ka}C_{ja}^*, \tag{2.25}$$

which is directly related to the charge density by

$$\begin{aligned}\rho(\mathbf{r}) &= 2\sum_i^{N/2} \psi_i\psi_i \\ &= \sum_{kj} P_{kj}\phi_k(\mathbf{r})\phi_j^*(\mathbf{r}).\end{aligned} \tag{2.26}$$

The matrix elements of the Fock matrix are given by

$$F_{kj} = \int d\mathbf{r}_1 \phi_k^*(1)\hat{f}(1)\phi_j(1) \tag{2.27}$$

$$= \int d\mathbf{r}_1 \phi_k^*(1)\hat{h}(1)\phi_j(1) + \sum_a^{N/2}\int d\mathbf{r}_1 \phi_k^*(1)\left[2\hat{\mathcal{J}}_a(1) - \hat{\mathcal{K}}_a(1)\right]\phi_j(1) \tag{2.28}$$

$$= H_{kj}^{\text{one}} + G_{kj}, \tag{2.29}$$

where the elements H_{kj}^{one} of the one electron Hamiltonian matrix involve one electron integrals and the elements G_{kj} involve two-electron integrals, the resulting matrix $\mathbf{G}$ depends on $\mathbf{P}$. The one electron integrals in $H_{kj}^{\text{core}} = T_{kj} + V_{kj}^{\text{nucl}}$ describe the kinetic energy

$$T_{kj} = \int d\mathbf{r}_1 \phi_k^*(1)\frac{1}{2}\left[-\hat{\nabla}^2\right]\phi_j(1) \tag{2.30}$$

and the attraction by nuclei,

$$V_{kj}^{\text{nucl}} = \sum_{A=1}^{M}\int d\mathbf{r}_1 \phi_k^*(1)\frac{Z_A}{r_{1A}}\phi_j(1). \tag{2.31}$$

The two-electron part is

$$G_{kj} = \sum_{lo} P_{lo}\left[(kj|ol) - \frac{1}{2}(kl|oj)\right] \tag{2.32}$$

where the dependence on $\mathbf{P}$ is shown, the expressions with the round brackets are the two-electron integrals

$$(kj|ol) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_k^*(1)\phi_j(1)r_{12}^{-1}\phi_o^*(2)\phi_l(2). \tag{2.33}$$

To solve the Roothaan equations, equation 2.23, they have to be unitary transformed so that

$$\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{1}. \tag{2.34}$$

This is done by a unitary transformation of the Fock matrix $\mathbf{F}$ represented in the non-orthogonal basis $\{\phi_u\}$,

$$\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X} \tag{2.35}$$

which gives the Fock matrix $\mathbf{F}'$. Subsequent matrix diagonalization of $\mathbf{F}'$ gives the coefficients matrix $\mathbf{C}'$ the orthogonal basis $\{\phi'_k\}$. Hence, we have solved the transformed Roothaan equations

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\epsilon, \tag{2.36}$$

where the coefficient matrix in the non-orthogonal basis is recovered by

$$\mathbf{C} = \mathbf{X}\mathbf{C}'. \tag{2.37}$$

In this way, the equations are solved in the orthogonal basis but the one and two-electron integrals have to be evaluated only once in the non-orthogonal basis.

The SCF procedure involves the following steps:

1. Provide a set of nuclei $\{R_A, Z_A\}$, basis functions $\{\phi_u\}$ and a number of electrons N.
2. Calculate the one and two-electron integrals.
3. Obtain $\mathbf{X}$ by diagonalization of $\mathbf{S}$.
4. Guess a density matrix $\mathbf{P}$.
5. Construct the Fock matrices $\mathbf{F}(\mathbf{P})$ and then $\mathbf{F}'$.
6. Diagonalize $\mathbf{F}'$ to obtain $\mathbf{C}'$ and ϵ.
7. Generate the coefficients in the original basis $\mathbf{C} = \mathbf{X}\mathbf{C}'$.
8. From $\mathbf{C}$, generate the new density matrix $\mathbf{P}$.
9. Calculate the total energy.
10. Check for convergence: a) the elements of $\mathbf{P}$ have not changed beyond a threshold and/or b) the change in total energy is below a threshold.
11. If the calculation is not converged return with $\mathbf{P}$ to step 5 to construct a new Fock matrix.
12. If the calculation is converged, $|\Psi_{\mathrm{SD}}\rangle$ is constructed with the obtained molecular orbitals. Then expectation values can be obtained, and the HF wave function can be used as a starting point for post-Hartree Fock calculations.

Given the electronic wave function of the molecule is well described by a single configuration and a balanced basis set is used, the HF method can be used to predict reasonable equilibrium structures, *i.e.* bond lengths and angles, charge densities and dipole moments. However, all these properties can be described more accurately by calculation of the correlation energy with post-Hartree-Fock methods. [47] For this purpose the HF method is vital, as all of the latter methods need the HF wave function as input.

2.2.2 Electron correlation methods

The Hartree-Fock energy in the complete basis set limit E_0^{HF} is an upper bound to the exact non-relativistic electronic groeund state energy E_0 in equation 2.3. The difference between the two energies

$$E_{\mathrm{corr}} = E_0 - E_0^{\mathrm{HF}} < 0 \tag{2.38}$$

defines the correlation energy E_{corr}. The methods described in this chapter approximate E_{corr} and the exact electronic wave function by improving upon the HF approximation.

Formally, the exact electronic wave function in a given finite basis can be described with the configuration interaction (CI) method. From the solution of Roothaan's equations a closed shell restricted HF determinant $|\Psi_0^{\mathrm{HF}}\rangle$ and a set of 2K spin orbitals $\{\chi_i\}$ are obtained, where $|\Psi_0^{\mathrm{HF}}\rangle$ is described by occupation of the N spin orbitals that are lowest in energy. From the spin orbitals, it is possible to construct $\binom{2K}{N}$ different Slater determinants, *i.e.* configurations, that can be described by the changes, called virtual excitations of electrons, with respect to $|\Psi_0^{\mathrm{HF}}\rangle$. For example, the replacement of the occupied spin orbital χ_a in $|\Psi_0^{\mathrm{HF}}\rangle$ by the unoccupied spin orbital χ_r leads to the singly excited determinant $|\Psi_a^r\rangle$. The ansatz in CI is to expand the electronic wave function $|\Psi_0\rangle$ in the basis of Slater determinants build from the spin orbitals $\{\chi_i\}$ as

$$|\Psi_0\rangle = c_0|\Psi_0^{\mathrm{HF}}\rangle + \left(\frac{1}{1!}\right)^2 \sum_{ar} c_a^r|\Psi_a^r\rangle + \left(\frac{1}{2!}\right)^2 \sum_{abrs} c_{ab}^{rs}|\Psi_{ab}^{rs}\rangle + \left(\frac{1}{3!}\right)^2 \sum_{abcrst} c_{abc}^{rst}|\Psi_{abc}^{rst}\rangle + \cdots \tag{2.39}$$

where the factors in front of the sums prevent multiple counting of excitations leading to the same determinants. The method of linear variation can be applied to solve the electronic SE using this trial function. Thus, one sets up the matrix representation of the electronic Hamiltonian in the basis of Slater-determinants. The lowest eigenvalue of this full CI matrix is an upper bound to the ground state energy and other eigenvalues are approximations to the energies of excited electronic states of the system. Albeit formally exact, the sheer size of the full CI matrix prohibits the application of the method for most many electron systems of interest. Therefore, a huge effort is devoted to the development of methods that approximate the correlation energy, for ever larger molecular systems, but bypass the problem of solving the full CI equations. One obvious choice of approximation to full CI, is to truncate the expansion in equation 2.39 at some level of excited configurations, say up to doubly excited configurations. The corresponding method is then called configuration interaction singles doubles (CISD). Indeed, this approach greatly reduces the scaling, while still giving a reasonable approximation to the correlation energy, but in terms of molecular interactions an error is introduced, because truncated CI expansions are not size extensive. In size extensive methods the energy of a super molecule consisting of a number of non interacting molecules, far away from each other, is exactly the sum of the energies of the individual monomers. HF for example is such a method, but CISD is not. In fact, the absolute correlation energy of CISD of the super molecule is always smaller than the sum of the absolute monomer correlation energies. This problem worsens with the systems size, such that in the limit of an infinitely large system, *i.e.* a crystal, the correlation energy per monomer is zero. The size extensivity error can be reduced by the inclusion of selected quadruple excitations, but is still present. [47]

In the following, the most important size extensive electron correlation methods, especially in terms of the calculation of weak non covalent interactions are presented, namely Møller-Plesset perturbation theory and coupled cluster theory. Then, explicit correlation schemes are discussed, that greatly improve the convergence to the complete basis set limit. Finally, the local correlation methods, including the method of increments are described, which all serve the purpose to treat large molecular systems.

2.2.2.1 Møller-Plesset perturbation theory

The Møller-Plesset perturbation theory (MPPT) is the application of the Rayleigh-Schrödinger perturbation theory (RSPT) variant of the general many-body perturbation theory (MBPT) to the electron correlation problem of N-electron systems, by using the Hartree-Fock Hamiltonian as the zero-order Hamiltonian. In MBPT one uses a set of known eigenfunctions and eigenvalues of a zero-order Hamiltonain $\hat{H}_0$,

$$\hat{H}_0|\Psi_i^{(0)}\rangle = E_i^{(0)}|\Psi_i^{(0)}\rangle \tag{2.40}$$

to approximate the unknown eigenfunctions and eigenvalues of a Hamiltonian $\hat{H}$

$$\hat{H}|\Phi_i\rangle = E_i|\Phi_i\rangle \tag{2.41}$$

that can be written by adding the perturbation $\hat{U}$ to $\hat{H}_0$ as

$$\hat{H} = \hat{H}_0 + \lambda\hat{U} \tag{2.42}$$

where λ controls the strength of the perturbation. This allows to expand E_i and $|\Phi_i\rangle$ as a power-series in λ,

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots \tag{2.43}$$

$$|\Phi_i\rangle = |\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \cdots \tag{2.44}$$

where $E_i^{(n)}$ and $|\Psi_i^{(n)}\rangle$ are the energies and wave function corrections of the nth order. Perturbation theory assumes that the perturbation $\hat{U}$ is small compared to $\hat{H}_0$, and that the power-series expansion of the energy and wave function, equations 2.43 and 2.44, converge. We consider the zero-order wave functions $|\Psi_i^{(0)}\rangle$ to be normalized and choose $|\Phi_i\rangle$ as intermediate normalized, *i.e.* $\langle\Psi_i^{(0)}|\Phi_i\rangle = 1$. The zero-order wave function is orthogonal to the functions of higher orders,

$$\langle\Psi_i^{(0)}|\Psi_i^{(n)}\rangle = 0 \quad n = 1, 2, 3, \cdots . \tag{2.45}$$

The ansatz from equations 2.42, 2.43 and 2.44 is inserted into equation 2.41,

$$\begin{aligned}\left(\hat{H}_0 + \lambda\hat{U}\right) \; &\left(|\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \cdots\right) \\ &= \left(E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots\right) \\ &\left(|\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \cdots\right)\end{aligned} \tag{2.46}$$

which can be separated into one equation per order

$$\hat{H}_0|\Psi_i^{(0)}\rangle = E_i^{(0)}|\Psi_i^{(0)}\rangle \qquad n=0 \tag{2.47}$$
$$\hat{H}_0|\Psi_i^{(1)}\rangle + \hat{U}|\Psi_i^{(0)}\rangle = E_i^{(0)}|\Psi_i^{(1)}\rangle + E_i^{(1)}|\Psi_i^{(0)}\rangle \qquad n=1 \tag{2.48}$$
$$\hat{H}_0|\Psi_i^{(2)}\rangle + \hat{U}|\Psi_i^{(1)}\rangle = E_i^{(0)}|\Psi_i^{(2)}\rangle + E_i^{(1)}|\Psi_i^{(1)}\rangle + E_i^{(2)}|\Psi_i^{(0)}\rangle \qquad n=2 \tag{2.49}$$
$$\vdots \qquad .$$

Multiplication of each equation from the left with $\langle\Psi_i^{(0)}|$ gives expressions for the energies of nth order,

$$E_i^{(0)} = \langle\Psi_i^{(0)}|\hat{H}_0|\Psi_i^{(0)}\rangle \tag{2.50}$$
$$E_i^{(1)} = \langle\Psi_i^{(0)}|\hat{U}|\Psi_i^{(0)}\rangle \tag{2.51}$$
$$E_i^{(2)} = \langle\Psi_i^{(0)}|\hat{U}|\Psi_i^{(1)}\rangle \tag{2.52}$$
$$E_i^{(3)} = \langle\Psi_i^{(0)}|\hat{U}|\Psi_i^{(2)}\rangle \tag{2.53}$$
$$\vdots\ .$$

To solve the first order equations one applies RSPT in that, the first order wave function correction is expressed in the basis of eigenstates of $\hat{H}_0$,

$$|\Psi_i^{(1)}\rangle = \sum_n c_n^{(1)}|\Psi_n^{(0)}\rangle, \tag{2.54}$$

where the coefficients can be projected out by multiplication with $\langle\Psi_n^{(0)}|$,

$$\langle\Psi_n^{(0)}|\Psi_i^{(1)}\rangle = c_n^{(1)}. \tag{2.55}$$

Thus,

$$|\Psi_i^{(1)}\rangle = \sum_{n,n\neq i} |\Psi_n^{(0)}\rangle\langle\Psi_n^{(0)}|\Psi_i^{(1)}\rangle, \tag{2.56}$$

where the sum runs not over $n = i$ since $c_i^{(1)} = 0$. Equation 2.48 is rearranged and $E_i^{(1)}$ substituted according to equation 2.51

$$\left(E_i^{(0)} - \hat{H}_0\right)|\Psi_i^{(1)}\rangle = \left(\hat{U} - \langle\Psi_i^{(0)}|\hat{U}|\Psi_i^{(0)}\rangle\right)|\Psi_i^{(0)}\rangle. \tag{2.57}$$

This equation is multiplied by $\langle\Psi_n^{(0)}|$ yielding,

$$\left(E_i^{(0)} - E_n^{(0)}\right)\langle\Psi_n^{(0)}|\Psi_i^{(1)}\rangle = \langle\Psi_n^{(0)}|\hat{U}|\Psi_i^{(0)}\rangle \tag{2.58}$$
$$\langle\Psi_n^{(0)}|\Psi_i^{(1)}\rangle = \frac{\langle\Psi_n^{(0)}|\hat{U}|\Psi_i^{(0)}\rangle}{E_i^{(0)} - E_n^{(0)}}. \tag{2.59}$$

Thus, the expansion coefficients for $|\Psi_i^{(1)}\rangle$ are expressed in terms of the matrix elements of the perturbation in terms of zero-order wave functions and their energy differences. This allows to evaluate the second order energy correction after inserting the expansion 2.56 into

equation 2.52,

$$E_i^{(2)} = \langle \Psi_i^{(0)} | \hat{U} \sum_{n,n\neq i} |\Psi_n^{(0)}\rangle\langle\Psi_n^{(0)}|\Psi_i^{(1)}\rangle \tag{2.60}$$

$$= \sum_{n,n\neq i} \frac{\langle \Psi_i^{(0)}|\hat{U}|\Psi_n^{(0)}\rangle\langle\Psi_n^{(0)}|\hat{U}|\Psi_i^{(0)}\rangle}{E_i^{(0)} - E_n^{(0)}} = \sum_{n,n\neq i} \frac{|\langle \Psi_i^{(0)}|\hat{U}|\Psi_n^{(0)}\rangle|^2}{E_i^{(0)} - E_n^{(0)}}. \tag{2.61}$$

Also the energies of third and higher orders can be expressed in terms of the zero-order solutions. For the third order energy the second order wave function correction is expanded,

$$|\Psi_i^{(2)}\rangle = \sum_{n,n\neq i} |\Psi_n^{(0)}\rangle\langle\Psi_n^{(0)}|\Psi_i^{(2)}\rangle, \tag{2.62}$$

and the above procedure is repeated using the second order and lower order equations. [47]

In MPPT the zero-order Hamilton operator is the HF Hamiltonian, which is the sum of the Fock operators for each electron,

$$\hat{H}_0 = \sum_{i=1}^{N} \hat{f}(i) \tag{2.63}$$

$$= \sum_{i=1}^{N} \left(\hat{h}(i) + \sum_{b=1}^{N} \hat{\mathcal{J}}_b(i) - \hat{\mathcal{K}}_b(i) \right) \tag{2.64}$$

$$= \sum_{i=1}^{N} \hat{h}(i) + 2\left\langle \hat{V}_{ee} \right\rangle, \tag{2.65}$$

where $\left\langle \hat{V}_{ee} \right\rangle$ is the total mean field electron-electron interaction, which is counted twice in the sum of Fock operators. The perturbation is defined by the difference of the HF to the exact electronic Hamiltonian,

$$\hat{U} = \hat{H}_{\mathrm{el}} - \hat{H}_0 = \hat{V}_{ee} - 2\left\langle \hat{V}_{ee} \right\rangle. \tag{2.66}$$

This perturbation is termed fluctuation potential, as it induces fluctuations to a mean field HF description of the electronic motions. As we know from the exact CI wave function, the fluctuations will be represented by excited determinants. The zero-order wave function is the HF ground state Slater-determinant

$$\langle \Psi^{(0)}| = \langle \Psi_0^{(0)}| = \langle \Psi_0^{\mathrm{HF}}| \tag{2.67}$$

and the sum of occupied spin orbital energies is the zero-order energy

$$\mathrm{MP0} = E^{(0)} = E_1^{(0)} = \sum_{i=1}^{N} \epsilon_i. \tag{2.68}$$

The first order correction $E^{(1)}$ is then

$$E^{(1)} = \langle \Psi_0^{\mathrm{HF}}|\hat{U}|\Psi_0^{\mathrm{HF}}\rangle = -\left\langle \hat{V}_{ee} \right\rangle, \tag{2.69}$$

which corrects the double counting of the electron repulsion and gives the MP1 energy as

$$\mathrm{MP1} = \mathrm{MP0} + E^{(1)} = E_0^{\mathrm{HF}}, \tag{2.70}$$

which is just the HF energy. Hence, the correlation energy is given by the sum over all orders starting from second order,

$$E_{\mathrm{corr}} = E^{(2)} + E^{(3)} + \cdots \tag{2.71}$$

The MP2 electron correlation method approximates

$$E_{\mathrm{corr}} \approx E^{(2)} = E_{\mathrm{corr}}^{\mathrm{MP2}} \tag{2.72}$$

and thus the MP2 energy is defined as

$$\mathrm{MP2} = \mathrm{MP0} + E^{(1)} + E^{(2)} = E_0^{\mathrm{HF}} + E^{(2)} = E_0^{\mathrm{HF}} + E_{\mathrm{corr}}^{\mathrm{MP2}}. \tag{2.73}$$

To evaluate the second order energy correction, the second order wave function correction is expanded in eigenfunctions of the HF Hamiltonian, as in the CI expansion, see equation 2.39. Then, according to equation 2.61 from RSPT, matrix elements of the perturbation operator with the HF determinant and singly, doubly, and higher excited determinants have to be evaluated. However, it can be shown, that $E^{(2)}$ depends only on matrix elements involving the HF determinant and doubly excited determinants. Higher than doubly excited determinants give zero, when operated on by a two electron operator as $\hat{V}_{ee}$. Furthermore the matrix elements with single excited determinants, are zero as

$$\langle \Psi_0^{\mathrm{HF}} | \hat{U} | \Psi_a^r \rangle = \langle \Psi_0^{\mathrm{HF}} | \hat{H}_{\mathrm{el}} | \Psi_a^r \rangle - \epsilon_r \langle \Psi_0^{\mathrm{HF}} | \Psi_a^r \rangle, \tag{2.74}$$

where $\langle \Psi_0^{HF} | \hat{H}_{\mathrm{el}} | \Psi_a^r \rangle = 0$, due to Brillouins theorem, which is a consequence of the orthogonality of the MO's, and $\langle \Psi_0^{HF} | \Psi_a^r \rangle = 0$ due to the orthogonality of the Slater determinants. Hence, in terms of MPPT equation 2.61 is

$$E^{(2)}(\mathrm{MP2}) = E_{\mathrm{corr}}^{\mathrm{MP2}} = -\sum_{ab}^{occ} \sum_{rs}^{virt} \frac{\langle \Psi_0^{HF} | \hat{U} | \Psi_{ab}^{rs} \rangle \langle \Psi_{ab}^{rs} | \hat{U} | \Psi_0^{HF} \rangle}{E_0^{\mathrm{HF}} - E_{ab}^{rs}} \tag{2.75}$$

$$= -\sum_{ab}^{occ} \sum_{rs}^{virt} \frac{\left[(ar|bs) - (as|br) \right]^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}, \tag{2.76}$$

where the MP2 correlation energy is calculated from two electron integrals and orbital energies. The equations of higher-order, MP3, MP4 etc. become increasingly complex. From the MP4 level on, singly, doubly, triply and quadruply excited determinants enter the respective equations. MP2 recovers about 80-90%, while MP3 and MP4 give about 90-95% to 95-98% of the correlation energy. The MPPT shows, that most of the ground state correlation energy is recovered by accounting for doubly excited determinants, while the single, double and triple excitations are of similar but lesser importance. As MPPT is not variational, the MPn energy is not an upper bound to the exact energy. Often, an oscillation of the energies of the different orders around the exact value is observed. Of great importance for energy differences, is the size extensivity of the MPn methods, as this does not introduce additional errors when the

system size is changed. And one can hope, that the errors in the correlation energy effectively cancel out when energy differences are calculated.[52]

In this work, we apply MP2 in chapter 3 to economically calculate pair potentials for H_2-F_2 and H_2-Br_2. The local variant of MP2, LMP2 is used in chapter 4 to calculate the major part of the correlation contribution to the intramolecular interactions between perfluoroalkyl groups in hairpin conformers. In all cases the MP2 results can be trusted, as only closed shell molecules are involved which are reliably described by single reference methods. However, when greater accuracy is desired the coupled cluster methods, described in the next section, are used.

2.2.2.2 Coupled cluster theory

Coupled cluster theory introduces electron correlation in a size extensive manner by expanding the electronic wave function as a product series of excitation operators working on the HF ground state. Excitation operators correlate electrons by excitation from occupied spin orbitals into unoccupied, *i.e.* virtual, spin orbitals. As a first example, we consider the double-excitation operator that works on the HF reference state, *i.e.* the uncorrelated state with spin orbital I and J already occupied,

$$\hat{\tau}_{IJ}^{AB} = a_A^\dagger a_I a_B^\dagger a_J, \tag{2.77}$$

where a_J and a_I annihilate electrons in the spin-orbitals J and I, whereas $a_A^\dagger$ and $a_B^\dagger$ create electrons in the unoccupied spin-orbitals A and B. Successive application of double-excitation operators correlating the same pair of electrons gives zero, as there is zero possibility to annihilate an electron in spin orbital J twice, *i.e.*

$$\hat{\tau}_{IJ}^{AB}\hat{\tau}_{IJ}^{CD} = 0. \tag{2.78}$$

Taking into account all possible double-excitation operators for a HF reference state $|\mathrm{HF}\rangle = |\Psi_0^{\mathrm{HF}}\rangle$ the cluster expansion becomes,

$$|\mathrm{CCD}\rangle = \left[\prod_{A>B,I>J}\left(1 + t_{IJ}^{AB}\hat{\tau}_{IJ}^{AB}\right)\right]|\mathrm{HF}\rangle, \tag{2.79}$$

the coupled cluster doubles (CCD) wave function, where t_{IJ}^{AB} denotes the amplitude of the respective double-excitation process. In general, also single, triple and higher excitation operators should be accounted for. An unspecified excitation operator be $\hat{\tau}_\mu$ with the related amplitude t_μ, then the CC wave function is

$$|\mathrm{CC}\rangle = \left[\prod_\mu (1 + t_\mu\hat{\tau}_\mu)\right]|\mathrm{HF}\rangle, \tag{2.80}$$

with it's product ansatz fundamentally different from an CI wave function

$$|\mathrm{CI}\rangle = \left[1 + \sum_\mu C_\mu\hat{\tau}_\mu\right]|\mathrm{HF}\rangle. \tag{2.81}$$

It is obvious, that amplitudes are different from the CI coefficients, as they are related to an excitation process involved in creation of several determinants. Still, an excited determinant $|\mu\rangle$ is generated by the action of $\hat{\tau}_\mu$ on $|\mathrm{HF}\rangle$,

$$|\mu\rangle = \hat{\tau}_\mu|\mathrm{HF}\rangle. \tag{2.82}$$

Using this expression we can expand the CC wave function

$$|\mathrm{CC}\rangle = \left(1 + \sum_\mu t_\mu \hat{\tau}_\mu + \sum_{\mu>\nu} t_\mu t_\nu \hat{\tau}_\mu \hat{\tau}_\nu + \cdots\right)|\mathrm{HF}\rangle \tag{2.83}$$

$$= |\mathrm{HF}\rangle + \sum_\mu t_\mu|\mu\rangle + \sum_{\mu>\nu} t_\mu t_\nu|\mu\nu\rangle + \cdots, \tag{2.84}$$

where it becomes clear, that due to the product ansatz cluster excitations not only generate $|\mu\rangle$ but also other higher excited determinates. In general, one distinguishes connected clusters and their amplitudes,

$$t_{\mu\nu}\hat{\tau}_{\mu\nu}|\mathrm{HF}\rangle = t_{\mu\nu}|\mu\nu\rangle \tag{2.85}$$

where one excitation operator generates a certain determinant $|\mu\nu\rangle$, and disconnected clusters,

$$t_\mu t_\nu \hat{\tau}_\mu \hat{\tau}_\nu|\mathrm{HF}\rangle = t_\mu t_\nu|\mu\nu\rangle, \tag{2.86}$$

where the same determinant $|\mu\nu\rangle$ is now generated by an other process represented by a product of excitation operators. The total amplitude to generate a certain determinant is then the sum of one connected and all possible disconnected cluster amplitudes. The solution to the coupled cluster equations is obtained by determining the amplitudes. Here, due to the non-linearity of the ansatz the variational principle leads to a set of equations that are more complicated than the FCI equations. Thus, in CC theory a set of equations for the amplitudes is obtained by projection of the Schrödinger equation against determinants,

$$\langle\mu| = \langle\mathrm{HF}|\hat{\tau}_\mu^\dagger \tag{2.87}$$

generated by connected clusters. These equations are the projected coupled cluster equations,

$$\langle\mu|\hat{H}|\mathrm{CC}\rangle = E\langle\mu|\mathrm{CC}\rangle, \tag{2.88}$$

where the energy E is given by projection against the HF wave function,

$$E = \langle\mathrm{HF}|\hat{H}|\mathrm{CC}\rangle. \tag{2.89}$$

The projected CC equations are non variational, however the difference to the exact CC energy is usually negligibly small [53]. The projected equations are solvable, as $\hat{H}$ couples only those determinants, that differ by more than two-fold excitations, thereby effectively truncating the CC expansion in the equations.

Instead of the product ansatz used so far, the commutation relations of the excitation operators and the fact that $\hat{\tau}_\mu^2 = 0$, allow to write the CC wave function in terms of an exponential ansatz,

$$|\mathrm{CC}\rangle = \exp(\hat{T})|\mathrm{HF}\rangle, \tag{2.90}$$

where $\hat{T}$ is the cluster operator given by the sum over all excitation operators and their amplitudes,

$$\hat{T} = \sum_{\mu} t_\mu \hat{\tau}_\mu. \tag{2.91}$$

Cluster operators of different hierarchies collect all single, double, triple, and higher excitation operators

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N. \tag{2.92}$$

In this way, it becomes possible to restrict the theory to a certain order of connected excitation processes. It is also interesting to determine which cluster operators are involved in creating a certain determinant,

$$\exp(\hat{T})|\mathrm{HF}\rangle = \sum_{i=0}^{N} \hat{C}_i |\mathrm{HF}\rangle \tag{2.93}$$

represented by an operator $\hat{C}_i$. By expanding $\exp(\hat{T})$ and equating with each $\hat{C}_i$ one finds the following relations,

$$\hat{C}_0 = 1 \tag{2.94}$$

$$\hat{C}_1 = \hat{T}_1 \tag{2.95}$$

$$\hat{C}_2 = \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \tag{2.96}$$

$$\hat{C}_3 = \hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3 \tag{2.97}$$

$$\hat{C}_4 = \hat{T}_4 + \hat{T}_3\hat{T}_1 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_2\hat{T}_1^2 + \frac{1}{24}\hat{T}_1^4. \tag{2.98}$$

Doubly excited determinants are reached via two processes, the connected excitation of two electrons and two disconnected simultaneous single excitations. For triply excited determinants one finds three excitation processes, the connected three electron excitation, the disconnected combination of a two electron and a single electron excitation and the three disconnected single excitations. A quadruply exited determinant is created by five distinct processes. We immediately see, that with the inclusion of only single and double excitation processes (CCSD) all single and doubly excited determinants are generated, as in CISD. Furthermore, also higher exited determinants, described by products of excitations, enter the amplitude equations. In truncated CC approximations not all processes are described, however higher connected excitation may only have small amplitudes. Solving the projected cluster equations and determining an amplitude then means finding the probability of the excitation process leading to all determinants that involve this type of excitation. In CCSD, the most important correlation processes are included, that is pair correlation via $\hat{T}_2$, and pair-pair interactions via $\hat{T}_2^2$. The single excitations have only a small effect, if canonical orbitals are used. Selected quadruplets are described in CCSD mostly via the process $\hat{T}_2^2$. The truncation leads to a neglect of the less probable processes $\hat{T}_4$ and $\hat{T}_3\hat{T}_1$. The main error in CCSD arises due to the neglect of the connected triple cluster operators $\hat{T}_3$, because $\hat{T}_1^3$ and $\hat{T}_2\hat{T}_1$ have a low weight for triples as $\hat{T}_1$ is less important. For the energy, the triples contribution can be approximated using MPPT, which is known as the CCSD(T) method. [52, 53]

The connection between CC and MPPT methods can be investigated by using the exponential cluster ansatz

$$|\Psi^{(n)}\rangle = [\exp(\hat{T})]^{(n)}|\mathrm{HF}\rangle, \tag{2.99}$$

instead of the linear ansatz in equation 2.54 to expand the n^{th} order correction to the the wave function. The resulting theory is the coupled cluster perturbation theory (CCPT). Accordingly, the cluster operator and the cluster amplitudes are expanded in terms of orders in the perturbation,

$$\hat{T} = \hat{T}^{(0)} + \hat{T}^{(1)} + \hat{T}^{(2)} + \cdots + \hat{T}^{(n)} \tag{2.100}$$

$$t_\mu = t_\mu^{(0)} + t_\mu^{(1)} + t_\mu^{(2)} + \cdots + t_\mu^{(n)} \tag{2.101}$$

where $\hat{T}^{(n)}$ and $t_\mu^{(n)}$ are the cluster operator and cluster amplitude corrections at the n^{th} order of perturbation theory. Up to second order the wave functions are,

$$|\Psi^{(0)}\rangle = |\mathrm{HF}\rangle \tag{2.102}$$

$$|\Psi^{(1)}\rangle = \hat{T}^{(1)}|\mathrm{HF}\rangle \tag{2.103}$$

$$|\Psi^{(2)}\rangle = \left(\hat{T}^{(2)} + \frac{1}{2}\hat{T}^{(1)}\hat{T}^{(1)}\right)|\mathrm{HF}\rangle. \tag{2.104}$$

The expressions can be further expanded in terms of excitations as

$$|\Psi^{(1)}\rangle = |\mathrm{MP1}\rangle = \hat{T}_2^{(1)}|\mathrm{HF}\rangle \tag{2.105}$$

$$|\Psi^{(2)}\rangle = |\mathrm{MP2}\rangle = \left(\hat{T}_1^{(2)} + \hat{T}_2^{(2)} + \hat{T}_3^{(2)} + \frac{1}{2}\hat{T}_2^{(1)}\hat{T}_2^{(1)}\right)|\mathrm{HF}\rangle. \tag{2.106}$$

The first order correction to the cluster operator is just the connected double cluster operator. At second order, connected single, triple and disconnected quadruple cluster excitations enter the description. The energy corrections are,

$$E^{(0)} = E_0 \tag{2.107}$$

$$E^{(1)}(\mathrm{MP1}) = \langle \mathrm{HF}|\hat{U}|\mathrm{HF}\rangle \tag{2.108}$$

$$E^{(2)}(\mathrm{MP2}) = \langle \mathrm{HF}|[\hat{U}, \hat{T}_2^{(1)}]|\mathrm{HF}\rangle \tag{2.109}$$

$$E^{(3)}(\mathrm{MP3}) = \langle \mathrm{HF}|[\hat{U}, \hat{T}_2^{(2)}]|\mathrm{HF}\rangle \tag{2.110}$$

$$E^{(4)} = \langle \mathrm{HF}|[\hat{U}, \hat{T}_2^{(3)}]|\mathrm{HF}\rangle, \tag{2.111}$$

where up to fourth order, only connected doubles amplitudes contribute directly. Thus, the energy expression for the MP2 energy correction, presented earlier in equation 2.75, can be written in terms of connected doubles amplitudes. The cluster operator appearing in the $E^{(2)}(\mathrm{MP2})$ term is,

$$\hat{T}_2^{(1)} = \sum_{ab}^{\mathrm{occ}} \sum_{rs}^{\mathrm{virt}} t_{ab}^{rs} \hat{\tau}_{ab}^{rs}, \tag{2.112}$$

This is inserted into the equation for the MP2 energy correction,

$$E^{(2)}(\mathrm{MP2}) = \langle \mathrm{HF}|[\hat{U}, \hat{T}_2^{(1)}]|\mathrm{HF}\rangle \tag{2.113}$$
$$= \langle \mathrm{HF}|\hat{U}\hat{T}_2^{(1)}|\mathrm{HF}\rangle \tag{2.114}$$
$$= \langle \mathrm{HF}|\hat{U}\sum_{ab}^{\mathrm{occ}}\sum_{rs}^{\mathrm{virt}} t_{ab}^{rs\,(1)}\hat{\tau}_{ab}^{rs}|\mathrm{HF}\rangle \tag{2.115}$$
$$= \sum_{ab}^{\mathrm{occ}}\sum_{rs}^{\mathrm{virt}} t_{ab}^{rs\,(1)}\langle \mathrm{HF}|\hat{U}|\psi_{ab}^{rs}\rangle, \tag{2.116}$$

Upon comparison with the expression in equations 2.75 and 2.76 the connected doubles amplitudes are,

$$t_{ab}^{rs\,(1)} = -\frac{\langle \psi_{ab}^{rs}|\hat{U}|\mathrm{HF}\rangle}{E_0^{\mathrm{HF}} - E_{ab}^{rs}} \tag{2.117}$$
$$= -\frac{\langle \psi_{ab}^{rs}|\hat{H}|\mathrm{HF}\rangle}{E_0^{\mathrm{HF}} - E_{ab}^{rs}} \tag{2.118}$$
$$= -\frac{(ar|bs) - (as|br)}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}. \tag{2.119}$$

Calculation of an MP2 energy thus involves in principle the calculation of the connected double amplitudes from CC theory to first order in the perturbation, *i.e.* the fluctuation potential. It is further possible to show, that the doubles amplitudes obtained in the CCSD approximation are correct up to second order in perturbation theory [53]. The energy corrections are described exactly up to the third order within CCSD. Thus, performing a CCSD calculation leads to amplitudes that can be used to also express the MP2 and MP3 energies.

In hybrid CC methods, CC and MPPT theories are combined, in that contributions for the highest excitations are approximated to a lower order in the fluctuation potential. Iterative methods are termed CCN, where for example CC2 is derived from CCSD by including the doubles cluster operators to first order only, CC3 is an approximation to CCSDT, in that triples cluster operators are included just to second order. The CCN methods generate improved wave functions at a reduced cost compared to the respective non hybrid CC method, which makes CC3 ideal to calculate highly accurate properties like polarizabilities. When only the energy has to be calculated accurately, non iterative hybrid CC methods like CCSD(T) are more economical than CC3. In CCSD(T) a correction $\Delta E^{\mathrm{CCSD(T)}}$ is added to the CCSD energy,

$$E^{\mathrm{CCSD(T)}} = E^{\mathrm{CCSD}} + \Delta E^{\mathrm{CCSD(T)}}. \tag{2.120}$$

The CCSD(T) correction is given by

$$\Delta E^{\mathrm{CCSD(T)}} = \langle \bar{t}|[\hat{U},^{*} \hat{T}_3^{(2)}]|\mathrm{HF}\rangle \tag{2.121}$$

and includes connected triples terms arising in the fourth and fifth order CCPT energy corrections. In the evaluation of this expression the amplitudes from the prior CCSD calculation are used. CCSD(T) describes the energy at least as accurate as MP4, at the same computational costs. Therefore, CCSD and CCSD(T) are used rather than MP3 and MP4, when a higher than MP2 accuracy is desired.

The order of the methods in terms of accuracy in the correlation energy is

$$\text{MP2} \approx \text{CC2} < \text{CISD} < \text{CCSD} < \text{MP4} < \text{CCSD(T)} \approx \text{CC3} < \text{CCSDT(Q)},$$

where substantial improvements upon MP2 come from the cluster model with CCSD, which in turn is substantially improved by the (T) correction in CCSD(T) [52, 53]. For weak intermolecular interactions of closed shell systems, CCSD(T) is a very accurate method. Hence, in this work CCSD(T) is used to benchmark the MP2 results for the hydrogen-halogen van der Waals interactions in chapter 3. The local variant LCCSD(T0) is used to calculate a correction term to the LMP2 correlation energy for the relative energy of alkanes and perfluoroalkylalkanes hairpin conformers in chapter 4.

2.2.2.3 Explicitly correlated methods

One speaks of explicitly correlated methods, when terms that explicitly depend linearly on the inter-electronic distance r_{12} are considered in the ansatz for the wave function. These methods greatly improve the convergence behaviour of the correlation energy of standard correlation methods with respect to basis set size. The reason for this improvement, is the inherent failure of standard correlation methods, here we assume finite basis sets, to describe the correct behaviour of spatially close electrons. The true electronic wave function has a Coulomb cusp, a discontinuous derivative, at $r_{12} = 0$, where the electron-electron repulsion operator has a singularity. At the Coulomb cusp, the derivative of the wave function has a constant value

$$\left(\frac{\partial \Psi}{\partial r_{12}}\right)_{r_{12}=0} = \frac{1}{2}\Psi\left(r_{12}=0\right). \tag{2.122}$$

To describe the Coulomb cusp of two electron systems accurately, Hylleraas type wave functions include orbital products and an expansion in all powers of r_{12},

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \exp(-\alpha_1 \mathbf{r}_1)\exp(-\alpha_2 \mathbf{r}_2)\sum_{klm} C_{klm}(\mathbf{r}_1 + \mathbf{r}_2)^k(\mathbf{r}_1 - \mathbf{r}_2)^l r_{12}^m. \tag{2.123}$$

This ansatz yields highly accurate results, but can only be applied to systems with few electrons. Conventional CI wave functions using Gaussian type basis sets depend only on even powers of the inter-electronic distance. Therefore, the derivatives of the CI wave function with respect to r_{12} at $r_{12} = 0$ yield zero, which leads to an inherently poor description of the the Coulomb cusp region of the electronic wave function [53]. In turn, to converge the correlation energy, very large basis sets are needed. In explicitly correlated methods, terms linear in r_{12} are added to the wave function. For the resulting R12 wave function the Coulomb cusp condition is then satisfied by construction, which improves the convergence behaviour of the correlation energy to the basis set limit. A simplified R12 wave function ansatz is,

$$\Psi_{\text{R12}} = \Psi_{\text{HF}} + \sum_{ijab} a_{ijab}\Psi_{ij}^{ab} + \sum_{ij} b_{ij} r_{ij}\Psi_{\text{HF}}, \tag{2.124}$$

where a sum of products over the HF wave function and the inter-electronic distance are included. The R12 ansatz in conjunction with CI, CC and MPPT methods, gives rise to matrix elements over the two electron operators, which involve three and four electron integrals such

as,

$$\langle\psi_i(1)\psi_j(2)\psi_k(3)|\frac{r_{12}}{r_{13}}|\psi_i(1)\psi_j(2)\psi_k(3)\rangle \tag{2.125}$$

$$\langle\psi_i(1)\psi_j(2)\psi_k(3)|\frac{r_{12}r_{23}}{r_{13}}|\psi_i(1)\psi_j(2)\psi_k(3)\rangle \tag{2.126}$$

$$\langle\psi_i(1)\psi_j(2)\psi_k(3)\psi_l(4)|\frac{r_{12}r_{13}}{r_{23}}|\psi_i(1)\psi_j(2)\psi_k(3)\psi_l(4)\rangle. \tag{2.127}$$

The sheer numbers in the order of K^6 and K^8 of such three and four electron integrals in the AO basis, make the R12 treatment very expensive. Introduced by Kutzelnigg *et al.* [54], the "resolution of the identity" (RI), or the equivalent density fitting (DF) techniques are key to evaluate the latter integrals, which makes R12 applicable to larger systems. Using the RI, four and three-electron integrals are approximated by products of two-electron integrals. In a complete basis the identity can be expressed by

$$\mathbf{1} = \sum_{p}^{\infty} |\psi_p\rangle\langle\psi_p| = \sum_{pqr}^{\infty} |\psi_p\psi_q\psi_r\rangle\langle\psi_p\psi_q\psi_r|. \tag{2.128}$$

Then, using the "second" identity, a three electron integral of a product of inter-electronic distances can be written as a product of two two electron integrals over two-electron operators,

$$\langle\psi_i(1)\psi_j(2)\psi_k(3)|r_{12}\mathbf{1}\frac{1}{r_{13}}|\psi_i(1)\psi_j(2)\psi_k(3)\rangle = \tag{2.129}$$
$$\sum_{pqr}^{\infty}\langle\psi_i\psi_j\psi_k|r_{12}|\psi_p\psi_q\psi_r\rangle\langle\psi_p\psi_q\psi_r|\frac{1}{r_{13}}|\psi_i\psi_j\psi_k\rangle.$$

Each integral reduces to a two-electron integral, because the two-electron operator acts only on the functions of two electrons, thus

$$\langle\psi_i\psi_j\psi_k|\frac{r_{12}}{r_{13}}|\psi_i\psi_j\psi_k\rangle = \sum_{pqr}^{\infty}\delta_{kr}\langle\psi_i\psi_j|r_{12}|\psi_q\psi_p\rangle\left(\delta_{qj}\langle\psi_q\psi_r|\frac{1}{r_{13}}|\psi_i\psi_k\rangle\right). \tag{2.130}$$

Since the MOs are orthonormal, we can evaluate the δ-functions and find, that the sum runs over p only,

$$\langle\psi_i\psi_j\psi_k|\frac{r_{12}}{r_{13}}|\psi_i\psi_j\psi_k\rangle = \sum_{p}^{\infty}\langle\psi_i\psi_j|r_{12}|\psi_q\psi_p\rangle\left(\langle\psi_q\psi_r|\frac{1}{r_{13}}|\psi_i\psi_k\rangle\right). \tag{2.131}$$

While exact in a complete basis, in a finite basis the RI gives only an approximation of the desired integral. This error can be controlled by the size of the RI basis. The RI and DF schemes are general and can be applied also to approximate integrals that occur in standard methods, thereby further decreasing the cost of HF, MPn and CCSD(T) methods. Very large basis sets may be avoided using R12 methods, but still large auxiliary basis set are needed to give accurate R12 results. [52]

Modern variants of R12 theory, are often termed F12 methods, for expample MP2-F12 and CCSD-F12. The name emphasizes the use of a general explicit correlation factor $\hat{F}_{12}$. If $\hat{F}_{12} = r_{12}$, the correlation factor is that of the original R12 method. In the implementations used in this work an exponential correlation function $F_{12}(r_{12}) = \exp(-\beta r_{12})$ is employed instead. To

increase the accuracy of the RI approximation with regular basis sets, complementary auxiliary basis functions (CABS) are used. The CABS are linear combinations of functions of the AO basis and of an additional auxiliary basis (ABS), which are orthogonalized upon the MOs. This gives additional flexibility in the RI, without the need to employ large basis sets in the non explicitly correlated terms or the initial HF calculation. Given the basis set incompleteness error $\Delta E^{\mathrm{corr}}_{\mathrm{BSIE}}$ for the correlation energy has been effectively reduced by the F12 treatment, the major part of the basis set incompleteness error for the total electronic energy is then due to the HF energy, which was evaluated in a smaller basis, *i.e.* without the CABS. The error $\Delta E^{\mathrm{HF}}_{\mathrm{BSIE}}$ can be reduced without the need for an new HF calculation in a larger basis, by the CABS singles correction. For this, the relaxation of the occupied MOs due to the presence of the CABS is calculated using RSPT in terms of single excitations.

Lastly, we take a look at how explicitly correlated terms are included in the wave function of the MP2-F12 and CCSD(T)-F12 methods [55, 56]. The MP2-F12 wave function is given by,

$$|\text{MP2-F12}\rangle = \underbrace{\sum_{i<j,a<b} t^{ij}_{ab}|\psi^{ab}_{ij}\rangle}_{|\text{MP2}\rangle} + \sum_{i<j,x<y} t^{ij}_{xy}|\psi^{xy}_{ij}\rangle, \tag{2.132}$$

where the first sum runs over conventional doubly excited configurations and the second sum runs over explicitly correlated doubly excited configurations. The latter terms are given by

$$|\psi^{xy}_{ij}\rangle = |\psi^{ab}_{ij}\rangle\mathcal{F}^{xy}_{\alpha\beta}, \tag{2.133}$$

with

$$\mathcal{F}^{xy}_{\alpha\beta} = \langle xy|\hat{F}_{12}\hat{Q}_{12}|\alpha\beta\rangle, \tag{2.134}$$

where $|\psi^{ab}_{ij}\rangle$ is the full space of doubly exited configurations, $\hat{F}_{12}$ is the correlation factor and $\hat{Q}_{12}$ is a projector that ensures strong orthogonality of the explicitly correlated functions to the reference function. The CABS approximation enters the equations via the strong orthogonality projector $\hat{Q}_{12}$. The CCSD-F12 ansatz includes the conventional single cluster operator and singles amplitudes. The doubles cluster operator however, includes not only the usual double excitations with corresponding amplitudes but also additional explicitly correlated double excitations with amplitudes $\mathcal{T}^{ij}_{\alpha\beta}$,

$$\hat{T}_2 = \sum_{i<j,a<b} T^{ij}_{ab}\hat{\tau}^{ab}_{ji} + \sum_{i<j,\alpha<\beta} \mathcal{T}^{ij}_{\alpha\beta}\hat{\tau}^{\alpha\beta}_{ji}. \tag{2.135}$$

The explicitly correlated doubles amplitudes

$$\mathcal{T}^{ij}_{\alpha\beta} = \mathcal{F}^{xy}_{\alpha\beta}T^{ij}_{xy} = \langle xy|\hat{F}_{12}\hat{Q}_{12}|\alpha\beta\rangle T^{ij}_{xy}, \tag{2.136}$$

depend via $\mathcal{F}^{xy}_{\alpha\beta}$ on $\hat{F}_{12}$ and on $\hat{Q}_{12}$.

In this work, we have used the local MP2-F12 [57] and local CCSD(T0)-F12 [58] methods using the ansätze **3***A(LOC) and **3***A(LOC,FIX), see chapter 4. This means, that ansatz 3 is used for local, domain specific ('LOC') projectors $\hat{Q}_{12}$, and explicitly correlated amplitudes are completely specified by the cusp condition ('FIX'). Further, the 'diagonal' ansatz for the excited configurations is used and all matrix elements involving exchange operators are neglected (A), which anyway are very small, due to the local basis. Last, the equations for

the conventional and the explicitly correlated amplitudes are decoupled (*), this is known as the extended Brillouin condition, thus they can be solved independently. This combination of F12-RI approximations is well suited for calculations of weak intermolecular interactions using local correlation methods, because it is free of geminal basis set superposition errors [59].

2.2.2.4 Local correlation methods

Inclusion of electron correlation in the description of the electronic energy, is a necessity for the accurate description of molecular interactions, and for molecular properties in general. One bottleneck of the canonical correlation methods like MP2 and CCSD(T) is their scaling behaviour with the system size. CCSD(T) and MP4 scale formally with K^7, CCSD with K^6, and MP2 scales with K^5, where K denotes the basis set size. This scaling limits the applicability of these methods, in conjunction with the large basis sets needed for accurate results, to small molecules. One approach to achieve lower order scaling variants of correlation methods is the local correlation method by Pulay and Saebø [60, 61], which was further developed by Werner, Hampel and Schütz [40]. The idea behind all local methods is to exploit the fact, that spatially close electrons are stronger correlated than electrons that are far away from each other. To achieve this, the correlation energy is evaluated in a basis of occupied localized molecular orbitals (LMO) to which non-orthogonal virtual correlation subspaces, termed domains, are generated from projected atomic orbitals (PAO). Compared to the canonical methods, where MOs are delocalized over the whole system, LMO domains considerably reduce the size of the virtual excitation space. Further, local methods allow to approximate or even neglect correlation between electron pairs based on distance criteria.

Domain based local correlation treatments involve the following procedure. In the fist step, m occupied local orbitals ψ_i^{LMO} are generated for a closed shell reference determinant $|\mathrm{HF}\rangle$,

$$|\psi_i^{\mathrm{LMO}}\rangle = \sum_{k=1}^{K} R_{ik}|\phi_k\rangle, \tag{2.137}$$

where ϕ_k are atom centred basis functions and R_{ik} are coefficients. Usually, the localization method of Pipek and Mezey [62] is used. In the case of saturated molecules, as treated in this work in chapter 4, the Foster-Boys localization method [63] gives similar LMOs and results in an equivalent treatment. Next, a virtual space is constructed by projecting the atomic orbitals onto the occupied local orbitals

$$|\tilde{\phi_k}\rangle = \left(1 - \sum_{i=1}^{m} |\psi_i^{\mathrm{LMO}}\rangle\langle\psi_i^{\mathrm{LMO}}|\right)|\phi_k\rangle = \sum_{\rho=1}^{K} \tilde{R}_{\rho k}|\phi_\rho\rangle. \tag{2.138}$$

The PAOs $\tilde{\phi}_k$ are by construction orthogonal to the occupied orbitals, but are non-orthogonal amongst themselves.

In the third step, orbital domains are constructed by selecting a subset of PAOs for each occupied LMO. The procedure to assign an orbital domain $[i]$ to an LMO ϕ_i^{LMO} goes as follows. For a given LMO ϕ_i^{LMO}, the atoms that dominate its charge density are found by ranking the atoms according to their Mulliken charges up to a cut-off value of 1.8. No more atoms are added to the list, if the Boughton and Pulay (BP) completeness criterion [64] is fulfilled. This

check for completeness works in the following way. The functional

$$f(\mathbf{R}') = \min \left[\int \left(\psi_i^{\mathrm{LMO}} - \psi_i^{'} \right)^2 d\tau \right], \tag{2.139}$$

computes how well the approximate orbital

$$\psi_i^{'} = \sum_{\rho \in [i]} |\phi_\rho\rangle R_{\rho i}^{'}, \tag{2.140}$$

built only from AOs at the atoms in orbital domain [i], can describe the original orbital ψ_i^{LMO}. A typical BP threshold for completeness is $f(\mathbf{R}') < 0.02$. The value of the threshold depends on the basis set. For example, using basis sets with more diffuse functions can make lower thresholds ($f(\mathbf{R}') < 0.01$) necessary.

Once the domains are defined, excitations from the LMOs to the PAOs can be considered. For local CCSD single and double excitations are needed. Singly excited local configurations $\Phi^{\tilde{\mu}}$ are defined by the action of one spin-coupled excitation operator onto the reference determinant,

$$\Phi_i^{\tilde{\mu}} = \hat{E}_{\tilde{\mu} i} |0\rangle. \tag{2.141}$$

The excitation operators are defined as,

$$\hat{E}_{\tilde{\mu} i} = |\tilde{\phi}_\mu^\alpha\rangle\langle\psi_i^\alpha| + |\tilde{\phi}_\mu^\beta\rangle\langle\psi_i^\beta|. \tag{2.142}$$

The usage of non-orthogonal PAOs lead to non-orthogonal configurations,

$$\langle \Phi_i^{\tilde{\mu}} | \Phi_j^{\tilde{\nu}} \rangle = 2\delta_{ij} \tilde{S}_{\mu\nu}. \tag{2.143}$$

All possible single excitations from a given LMO are collected in the function

$$\Psi_i = \sum_\mu \tilde{t}_\mu^i \Phi_i^{\tilde{\mu}}, \tag{2.144}$$

with amplitudes $\tilde{t}_\mu^i$. Local double excitations are given by

$$\Phi_{ij}^{\tilde{\mu}\tilde{\nu}} = \hat{E}_{\tilde{\mu} i} \hat{E}_{\tilde{\nu} j} |0\rangle. \tag{2.145}$$

The double excitations are restricted to a joint virtual space $[ij]$, thus double excitations from ψ_i^{LMO} are possible into $[i]$ and $[j]$, which also holds for ψ_j^{LMO}. This joint virtual space $[ij]$ is called a pair domain. The pair functions that collect all pair excitations within one pair domain are given by

$$\Psi_{ij} = \sum_{\mu \in [ij]} \sum_{\nu \in [ij]} \tilde{T}_{\mu\nu}^{ij} \Phi_{ij}^{\tilde{\mu}\tilde{\nu}} \quad (i \geq j), \tag{2.146}$$

where $\tilde{T}_{\mu\nu}^{ij}$ are the local CCSD amplitudes. Some functions within the pair domain $[ij]$ might be redundant and they or their contribution at some stage in the determination of the amplitudes have to be discarded. [40]

The pair domain correlation energy ϵ_{ij} decreases with $\frac{1}{r^6}$, where r is the distance between the charge centers of the LMOs ψ_i^{LMO} and ψ_j^{LMO}. This allows to define distance criteria for

each pair domain, for which different levels of approximations are used to calculate the pair correlation energy or even neglect pair correlation all together. The implementation of Werner *et al.* [41] distinguishes up to five levels of approximations for pair correlation energies. Pairs are classified as either strong, close, weak, distant or very distant. Strong pairs, are treated at the full level of theory, say LCCSD, and LMP2 doubles amplitudes from close pairs are also included in the CC equations of the strong pairs. Close pairs are treated at the LMP2 level of theory and affect the strong pair amplitudes as described. Weak and distant pairs are treated at the LMP2 level of theory. Distant pairs differ from weak pairs, in that two electron integrals can be treated by multi-pole approximations. Last, very distant pairs are neglected. Either distance criteria, r_s, r_c, r_w, r_d and r_{vd}, or connectivity criteria, *i.e.* the number of bonds between atoms of the two LMOs, are used to invoke the pair approximations.

The exponential ansatz for the LCCSD wave function is

$$|\mathrm{LCCSD}\rangle = \exp\left(\hat{T}_\mathrm{loc}\right)|\mathrm{HF}\rangle, \tag{2.147}$$

where the cluster operator $\hat{T}_\mathrm{loc}$ includes domain and pair domain restricted single and double excitations, respectively. The cluster operator is given by

$$\hat{T}_\mathrm{loc} = \sum_i \sum_{\mu\in[i]} \tilde{t}^i_\mu \hat{E}_{\tilde{\mu}i} + \frac{1}{2}\sum_{ij}\sum_{\mu\in[ij]}\sum_{\nu\in[ij]} \tilde{T}^{ij}_{\mu\nu}\hat{E}_{\tilde{\mu}i}\hat{E}_{\tilde{\nu}j} \quad ;\tilde{T}^{ij}_{\mu\nu} = \tilde{T}^{ji}_{\nu\mu}. \tag{2.148}$$

The projected Schrödinger equations are

$$E = \langle \mathrm{HF}|\hat{H}|\mathrm{LCCSD}\rangle, \tag{2.149}$$

$$\tilde{v}^i_\mu = \langle \Phi^{\tilde{\mu}}_i|\hat{H} - E|\mathrm{LCCSD}\rangle = 0, \tag{2.150}$$

$$\tilde{V}^{ij}_\mu = \langle \tilde{\Phi}^{\tilde{\mu}\tilde{\nu}}_{ij}|\hat{H} - E|\mathrm{LCCSD}\rangle = 0 \qquad (i \geq j). \tag{2.151}$$

These non-linear equations are solved iteratively for the amplitudes $\tilde{t}^i_\mu$ and $\tilde{T}^{ij}_{\mu\nu}$. [40]

The perturbative treatment of local triples involves the following main approximations. Connected triples excitations are restricted to triples domains [ijk], build by uniting three strong pair domains [ij], [jk] and [ik]. The (T0) approximation further neglects certain couplings between triples and is non iterative. It is used the as a first guess of full iterative schemes (T), and recovers 97% of the full local triples energy [65, 66]. In this work, the local (T0) approximation is used exclusively.

Local methods can be used with explicitly correlated terms as described in the section 2.2.2.3. Other techniques to construct virtual excitation spaces for local correlation methods are the pair natural orbital (PNO) method [67] and the orbital specific virtuals (OSV) [68]. In the PNO method, a set of optimal virtual orbitals are found for each pair [ij], that describe the pair correlation energy with the least number of orbitals. In other words, optimally compact pair domains are found. This has the advantage, that potential energy surfaces have no discontinuities due to inconsistent pair domains, and the number of orbitals within pair domains are much lower compared to PAOs. However, one drawback is that the total number of distinct virtual orbitals in systems with many pairs, is much larger. In the OSV approach, orbital specific virtuals are generated, that are optimal for pair correlation, and pair domains are formed as the union of orbital spaces. OSV thereby can be seen as an intermediate between the PAO and PNO methods. A big benefit of OSVs over PAOs is that only one parameter controls the

accuracy of the local approximation, and that continuous potential energy surfaces are easier to obtain. [69] For the calculation of the relative energy of alkanes and perfluoroalkylalkanes hairpin conformers in chapter 4, we carefully applied PAO based local correlation methods.

2.2.2.5 Method of increments

In the method of increments (MoI), the correlation energy of a N electron system is systematically approximated by a number of correlation energy contributions obtained from smaller sub-sets of electrons [44, 70–72]. Similar earlier approaches have been made by Faddeev for scattering theory [73], Nesbet [74], who used Bethe-Goldstone equations of nth order and Stollhoff and Fulde [42, 43] with their Local Ansatz.

For an MoI treatment, the molecular orbitals are first localized. Then, the resulting LMOs are grouped together into bodies. Each body holds at minimum one LMO or a number of spatially close LMOs. However, then the procedure is a general one. The correlation energy is approximated by a MoI many body expansion,

$$E_{\text{corr.}} = \sum_{i} \epsilon_i + \sum_{\substack{i<j \\ i,j}} \Delta\epsilon_{ij} + \sum_{\substack{i<j<k \\ i,j,k}} \Delta\epsilon_{ijk} + \cdots , \tag{2.152}$$

as the sum over unique one-body ϵ_i, two-body $\Delta\epsilon_{ij}$, three-body $\Delta\epsilon_{ijk}$ and higher order increments. The one-body increment ϵ_i represents the energy due to correlation of only the electrons within the LMOs that make up the i^{th} body. The two-body increment

$$\Delta\epsilon_{ij} = \epsilon_{ij} - (\epsilon_i + \epsilon_j) \tag{2.153}$$

describes the part of the correlation energy ϵ_{ij} of only the electrons in two bodies i and j, which is not already described by ϵ_i and ϵ_j. The three-body increment $\Delta\epsilon_{ijk}$ accounts only for pure three-body effects and is given by

$$\Delta\epsilon_{ijk} = \epsilon_{ijk} - (\epsilon_i + \epsilon_j + \epsilon_k) - (\Delta\epsilon_{ij} + \Delta\epsilon_{ik} + \Delta\epsilon_{jk}). \tag{2.154}$$

Higher order increments are described accordingly. The correlation energies of individual bodies, $\epsilon_i, \epsilon_{ij}, \cdots$ can be calculated using the local correlation methods, for example LMP2 or LCCSD, described in the previous subsection. After re-canonization of the LMOs within a given body, also canonical correlation methods, (MP2, CCSD(T)) can be applied.

Because in extended systems the number of increments increases rapidly, truncation of the expansion at some point becomes mandatory. The MoI expansion of the correlation energy for saturated systems usually converges at the three-body level. Additionally distance criteria may be used to discard bodies with local orbitals that are well separated. Developments of the MoI go into the following directions. Application of the MoI for non-metallic solids and molecular crystals has become feasible, by clever embedding schemes that mimic the crystal environment around the small correlated clusters cut out from the crystal [45, 75, 76]. The description of metals with the MoI still poses a challenge, because of the high degree of static correlation present. To tackle this problem multi-reference incremental schemes are actively developed. In multi-reference approaches, the bodies are augmented with localized virtual orbitals, and subjected to an complete active space calculation to recover the static correlation

energy [77, 78]. For molecules powerful automatic body generation and truncation schemes have been developed, that result in a basically black box variant of the MoI [46]. In this work the MoI was applied in conjunction with LMP2 and LCCSD(T0) for the determination of the correlation energy of alkane chain molecules in chapter 4.

2.2.3 Density functional theory

Wave function based electronic structure methods employ exact Hamiltonians and invoke approximations in the description of the wave function. The wave function can be systematically improved, in principle up to the known exact FCI form. The improvements come at ever increasing costs, as with every correction additional complicated terms arise, as demonstrated in the preceding sections on electron correlation methods. Hohenberg and Kohn [79] showed that the exact ground-state electronic energy is a functional $E[\rho]$ of the exact one electron density $\rho(\mathbf{r})$. Hence, the energy is directly connected to the density, which is a function of just three coordinates. In density functional theory (DFT) approximations are made to the unknown exact energy functional $E[\rho]$, which is equivalent to saying that approximate Hamiltonians are used. This is the inherent difference between DFT and wave function (WF) theory. The energy functional,

$$E[\rho] = T[\rho] + E_{\mathrm{ne}}[\rho] + E_{\mathrm{ee}}[\rho], \tag{2.155}$$

can be separated into functionals for the kinetic energy $T[\rho]$, the electron-nuclei attraction $E_{\mathrm{ne}}[\rho]$ and the electron-electron repulsion $E_{\mathrm{ee}}[\rho]$, as motivated by the form of the electronic Hamilton operator. The only term in the equation above that's known explicitly without further assumptions is the electron-nuclei attraction

$$E_{\mathrm{ne}}[\rho] = -\sum_a \frac{Z_a \rho(\mathbf{r})}{|\mathbf{R}_a - \mathbf{r}|}. \tag{2.156}$$

The $E_{\mathrm{ee}}[\rho]$ functional may be separated into a Coulomb $J[\rho]$ and an exchange-correlation term $K[\rho]$. The Coulomb term is also known as a functional of the density:

$$J[\rho] = \frac{1}{2} \int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{2.157}$$

The exact kinetic energy functional is known, if the density is expressed in terms of natural orbitals ψ_i^{NO} and occupation numbers n_i as

$$T[\rho_{\mathrm{exact}}] = \sum_{i=1}^{\infty} n_i \langle \psi_i^{\mathrm{NO}} | -\frac{1}{2}\nabla^2 | \psi_i^{\mathrm{NO}} \rangle, \tag{2.158}$$

where

$$\rho_{\mathrm{exact}} = \sum_{i=1}^{\infty} n_i |\psi_i^{\mathrm{NO}}|^2. \tag{2.159}$$

The natural orbitals are the eigenvectors of the exact density matrix. However, since the exact density matrix is unknown, a simpler representation of the density in terms of a non-interacting

system with orbitals ψ_i forming a Slater-determinant is given by

$$\rho = \sum_{i=1}^{N} |\psi_i|^2. \tag{2.160}$$

The kinetic energy functional for a density of a Slater-determinant is also known

$$T_S[\rho] = \sum_{i=1}^{N} \langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \rangle. \tag{2.161}$$

Hence, in Kohn-Sham (KS) theory [80] orbitals are introduced into DFT, in order to calculate $T_S[\rho]$, which is assumed to cover a large part of the true kinetic energy. Then, the DFT energy functional in the KS formulation is written as

$$E_{\mathrm{DFT}}[\rho] = T_S[\rho] + E_{\mathrm{ne}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho], \tag{2.162}$$

where the first three terms are known exactly for non-interacting electrons. Thus, all approximations with the goal to describe the electron-electron interaction are collected in the exchange-correlation functional $E_{\mathrm{xc}}[\rho]$. Although its name indicates otherwise, $E_{\mathrm{xc}}[\rho]$ not only describes exchange and correlation, but also the correlation correction ΔT to the kinetic energy. Analogous to HF theory, a set of orthogonal canonical Kohn-Sham orbitals that minimize $E_{\mathrm{DFT}}[\rho]$ are determined by a set of one-electron equations, called the Kohn-Sham equations,

$$\hat{h}_{\mathrm{KS}} \psi_i = \epsilon_i \psi_i, \tag{2.163}$$

where Kohn-Sham operator

$$\hat{h}_{\mathrm{KS}} = -\frac{1}{2}\nabla^2 + V_{\mathrm{eff}} \tag{2.164}$$

is written in terms of a kinetic energy operator and an effective potential

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{ne}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{\mathrm{xc}}(\mathbf{r}). \tag{2.165}$$

For molecular systems, the Kohn-Sham orbitals are expanded in a basis of AOs, which coefficients are determined variationally by solving the resulting Kohn-Sham matrix equations in a SCF procedure. The cost of KS-DFT calculations is at least similar and often lower compared to that of a HF calculation. This depends on the complexity of the integrals over the effective potential, and hence on the terms that enter the exchange correlation functional.

The explicit form of the exchange correlation functional determines the type of DFT method. To describe the functional in terms of quantities familiar from wave function theory, it is commonly separated into

$$E_{\mathrm{xc}}[\rho] = E_{\mathrm{x}}[\rho] + E_{\mathrm{c}}[\rho], \tag{2.166}$$

a pure exchange part $E_{\mathrm{x}}[\rho]$ and a pure correlation part $E_{\mathrm{c}}[\rho]$. Here, it is important to note, that the exchange functional and the correlation functional do not represent exchange and correlation energies as defined in wave function theory. Both functionals can be expressed in

terms of energy densities ϵ_{x} and ϵ_{c} as

$$E_{\mathrm{x}}[\rho] = \int \rho(\mathbf{r})\epsilon_{\mathrm{x}}[\rho(\mathbf{r})]d\mathbf{r}, \tag{2.167}$$

$$E_{\mathrm{c}}[\rho] = \int \rho(\mathbf{r})\epsilon_{\mathrm{c}}[\rho(\mathbf{r})]d\mathbf{r}, \tag{2.168}$$

$$\epsilon_{\mathrm{xc}} = \epsilon_{\mathrm{x}} + \epsilon_{\mathrm{c}}. \tag{2.169}$$

Energy densities are used in the definition of the exchange correlation potential as

$$V_{\mathrm{xc}} = \frac{\partial E_{\mathrm{xc}}}{\partial \rho(\mathbf{r})} = \epsilon_{\mathrm{xc}}[\rho(\mathbf{r})] + \rho(\mathbf{r})\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho(\mathbf{r})}. \tag{2.170}$$

The exchange and correlation between electrons of different spin is handled by the dependence of E_{xc} on spin-densities ρ^{α} and ρ^{β} or alternatively on the spin-polarization ζ given by

$$\zeta = \frac{\rho^{\alpha} - \rho^{\beta}}{\rho^{\alpha} + \rho^{\beta}} \tag{2.171}$$

and the radius of the effective volume containing one electron

$$r_{\mathrm{S}} = \sqrt[3]{\frac{3}{4\pi\rho}}. \tag{2.172}$$

The different approximations to $E_{\mathrm{xc}}[\rho]$ fall into the following groups, local density approximations (LDA), generalized gradient approximation (GGA), meta-GGA, hybrid functionals and double hybrid functionals. [52]

2.2.3.1 LDA functionals

The LDA builds on the description of the density as a uniform electron gas. The exchange energy of a uniform electron gas is given by the Dirac formula in terms of local spin densities as

$$\epsilon_{\mathrm{x}}^{\mathrm{LSDA}}[\rho] = -\sqrt[3]{2}C_{\mathrm{x}}[\sqrt[3]{\rho_{\alpha}} + \sqrt[3]{\rho_{\alpha}}] = -\frac{1}{2}C_{\mathrm{x}}\sqrt[3]{\rho}[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}]. \tag{2.173}$$

The functional for the correlation energy density of a uniform electron gas is given by the analytical formula found by Vosko, Wilk and Nusair (VWN) [81] as

$$\epsilon_{\mathrm{c}}^{\mathrm{VWN}}(r_{\mathrm{S}},\zeta) = \epsilon_{\mathrm{c}}(r_{\mathrm{S}},0) + \epsilon_{\mathrm{a}}(r_{\mathrm{S}})\left[\frac{f(\zeta)}{f''(0)}\right][1-\zeta^4] + [\epsilon_{\mathrm{c}}(r_{\mathrm{S}},1) - \epsilon_{\mathrm{c}}(r_{\mathrm{S}},0)]f(\zeta)\zeta^4, \tag{2.174}$$

with

$$f(\zeta) = \frac{(1-\zeta)^{4/3} + (1-\zeta)^{4/3} - 2}{2(2^{1/3}-1)} \tag{2.175}$$

and

$$\epsilon_{\mathrm{c/a}}(x) = A\begin{bmatrix} \ln\left(\frac{x^2}{X(x)}\right) + \frac{2b}{Q}\tan^{-1}\left(\frac{Q}{2x+b}\right) \\ -\frac{bx_0}{X(x_0)}\left[\ln\left(\frac{(x-x_0)^2}{X(x)}\right) + \frac{2(b+2x_0)}{Q}\tan^{-1}\left(\frac{Q}{2x+b}\right)\right] \end{bmatrix} \tag{2.176}$$

$$x = \sqrt{r_{\mathrm{S}}} \tag{2.177}$$

$$X(x) = x^2 + bx + c \tag{2.178}$$

$$Q = \sqrt{4c - b^2}. \tag{2.179}$$

The unpolarized $\epsilon_{\mathrm{c}}(r_{\mathrm{S}}, 0)$ and the polarized $\epsilon_{\mathrm{c}}(r_{\mathrm{S}}, 1)$ correlation energy densities and $\epsilon_{\mathrm{a}}(r_{\mathrm{S}})$ differ in the fitting parameters A, x_0, b and c.

LDA underestimates the exchange energy and overestimates the correlation energy, which leads to a general overestimation of bond strengths. [52]

2.2.3.2 GGA functionals

The LDA functionals depend only on the local density, one way to improve on this, is to also take into account local gradients of the density $\nabla\rho$, which is known as the generalized gradient approximation (GGA) [52]. Several gradient corrected exchange functionals have been proposed. The exchange functional PW86 introduced by Perdew and Wang [82] is given by,

$$\epsilon_{\mathrm{x}}^{\mathrm{PW86}} = \epsilon_{\mathrm{x}}^{\mathrm{LDA}}(1 + ax^2 + bx^4 + cx^6)^{1/15}, \tag{2.180}$$

where x depends on the gradient,

$$x = \frac{|\nabla\rho|}{\rho^{4/3}}, \tag{2.181}$$

and a, b and c are constants. Beckes functional B88 is a correction to the LDA exchange,

$$\epsilon_{\mathrm{x}}^{\mathrm{B88}} = \epsilon_{\mathrm{x}}^{\mathrm{LDA}} + \delta\epsilon_{\mathrm{x}}^{\mathrm{B88}} \tag{2.182}$$

$$\delta\epsilon_{\mathrm{x}}^{\mathrm{B88}} = -\beta\rho^{1/3}\frac{x^2}{1 + 6\beta x \sinh^{-1}(x)}, \tag{2.183}$$

where β is a constant and x is the same gradient variable as in PW86. Another GGA exchange functional is PW91 by Perdew and Wang [83] given by,

$$\epsilon_{\mathrm{x}}^{\mathrm{PW91}} = \epsilon_{\mathrm{x}}^{\mathrm{LDA}}\left(\frac{1 + xa_1\sinh^{-1}(xa_2) + (a_3 + a_4 e^{-bx^2})x^2}{1 + xa_1\sinh^{-1}(xa_2) + a_5 x^2}\right), \tag{2.184}$$

where a_i and b are constants, and x is defined as in PW86. Together with this exchange functional, the Perdew and Wang PW91 correlation functional is used.

Lee, Yang and Paar proposed the GGA correlation functional LYP [84] as,

$$\begin{aligned}\epsilon_c^{\mathrm{LYP}} = &-a\frac{\gamma}{(1 + d\rho^{-1/3})} - ab\frac{\gamma e^{-c\rho^{-1/3}}}{9\left(1 + d\rho^{-1/3}\rho^{8/3}\right)} \\ &\times\left[\begin{array}{c} 18(2^{2/3})C_F(\rho_\alpha^{8/3} + \rho_\beta^{8/3}) - 18\rho t_W \\ +\rho_\alpha\left(2t_W^\alpha + \nabla^2\rho_\alpha\right) + \rho_\beta\left(2t_W^\beta + \nabla^2\rho_\beta\right)\end{array}\right],\end{aligned} \tag{2.185}$$

$$\gamma = 2\left[1 - \frac{\rho_\alpha^2 + \rho_\beta^2}{\rho^2}\right], \tag{2.186}$$

$$t_W^\sigma = \frac{1}{8}\left(\frac{|\nabla\rho_\sigma|^2}{\rho_\sigma} - \nabla^2\rho_\sigma\right). \tag{2.187}$$

The parameters a, b, c and d are fitted to exact data from the helium atom and t_W^σ denotes the local Weizäcker kinetic energy density. [52]

The functional by Perdew, Burke and Enzerhof (PBE) is a correction to LSDA, in which all parameters are fundamental constants [85]. PBE is designed as a simplification of the PW91 functional. Both are analytical functions that describe the numerical GGA for the exchange correlation hole around an electron in an uniform electron gas. Both functionals essentially give the same results. The PBE correlation functional has the following form,

$$\epsilon_{\mathrm{c}}^{\mathrm{PBE}} = \epsilon_{\mathrm{c}}^{\mathrm{LSDA}}(r_{\mathrm{S}}, \zeta) + H(r_{\mathrm{S}}, \zeta, x) \tag{2.188}$$

where the dimensionless density gradient x is given by,

$$x = \frac{|\nabla\rho|}{2\phi_\zeta k_s \rho}, \tag{2.189}$$

$$\phi_\zeta(\zeta) = \frac{(1+\zeta)^{2/3} + (1-\zeta)^{2/3}}{2}, \tag{2.190}$$

$$k_s = \sqrt{4k_F/\pi}, \tag{2.191}$$

which depends on a spin-scaling factor ϕ_ζ and the Thomas-Fermi screening wave number k_s, with the Fermi wave number given by $\rho = k_F^3/3\pi^2$. The form of the gradient correction H is,

$$H = \gamma\phi_\zeta^3 \times \ln\left(1 + \frac{\beta}{\gamma}x^2\left[\frac{1+Ax^2}{1+Ax^2+A^2x^4}\right]\right), \tag{2.192}$$

$$A = \frac{\beta}{\gamma}\left[\exp\left(-\frac{\epsilon_{\mathrm{c}}^{\mathrm{LSDA}}}{(\gamma\phi_\zeta^3)}\right) - 1\right]^{-1}. \tag{2.193}$$

The correlation functional and the parameters β and γ are determined from the behaviour of the numerical GGA in certain limits. The PBE exchange functional is given by,

$$\epsilon_{\mathrm{x}}^{\mathrm{PBE}} = \epsilon_{\mathrm{x}}^{\mathrm{LSDA}} F_{\mathrm{x}}(s), \tag{2.194}$$

$$F_{\mathrm{x}}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa}, \tag{2.195}$$

$$s = \frac{|\nabla\rho|}{2k_F\rho}, \tag{2.196}$$

$$\mu = \beta(\pi^2/3), \tag{2.197}$$

$$\kappa = 0.804. \tag{2.198}$$

GGA approaches are considered as the second rung of the "Jacobs ladder" of DFT, while LDA constitutes the first rung. The third rung are the so called meta-GGA functionals, which are described in the next part.

2.2.3.3 Meta-GGA functionals

The family of meta-GGA functionals depend like GGAs on the density ρ, the local gradient of the density $\nabla\rho$, and additionally on the kinetic energy density τ [86]. The kinetic energy density is given by,

$$\tau_\sigma(\mathbf{r}) = \sum_i^{\text{occ}} \frac{1}{2}|\nabla\psi_{i\sigma}(\mathbf{r})|^2, \tag{2.199}$$

and depends on the occupied Kohn-Sham orbitals $\psi_{i\sigma}$. The meta-GGA functional by Tao, Perdew, Saroverov and Scuseria (TPSS) contains no empirical parameters and is an improvement upon the Perdew-Kurth-Zupan-Blaha (PKZB) meta-GGA functional [86]. TPSS improves upon the PBE functional, by additionally considering the kinetic energy density. The TPSS exchange scaling $F_\text{x}(p,z)$ is given by,

$$\epsilon_\text{x}^\text{TPSS} = \epsilon_\text{x}^\text{LSDA} \cdot F_\text{x}(p,z), \tag{2.200}$$

$$F_\text{x}(p,z) = 1 + \kappa - \frac{\kappa}{1 + x(p,z)/\kappa}, \tag{2.201}$$

$$p = s^2 = \frac{|\nabla\rho|^2}{4(2\pi^2)^{2/3}\rho^{8/3}}, \tag{2.202}$$

$$z = \frac{\tau^W}{\tau}, \tag{2.203}$$

$$\tau^W = \sum_\sigma \tau_\sigma, \tag{2.204}$$

$$x(p,z) = \left\{ \left[\frac{10}{81} + c\frac{z^2}{(1+z^2)^2}\right] p + \frac{146}{2025}\tilde{q}_b^2 \right. \tag{2.205}$$

$$-\frac{73}{405}\tilde{q}_b\sqrt{\frac{1}{2}\left(\frac{3}{5}z\right)^2 + \frac{1}{2}p^2} + \frac{1}{\kappa}\left(\frac{10}{81}\right)^2 p^2 \tag{2.206}$$

$$\left. +2\sqrt{e}\frac{10}{81}\left(\frac{3}{5}z\right)^2 + e\mu p^3 \right\} / (1 + \sqrt{d}p)^2,$$

$$\tilde{q}_b = (9/20)(\alpha - 1)/[1 + b\alpha(\alpha - 1)]^{1/2} + 2p/3, \tag{2.207}$$

$$\alpha = \frac{\tau - \tau^W}{\tau^\text{unif}} = (5p/3)(z^{-1} - 1), \tag{2.208}$$

$$\tau^\text{unif} = \frac{3}{10}(2\pi^2)^{2/3}\rho^{5/3}. \tag{2.209}$$

The functional form of $F_\text{x}(p,z)$ ensures, that the exchange potential has finite values at the nuclei, for ground-state one and two-electron densities. The constants c and e are thus chosen such, that for the exact value of s for an exponential two-electron density ($z = 1, \alpha = 0$) at a nucleus $dF_\text{x}(s^2,1)/ds|_{s=0.376} = 0$ holds. The TPSS correlation functional is very similar to the

respective PKZB functional,

$$\epsilon_{\mathrm{c}}^{\mathrm{TPSS}} = \epsilon_{\mathrm{c}}^{\mathrm{revPKZB}} \left[1 + d\epsilon_{\mathrm{c}}^{\mathrm{revPKZB}} (\tau^W/\tau)^3\right], \tag{2.210}$$

$$\epsilon_{\mathrm{c}}^{\mathrm{revPKZB}} = \epsilon_{\mathrm{c}}^{\mathrm{PBE}} \left[1 + C(\zeta,\xi)(\tau^W/\tau)^2\right] \tag{2.211}$$

$$- \left[1 + C(\zeta,\xi)\right] \sum_\sigma \frac{\rho_\sigma}{\rho} \tilde{\epsilon}_{\mathrm{c}}. \tag{2.212}$$

The term $\tilde{\epsilon}_{\mathrm{c}}$ is given by $\tilde{\epsilon}_{\mathrm{c}} = \max\left[\epsilon_{\mathrm{c}}^{\mathrm{PBE}}(\rho_\sigma, 0, \nabla\rho_\sigma, 0), \epsilon_{\mathrm{c}}^{\mathrm{PBE}}(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)\right]$ to ensure $E_{\mathrm{c}} \leq 0$ for any ρ. The function C has the form,

$$C(\zeta,\xi) = \frac{C(\zeta,0)}{\left(1 + \xi^2\left[(1+\zeta)^{-4/3} + (1-\zeta)^{-4/3}\right]/2\right)^4}, \tag{2.213}$$

$$C(\zeta,0) = 0.53 + 0.87\zeta^2 + 0.50\zeta^4 + 2.26\zeta^6, \tag{2.214}$$

$$\xi = \frac{|\nabla\zeta|}{2(3\pi^2\rho)^{1/3}}. \tag{2.215}$$

The TPSS correlation functional eliminates a self-correlation error for spin-unpolarized densities, which is present in LSDA and GGA. Thereby, correlation energies for respective atomic densities are improved considerably over the LDA, and GGA methods. By construction all correct features of LSDA and PBE GGA are retained in TPSS. Thus, the functionals, PBE and TPSS can be seen as systematic non-empirical improvements over the LSDA description [86].

2.2.3.4 Hybrid functionals

Hybrid functionals include the exact exchange of a Slater-determinant formed by Kohn-Sham orbitals. This approach is justified by the Adiabatic Connection Formula (ACF), which couples the exchange correlation energy of the non-interacting reference system to the interacting many electron system [52]. The electron-electron interaction is scaled by the parameter λ and the exchange correlation energy can be expressed by an integral over λ,

$$E_{\mathrm{xc}} = \int_0^1 \langle\Psi_\lambda|\mathbf{V}_{\mathrm{xc}}(\lambda)|\Psi_\lambda\rangle d\lambda, \tag{2.216}$$

that connects the limits of non-interacting ($\lambda = 0$) and fully interacting ($\lambda = 1$) electrons. A simple approximation to this integral is a linear interpolation between the two limits,

$$E_{\mathrm{xc}} \approx \frac{1}{2}\langle\Psi_0|\mathbf{V}_{xc}(0)|\Psi_0\rangle + \frac{1}{2}\langle\Psi_1|\mathbf{V}_{xc}(\lambda)|\Psi_1\rangle. \tag{2.217}$$

The first term for non-interacting electrons is correctly described by the exchange energy of a Slater-determinant with KS orbitals, which is the exact wave function for $\lambda = 0$. The second term, for fully interacting electrons, is unknown and has to be approximated, for example by LSDA. Which gives the half-and-half (H+H) method [87],

$$E_{\mathrm{xc}}^{\mathrm{H+H}} = \frac{1}{2}E_{\mathrm{x}}^{\mathrm{exact}} + \frac{1}{2}E_{\mathrm{x}}^{\mathrm{LSDA}} + E_{\mathrm{c}}^{\mathrm{LSDA}}. \tag{2.218}$$

This method can of course be extended, by the introduction of GGA correction terms as in Beckes three parameter functional (B3) [88],

$$E_{\mathrm{xc}}^{\mathrm{B3}} = (1-a)E_{\mathrm{x}}^{\mathrm{LSDA}} + aE_{\mathrm{x}}^{\mathrm{exact}} + b\Delta E_{\mathrm{x}}^{\mathrm{B88}} + E_{\mathrm{c}}^{\mathrm{LSDA}} + c\Delta E_{\mathrm{c}}^{\mathrm{GGA}} \tag{2.219}$$

where the parameters $a = 0.2$, $b = 0.72$ and $c = 0.81$ are determined by fitting to empirical (experimental) and ab inito data, and can vary depending on the actual GGA correlation correction. The GGA correlation correction is given in the name of the B3 methods, prominent examples are B3LYP and B3PW91 [88]. The formula for the meta-GGA hybrid functional TPSSh [89] is straightforward,

$$E_{\mathrm{xc}}^{\mathrm{TPSSh}} = (1-a)E_{\mathrm{x}}^{\mathrm{TPSS}} + aE_{\mathrm{x}}^{\mathrm{exact}} + E_{\mathrm{c}}^{\mathrm{TPSS}}, \tag{2.220}$$

where the empirical parameter $a = 0.10$ is determined by minimizing the mean absolute deviations of enthalpies of formations of a test set of molecules.

In double hybrid functionals [90, 91], in addition to the hybrid description of E_{x} by the exact exchange, a correlation energy $E_{\mathrm{c}}^{\mathrm{PT2}}$ obtained from second order perturbation theory is mixed to the GGA correlation energy,

$$E_{\mathrm{xc}}^{\mathrm{DH}} = (1-a_{\mathrm{x}})E_{\mathrm{x}}^{\mathrm{GGA}} + a_{\mathrm{x}}E_{\mathrm{x}}^{\mathrm{exact}} + (1-a_{\mathrm{c}})E_{\mathrm{c}}^{\mathrm{GGA}} + a_{\mathrm{c}}E_{\mathrm{x}}^{\mathrm{PT2}}. \tag{2.221}$$

$E_{\mathrm{c}}^{\mathrm{PT2}}$ is calculated using the formula for MP2 with KS orbitals.

2.2.3.5 Dispersion correction methods

Standard DFT methods like LDA, GGA, meta-GGA and hybrids, as well as the HF WF method, fail to describe London dispersion interactions. The reason for this is, that dispersion is a non-local dynamic correlation effect, which is impossible to describe in form of a local correlation functional, or with an uncorrelated wave function (HF). Dispersion is accounted for in MP2 and CC theory. Hence, double hybrids underestimate dispersion interactions, which are mixed in only by the fraction of $E_{\mathrm{c}}^{\mathrm{PT2}}$ present in the correlation functional. To extend the application of DFT methods and increase their accuracy, dispersion corrections are added to standard DFT results. In the following, three distinct schemes for dispersion corrections are introduced. All three schemes follow different approaches. The D3 method [9] is based on pre-calculated reference dispersion coefficients and uses only the nuclear geometry of the systems as input. The vdW-QHO-WF method [92] on the other hand, describes the dispersion interactions of localized KS orbitals using a coupled system of quantum harmonic oscillators. At last VV10 [38] is a non-local functional of the density, that is able to describe dispersion interactions.

DFT-D3

One of the most widely used dispersion corrections is the DFT-D3 method by Grimme [9]. D3 is an "on top" dispersion correction E_{disp}, that is subtracted from the uncorrected KS-DFT energy $E_{\text{KS-DFT}}$ to yield the total DFT-D3 energy

$$E_{\text{DFT-D3}} = E_{\text{KS-DFT}} - E_{\mathrm{disp}}. \tag{2.222}$$

The dispersion energy is given by the sum of two and three body terms,

$$E_{\text{disp}} = E^{(2)} + E^{(3)}, \tag{2.223}$$

where the two body dispersion, between pairs of atoms, typically accounts for at least 90% of the total dispersion energy. This term is given by the following sum over all atom pairs AB by,

$$E^{(2)} = \sum_{\text{AB}} \sum_{n=6,8,10,\cdots} s_n \frac{C_n^{\text{AB}}}{r_{\text{AB}}^n} f_{d,n}(r_{\text{AB}}), \tag{2.224}$$

where the second sum runs over even orders of n, beginning with $n = 6$, the isotropic nth-order dispersion coefficient, specific for the atomic pair AB, is denoted C_n^{AB}, the internuclear distance is given by r_{AB}. The global scaling factors s_n for $n > 6$ adjust the dispersion correction to a given DFT functional. The range of the dispersion correction is controlled by a damping function $f_{d,n}(r_{\text{AB}})$. [9]

The 6th-order dispersion coefficients are given by the Casimir-Polder formula,

$$C_6^{\text{AB}} = \frac{3}{\pi} \int_0^\infty \alpha^{\text{A}}(i\omega)\alpha^{\text{B}}(i\omega) d\omega \tag{2.225}$$

where $\alpha^{\text{A}}(i\omega)$ is the averaged dipole polarizability of atom A at imaginary frequency $i\omega$. The polarizabilities are calculated from *ab initio* using time dependent DFT (TD-DFT). Atoms in molecules have lower polarizabilities due to bond formation, when compared to free atoms. Therefore, polarizabilities are calculated for hydride reference molecules $\text{A}_\text{m}\text{H}_\text{n}$ and $\text{B}_\text{k}\text{H}_\text{l}$. The polarizability for A bound in the hydride is then calculated by subtraction of half the polarizability for H_2 per hydrogen in the hydride and division by m. The formula for the dispersion coefficients then becomes,

$$C_6^{\text{AB}} = \frac{3}{\pi} \int_0^\infty \frac{1}{\text{m}} \left[\alpha^{\text{A}_\text{m}\text{H}_\text{n}}(i\omega) - \frac{n}{2}\alpha^{\text{H}_2}(i\omega)\right] \times \frac{1}{\text{k}} \left[\alpha^{\text{B}_\text{k}\text{H}_\text{l}}(i\omega) - \frac{l}{2}\alpha^{\text{H}_2}(i\omega)\right] d\omega. \tag{2.226}$$

This approach yields dispersion coefficients that depend on the coordination number CN. Hence, a to each atom in the actual structure a coordination number CN^{A} has to be assigned. In D3, this is accomplished by computing a fractional coordination number for each atom, based on the sum of tabulated covalent radii and the internuclear distance via the "counting" function

$$\text{CN}^{\text{A}} = \sum_{\text{B}\neq\text{A}}^{N_{\text{at}}} \frac{1}{1 + e^{-k_1(k_2(R_{\text{A,cov}}+R_{\text{B,cov}})/r_{\text{AB}-1})}}. \tag{2.227}$$

The covalent radii $R_{\text{A,cov}}$ are scaled by the factor $k_2 = 4/3$. The parameter $k_1 = 16$ ensures that distant atoms do not artificially increase the CN in large molecules and that coordination number contributions for stretched C-C single bonds approach 0 within the bond breaking range of 1.5-3 Å. For organic molecules the fractional coordination numbers are close to the expected values. For metallic systems the CN are too high, to balance this the radii of metal atoms are decreased by 10%. [9]

For each element polarizabilities are calculated, for the bare atom and reference hydride systems with different CN. Then a data base of reference dispersion coefficients $C_{6,\text{ref}}^{\text{AB}}(\text{CN}^{\text{A}}, \text{CN}^{\text{B}})$ for the atom pairs in a given coordination is created using eq. 2.226. The actual dispersion

coefficient is determined by interpolating between reference coefficients using the following formula,

$$C_6^{\mathrm{AB}}(\mathrm{CN}^{\mathrm{A}}, \mathrm{CN}^{\mathrm{B}}) = \frac{Z}{W}, \tag{2.228}$$

$$Z = \sum_i^{N_{\mathrm{A}}} \sum_j^{N_{\mathrm{B}}} C_{6,\mathrm{ref}}^{\mathrm{AB}}(\mathrm{CN}^{\mathrm{A}}, \mathrm{CN}^{\mathrm{B}}) L_{ij}, \tag{2.229}$$

$$W = \sum_i^{N_{\mathrm{A}}} \sum_j^{N_{\mathrm{B}}} L_{ij}, \tag{2.230}$$

$$L_{ij} = e^{-k_3[(\mathrm{CN}^{\mathrm{A}} - \mathrm{CN}_i^{\mathrm{A}})^2 + (\mathrm{CN}^{\mathrm{B}} - \mathrm{CN}_j^{\mathrm{B}})^2]}, \tag{2.231}$$

where the number of reference coordinations per atom is given by N_{A} and N_{B}. The reference C_6 are weighted by the Gaussian distance L between actual and reference CNs. L is dependent on parameter $k_3 = 4$. The value for k_3 was chosen such, that the interpolation is smooth. [9]

Higher order dispersion coefficients than C_6 are calculated recursively. In practice only C_8 is included,

$$C_8^{\mathrm{AB}} = 3 C_6^{\mathrm{AB}} \sqrt{Q^A Q^B}, \tag{2.232}$$

$$Q^A = s_{42} \sqrt{Z^{\mathrm{A}}} \frac{\langle r^4 \rangle^{\mathrm{A}}}{\langle r^2 \rangle^{\mathrm{A}}}, \tag{2.233}$$

where $\langle r^2 \rangle$ and $\langle r^4 \rangle$ are derived from averaged atomic densities. [9]

The three-body term takes into account the Axilrod-Teller-Muto dispersion,

$$E^{\mathrm{ABC}} = \frac{C_9^{\mathrm{ABC}} 3 \cos\theta_a \cos\theta_b \cos\theta_c}{(r_{\mathrm{AB}} r_{\mathrm{BC}} r_{\mathrm{CA}})^3}, \tag{2.234}$$

where θ_a, θ_b, θ_c and r_{AB}, r_{BC}, r_{CA} define the triangle of three atoms A, B, C, with triple-dipole constant C_9^{ABC}. The triple-dipole constant is approximated by,

$$C_9^{\mathrm{ABC}} \approx -\sqrt{C_6^{\mathrm{AB}} C_6^{\mathrm{AC}} C_6^{\mathrm{BC}}}. \tag{2.235}$$

The total three body contribution is given by a sum over all E^{ABC}, where each is damped by the function $f_{d,(3)}$ [9],

$$E^{(3)} = \sum_{\mathrm{ABC}} f_{d,(3)}(\bar{r}_{\mathrm{ABC}}) E^{\mathrm{ABC}}. \tag{2.236}$$

Gimme *et al.* investigated the effect of the damping function and compared the effect of different damping schemes [36]. Damping of the dispersion correction at $r_{\mathrm{AB}} = 0$ to zero, so called zero damping is achieved by the following damping function,

$$f_{\mathrm{damp}}(r_{\mathrm{AB}}) = \frac{1}{1 + 6(r_{\mathrm{AB}}/(s_{r,n} r_0^{\mathrm{AB}}))^{-\gamma}}, \tag{2.237}$$

where r_0^{AB} is a cut-off radius, determined from averaged vdW-radii, and $s_{r,n}$ is a scaling factor. A complementary damping scheme, where the dispersion correction at $r_{\mathrm{AB}} = 0$ is damped up

to a constant value is the Becke-Johnson (BJ) damping given by,

$$E_{\text{disp}}^{\text{D3(BJ)}} = -\frac{1}{2}\sum_{\text{A}\neq\text{B}} s_6 \frac{C_6^{\text{AB}}}{R_{\text{AB}}^6 + [f(R_{\text{AB}}^0)]^6} + s_8 \frac{C_8^{\text{AB}}}{R_{\text{AB}}^8 + [f(R_{\text{AB}}^0)]^8}, \tag{2.238}$$

with

$$f(R_{\text{AB}}^0) = a_1 R_{\text{AB}}^0 + a_2. \tag{2.239}$$

BJ-damping introduces the fit parameters a_1 and a_2, where R_{AB}^0 is different from the cut-off radius used in zero damping,

$$R_{\text{AB}}^{0,(\text{BJ})} = \sqrt{\frac{C_8^{\text{AB}}}{C_6^{\text{AB}}}}. \tag{2.240}$$

The comparison between zero and BJ damping showed, that the overall effect of the damping scheme is small, but that BJ damping improves on non-covalent equilibrium distances. It is recommended to use the BJ damping scheme as the default in DFT-D3. [36]

The advantages of the DFT-D3 method are low empiricism, as most parameters are calculated from first principles, it has the correct asymptotic r^{-6} behaviour for molecules, it describes many relevant elements of the periodic table ($Z = 1 - 94$), CN dependent dispersion coefficients are determined from the nuclear structure of the molecules, it provides good accuracy, it is very fast and can be used with many standard density functional methods. [9]

Quantum harmonic oscillator with maximally localized Wannier functions method

In the Quantum Harmonic Oscillator (QHO) model, dispersion is described by a coupled system of QHOs [39, 92, 93]. The Hamiltonian of N isotropic coupled 3-dimensional QHOs is given by,

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_{\chi_i}^2 + \frac{1}{2}\sum_{i=1}^{N} \omega_i^2 \chi_i^2 + \sum_{i>j=1}^{N} \omega_i \omega_j \sqrt{\alpha_i \alpha_j}\, \chi_i T_{ij} \chi_j, \tag{2.241}$$

where $\chi_i = \sqrt{m_i}\zeta_i$ denote mass weighted displacement coordinates, for mass m_i and displacement from equilibrium ζ_i. The characteristic frequency of the QHO i is given by ω_i, the polarizability is denoted α_i and the dipole-dipole interaction tensor is given by T_{ij}. The sums in the QHO Hamiltonian describe, from the left to the right, the kinetic energy, the potential energy and the interaction energy between the QHOs. The interaction tensor is given by,

$$T_{ij}^{ab} = -\frac{3r_{ij}^a r_{ij}^b - r_{ij}^2 \delta_{ab}}{r_{ij}^5} \left(\text{erf}\left(\frac{r_{ij}}{\sigma_{ij}}\right) - \frac{2}{\sqrt{\pi}} \frac{r_{ij}}{\sigma_{ij}} e^{-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^2} \right) + \frac{4}{\sqrt{\pi}} \frac{1}{\sigma_{ij}^3} \frac{r_{ij}^a r_{ij}^b}{r_{ij}^2} e^{-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^2}, \tag{2.242}$$

where the distance between two QHOs i and j is denoted r_{ij}, and their effective width is given by σ_{ij}. The Gauß-error function is part of a damping function $V_{ij} = \text{erf}\left(\frac{r_{ij}}{\sigma_{ij}}\right)/r_{ij}$. The energy of the interacting QHOs is obtained by diagonalization of the $3N \times 3N$ matrix $\mathbf{C}^{\text{QHO}}$, with N^2 3×3 blocks $\mathbf{C}_{ij}^{\text{QHO}}$ for each QHO pair,

$$\mathbf{C}_{ij}^{\text{QHO}} = \omega_i^2 \mathbf{1} \tag{2.243}$$

$$\mathbf{C}_{i\neq j}^{\text{QHO}} = \omega_i \omega_j \sqrt{\alpha_i \alpha_j} T_{ij}. \tag{2.244}$$

$$\tag{2.245}$$

Silvestrelli proposed to parametrize the QHO model dependent on the localized orbitals of a molecule, *i.e.* localized Wannier functions in periodic systems [92]. This model is termed DFT/vdW-QHO-WF. Maximally localized Wannier functions are obtained by unitary transformation of the KS orbitals under the constraint to minimize the spread S_i. To each local orbital center one QHO is assigned, which properties are derived from the spread of the local orbital. The effective width is then derived from the spreads,

$$\sigma_{ij} = \beta\sqrt{S_i^2 + S_j^2}, \tag{2.246}$$

as is the polarizability

$$\alpha_i = \gamma S_i^3, \tag{2.247}$$

and the characteristic frequency

$$\omega_i = \sqrt{\zeta n_i/\alpha_i}. \tag{2.248}$$

Beside the spread also the number of electrons in the LMO $n_i = 2$ and global parameters β, γ and ζ are considered. The global parameters are determined by minimizing the mean absolute relative error of PBE/vdW-QHO-WF with respect to the CCSD(T) reference data of the S22 test set. The dispersion correction is given by the correlation energy of the QHO system,

$$E_{c,\mathrm{QHO}} = \frac{1}{2}\sum_{p=1}^{3N}\sqrt{\lambda_p} - \frac{3}{2}\sum_{i=1}^{N}\omega_i, \tag{2.249}$$

where the first sum goes over the eigenvalues of the coupled QHO system and the second sum over the uncoupled, *i.e.* uncorrelated, eigenvalues of the QHOs. In the most recent parametrization the following parameters have been obtained: $\beta = 1.65$, $\gamma = 0.88$ and $\zeta = 1.30$ [39].

The advantage of the vdW-QHO-WF method is that the dispersion model is obtained directly from the electronic structure, *i.e.* the spreads of the Wannier functions. Thus, no atom specific reference calculations have to be performed. It was noticed, that the spread functional can be very sensitive to simulation box sizes. In periodic calculations one should therefore converge the dispersion correction with respect to the box sizes [39].

The nonlocal VV10 density functional method

Vydrov and Van Voorhis developed a nonlocal correlation functional (VV10) that is able to describe dispersion interactions using only the electronic density as an input [38]. The nonlocal VV10 functional is used in combination with standard semi-local exchange and correlation functionals. The nonlocal part of the correlation energy is given by,

$$E_{\mathrm{c}}^{\mathrm{nl}} = \frac{\hbar}{2}\int\int d\mathbf{r}d\mathbf{r}'\rho(\mathbf{r})\epsilon(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}') \tag{2.250}$$

where the kernel is,

$$\epsilon(\mathbf{r},\mathbf{r}') = -\frac{3e^4}{2m^2gg'(g+g')}, \tag{2.251}$$

with

$$g = \omega_0(\mathbf{r})R^2 + \kappa(\mathbf{r}) \tag{2.252}$$

$$g' = \omega_0(\mathbf{r}')R^2 + \kappa(\mathbf{r}') \tag{2.253}$$

$$R = |\mathbf{r} - \mathbf{r}'| \tag{2.254}$$

$$\omega_0 = \sqrt{\omega_g^2 + \frac{\omega_p^2}{3}}, \tag{2.255}$$

where the quantity ω_p is called the local plasma frequency which depends on the density, ω_g denotes the local band gap, which depends also on the gradient of the density, the equations for both are,

$$\omega_p^2 = 4\pi\rho e^2/m, \tag{2.256}$$

$$\omega_g^2 = C\frac{\hbar^2}{m^2}\left|\frac{\nabla\rho}{\rho}\right|^4. \tag{2.257}$$

The term κ is given by,

$$\kappa = b\frac{\nu_{\mathrm{F}}^2}{\omega_p} \tag{2.258}$$

where $\nu_{\mathrm{F}} = (2\pi\rho)^{1/3}\hbar/m$ is the local Fermi velocity, and b is a damping parameter. The VV10 functional is defined as

$$E_{\mathrm{c}}^{\mathrm{VV10}} = E_{\mathrm{c}}^{\mathrm{nl}} + \beta N, \tag{2.259}$$

where β is the constant absolute value of the nonlocal correlation energy density of the uniform electron gas and N is the number of electrons. It can be shown, that β depends only on fundamental constants and on b. Thus only two adjustable parameters C and b have to be determined. C was optimized to reduce the average error of C_6 coefficients and the damping parameter b was fitted using the S22 set of binding energies. The final complete exchange correlation functional is given by,

$$E_{\mathrm{xc}}^{\text{DFT-VV10}} = E_{\mathrm{xc}}^{\mathrm{DFT}} + E_{\mathrm{c}}^{\mathrm{VV10}}, \tag{2.260}$$

where $E_{\mathrm{xc}}^{\mathrm{DFT}}$ is a semi-local DFT functional, for example TPSS. For each functional the damping parameter b has to be adjusted individually. The advantage of VV10 is, that no atoms have to be identified, as in D3, nor local orbitals have to be constructed, as in the vDW-QHO-WF. Comparisons between D3 and VV10 show, that both methods are similar in accuracy in terms of interaction energies and bond lenghts [94, 95].

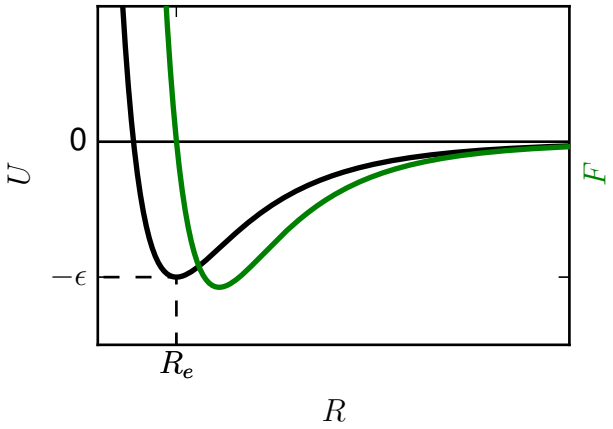


Figure 2.1: Potential (black) and the resulting force (green) between two molecules.

2.3 Intermolecular interactions

The different phases of matter, *i.e.* the solid, liquid and gas phase, and the transitions between them can be described on the basis of the forces acting between the molecules. Attractive forces between molecules describe the transition from the gas to the condensed phases, *i.e.* liquid or solid. This force has to cease for molecules that are distant from each other, in order to explain the occurrence of the gas phase, where the molecules move freely. The low compressibility of condensed phases, is evidence for a strongly repulsive force acting between molecules at short ranges. For a pair of molecules A and B with intermolecular distance R, the force F at any distance is given by

$$F = -\frac{\partial}{\partial R}U, \tag{2.261}$$

where U denotes the interaction energy. A typical potential and the resulting force for the interaction of two molecules or atoms, that has the described properties is shown in figure 2.1. For small values of R, the forces are strongly repulsive. At R_e at the potential has a minimum, which is called the well depth ϵ, where the intermolecular force is zero. In direction of larger distances the potential approaches zero, which results in decreasing attractive forces. To separate the two molecules, which are in equilibrium at R_e, classically an energy corresponding to ϵ will be necessary. The interactions of many molecules, can be expressed in a many body expansion,

$$U = \sum_{i>j} U_{ij} + \sum_{i>j>k} \Delta U_{ijk} + \sum_{i>j>k>l} \Delta U_{ijkl} + \cdots, \tag{2.262}$$

where U_{ij} denote pair interactions, ΔU_{ijk} are three-body corrections to the pair interactions and ΔU_{ijkl} are four-body corrections to three-body interactions. In general, pair interactions are assumed to be the most important and that successive higher many body corrections are of decreasing importance. Van der Waals developed a law for real gases on the basis of intermolecular interactions. Due to the repulsive part of U, a real gas occupies a volume b, which reduces the volume available for free movement. The attractive potential well exerts attractive forces between the molecules in the gas, which reduces the pressure on the container walls by a/V^2, in comparison to the pressure of an ideal gas. The Van der Waals real gas law

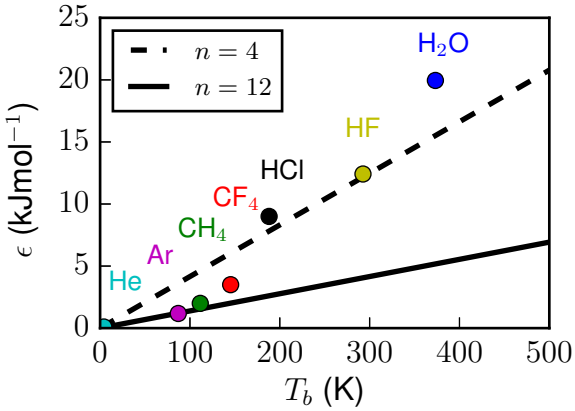


Figure 2.2: Magnitude of van der Waals interactions of small molecules. Well depths ϵ of pair potentials, denoted by lines, are estimated from boiling points T_b and the number of neighbouring molecules n using equation 2.265. Experimental results for small molecules are denoted by coloured circles. The CF_4 dimer well depth from *ab inito* CCSD(T) calculation. Data for He, Xe, CH_4 and H_2O taken from reference [3]. Spectroscopic well depths for HCl and HF dimers taken from [96], and [97] respectively.

is given by,

$$(P + a/V^2)(V - b) = \mathcal{R}T, \tag{2.263}$$

$$P = \frac{\mathcal{R}T}{V - b} - \frac{a}{V^2}, \tag{2.264}$$

where P and V are the measured pressure and volume and $\mathcal{R}$ is the universal gas constant. This modification of the ideal gas law was very successful, and established the term 'van der Waals' forces as a synonym for general intermolecular forces and interactions. The order of magnitude of intermolecular interactions, *i.e.* the well depths ϵ of pair potentials, can be estimated form the enthalpy of vaporization at ambient pressure, which is related the boiling point T_b, by $\Delta H_{\text{vap}} \approx 10\mathcal{R}T_b$. The total energy of N_A molecules in the liquid phase is $1/2N_A n\epsilon$, where only pair interactions between the n nearest neighbours of each molecule are considered. This results in the following approximative relation

$$1/2N_A n\epsilon \approx 10\mathcal{R}T. \tag{2.265}$$

Using this relation, we find that the van der Waals interactions of small molecules, in terms of well depths of pair potentials, are in the range of 1-25 kJ/mol. Figure 2.2 shows that this approximation works surprisingly well, and that measured well depths fall into this range. Rare gas atoms and molecules like CH_4, CF_4 show very weak interactions below 5 kJ/mol. Stronger interactions, $>$ 10 kJ/mol, are found for molecules with permanent dipole moments like HCl, HF and H_2O. However, the van der Waals interactions are weak in comparison with chemical bonds, which are typically $>$ 200 kJ/mol.[3]

2.3.1 Perturbation theory for intermolecular interactions

The main physics behind intermolecular interactions can be described using perturbation theory [3]. In a first approximation one considers only the interactions at long range, that can be obtained by neglecting the correct antisymmetry of the total wavefunction of the pair of molecules due to electron exchange. The zero-order Hamiltonian is that for a pair of non-interacting molecules A and B. It is given by

$$\hat{H}^0 = \hat{H}^A + \hat{H}^B, \tag{2.266}$$

where $\hat{H}^A$ is the Hamiltonian of A, where only its electrons and nuclei interact, and likewise $\hat{H}^B$ for B. The unperturbed wave functions for A and B are Ψ_m^A and Ψ_n^B, with energies E_m^A and E_n^B, respectively. For the total system, the unperturbed solutions are given by,

$$\Psi_m^A \Psi_n^B = |mn\rangle \tag{2.267}$$
$$\hat{H}^0|mn\rangle = (\hat{H}^A + \hat{H}^B)|mn\rangle \tag{2.268}$$
$$= (E_m^A + E_n^B)|mn\rangle \tag{2.269}$$
$$= E_{mn}^{(0)}|mn\rangle. \tag{2.270}$$

The perturbation is given by the Coulomb interactions between the electrons and nuclei, of A with those of B,

$$\hat{H}' = \sum_{a\in A}\sum_{b\in A}\frac{z_a z_b}{r_{ab}}, \tag{2.271}$$

where each sums runs over all particles with charges z_i, i.e electrons and nuclei, of one molecule. Up to second order the energy in non-degenerate RSPT of the electronic ground state of a closed shell system, where $m = n = 0$ is given by,

$$E_{00} = E_{00}^{(0)} + E_{00}^{(1)} + E_{00}^{(2)}, \tag{2.272}$$

where

$$E_{00}^{(0)} = E_0^A + E_0^B, \tag{2.273}$$
$$E_{00}^{(1)} = \langle 00|\hat{H}'|00\rangle, \tag{2.274}$$
$$E_{00}^{(2)} = -\sum_{mn}{}' \frac{\langle 00|\hat{H}'|mn\rangle\langle mn|\hat{H}'|00\rangle}{E_{mn}^{(0)} - E_{00}^{(0)}}. \tag{2.275}$$

The term where both molecules are simultaneously in their ground states, *i.e.* $n = 0$ and $m = 0$, is excluded from the double sum in the expression for the second-order energy. The first-order energy is the electrostatic interaction energy,

$$E_{00}^{(1)} = U_{\text{es}} = \int \frac{\rho^A(\mathbf{r})\rho^B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}', \tag{2.276}$$

of the individual charge densities ρ^A and ρ^B. The second-order energy can be separated in parts, where the sum runs over excited electronic states of only one molecule, say A, while B is in the ground state, and vice versa for B and parts where both molecules are excited. The first type is the induction interaction and the second type is the dispersion interaction. The

second order energy is then,

$$E_{00}^{(2)} = U_{\text{ind}}^{A} + U_{\text{ind}}^{B} + U_{\text{disp}} \tag{2.277}$$

$$U_{\text{ind}}^{A} = -\sum_{m\neq 0} \frac{\langle 00|\hat{H}'|m0\rangle\langle m0|\hat{H}'|00\rangle}{E_m^A - E_0^A}, \tag{2.278}$$

$$U_{\text{ind}}^{B} = -\sum_{n\neq 0} \frac{\langle 00|\hat{H}'|0n\rangle\langle 0n|\hat{H}'|00\rangle}{E_n^B - E_0^B}, \tag{2.279}$$

$$U_{\text{disp}} = -\sum_{m\neq 0}\sum_{n\neq 0} \frac{\langle 00|\hat{H}'|mn\rangle\langle mn|\hat{H}'|00\rangle}{E_{mn}^{(0)} - E_{00}^{(0)}}. \tag{2.280}$$

In total, the result from the perturbation theory is, that the van der Waals interaction in the long range limit is given by

$$U = U_{\text{es}} + U_{\text{ind}}^{A} + U_{\text{ind}}^{B} + U_{\text{disp}}, \tag{2.281}$$

an electrostatic interaction between the unperturbed charge densities of A and B, the induction energy of A, the induction energy of B and the dispersion energy. The second order terms, induction and dispersion, depend on the excited electronic states of A and B. Here, we have for simplicity assumed a separation of the electronic and nuclear states, and only performed the RSPT for the electronic states, and assumed static nuclei. In principle, the molecular Hamiltonians have to be used, which then involves averaging of the electronic interactions over the nuclear distributions of the present vibrational and rotational states of A and B. Further, the summations for the second order energies then involve molecular states, *i.e.* for example contributions from excited rotational states. The application of RSPT to the problem of intermolecular interaction was pioneered by Fritz London [1].[1] He showed, that besides the electrostatic and induction interactions, which were already known from classical theory, a third type of interaction exists, namely the dispersion interaction, which can be described only by quantum mechanics [1]. Hence, the dispersion forces are also called London forces.

2.3.1.1 Electrostatic interactions

From equation 2.276 we know that the electrostatic interaction energy between two molecules, is exactly defined by the individual charge distributions. The electrostatic interaction can be further analysed, if the charge distributions are described in terms of multipole moments. For two neutral molecules one can show [3] that this leads to the following expression,

$$\begin{aligned} U_{\text{es}} = &-T_{\alpha\beta}\mu_\alpha^A\mu_\beta^B - \frac{1}{3}T_{\alpha\beta\gamma}\left(\mu_\alpha^A\Theta_{\beta\gamma}^B - \Theta_{\beta\gamma}^A\mu_\gamma^B\right) \\ &-T_{\alpha\beta\gamma\delta}\left(\frac{1}{15}\mu_\alpha^A\Omega_{\beta\gamma\delta}^B - \frac{1}{9}\Theta_{\alpha\beta}^A\Theta_{\gamma\delta}^B + \frac{1}{15}\Omega_{\alpha\beta\gamma}^A\mu_\delta^B\right) + \cdots, \end{aligned} \tag{2.282}$$

[1] London was Privatdozent at the University of Berlin when he (together with Eisenschitz) published work on the quantum mechanical description of intermolecular forces in 1930. Prior to that he worked with Heitler on the theory of the chemical bond and on chemical reactions. Being of Jewish origin, he lost his position in 1933 due to the racial laws implemented by the Nazis. He was able to leave Germany and to work (with Heinz London) on superconductivity and superfluidity as macroscopic quantum phenomena in England, France and in the USA.[98–100]

where summations are implied over repeated indices. Here, μ_α^A are the components $\alpha = \{x, y, z\}$ of the Cartesian dipole moment tensor of molecule A, likewise $\Theta_{\alpha\beta}^A$ denote the components of its quadrupole moment tensor and $\Omega_{\alpha\beta\gamma}^A$ those of the octopole moment tensor. The tensors T give the orientation and distance dependence of the products of multipole moment tensor elements. In general,

$$T_{\alpha\beta\cdots\nu} = \nabla_\alpha \nabla_\beta \cdots \nabla_\nu \frac{1}{R}. \tag{2.283}$$

Thus, the electrostatic interaction can be described in terms of dipole-dipole, dipole-quadrupole, quadrupole-quadrupole, and interactions involving higher multipole moments. The explicit form for the dipole-dipole interaction can be written in a convenient form, by setting the origin at the center of molecule A at $\mathbf{A}$ and $\mathbf{R} = \mathbf{B} - \mathbf{A}$ along the z axis. The direction of the dipole moments $\boldsymbol{\mu}^A$ and $\boldsymbol{\mu}^B$ are defined by the polar angels θ_A, ϕ_A and θ_B, ϕ_B. The dipole-dipole interaction is then,

$$U_{\mu\mu} = -\frac{\mu^A \mu^B}{R^3} \left(2\cos\theta_A \cos\theta_B - \sin\theta_A \sin\theta_B \cos(\phi_B - \phi_A)\right). \tag{2.284}$$

Thus, the dipole-dipole interaction is proportional to R^{-3} and the orientation dependence is simple, see table 2.1. The most attractive interaction occurs for a head to tail orientation, where both dipole moments point into the same direction, $\theta_A = \theta_B = 0$. If one dipole moment is then turned around, $\theta_A = \pi$, $\theta_B = 0$ the interaction becomes most repulsive. [3]

$\boldsymbol{\mu}^A$	$\uparrow$	$\downarrow$	$\rightarrow$	$\leftarrow$	$\nearrow$
$\boldsymbol{\mu}^B$	$\uparrow$	$\uparrow$	$\rightarrow$	$\rightarrow$	$\nearrow$
$U_{\mu\mu}(\frac{\mu^A\mu^B}{R^3})$	-2	$+2$	$+1$	-1	0

Table 2.1: Orientation dependence of the dipole-dipole interaction. Adapted from [3].

The quadrupole-quadrupole interaction becomes the first non-zero term in equation 2.282, if none of the molecules has a dipole moment and both have quadrupole moments, for example two linear homonuclear diatomic molecules. For two linear molecules only one non-zero independent component of the quadrupole tensor exists. If the molecular axis is oriented along the z-axis the components are $\Theta_{zz} = \Theta$, $\Theta_{xx} = \Theta_{yy} = -\frac{1}{2}\Theta$ and all other components are zero. If both molecules lie on the z-axis and the orientation of their molecular axis are described by polar angles as previously defined for the dipole-dipole interaction, the explicit quadrupole-quadrupole interaction is given by,

$$\begin{aligned} U_{\Theta\Theta} = \frac{\Theta^A \Theta^B}{R^5} \frac{3}{4} \big[& 1 - 5\cos^2\theta_A - 5\cos^2\theta_B - 15\cos^2\theta_A \cos^2\theta_B \\ & + 2(4\cos\theta_A \cos\theta_B - \sin\theta_A \sin\theta_B \cos(\phi_B - \phi_A))^2 \big]. \end{aligned} \tag{2.285}$$

Hence, the quadrupole-quadrupole interaction is proportional to R^{-5} and thus decays faster than the dipole-dipole interaction with the intermolecular distance. In general, the interaction between two multipole moments of rank l and l' is proportional to $R^{-l-l'-1}$. For example, the dipole-quadrupole interaction is proportional to R^{-4}. Because of their faster decay, interactions between higher moments generally become more important at shorter intermolec-

ular distances compared to those those of lower ranks. The orientation dependence of the quadrupole-quadrupole interaction between two linear molecules gives repulsive interactions for linear and parallel orientations, attractive interactions are found for T shaped structures and slipped parallel orientations, see table 2.2. Since, the orientation factors are quite similar for the latter most attractive orientations, the shape of the molecules, and thus how close they are in these structures determines which is favoured. For spherical molecules the T-shape and for long molecules the slipped parallel orientation leads to more attractive quadrupole-quadrupole interactions [3].

$\boldsymbol{\Theta}^A_{zz}$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\circ$	$\nearrow\!\!\swarrow$
$\boldsymbol{\Theta}^B_{zz}$	$\updownarrow$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\nearrow\!\!\swarrow$
$U_{\Theta\Theta}(\frac{\Theta^A\Theta^B}{R^5})$	$+6$	-3	$+2\frac{1}{4}$	$+\frac{3}{4}$	$-2\frac{7}{16}$

Table 2.2: Orientation dependence of the quadrupole-quadrupole interaction for $\Theta > 0$. Adapted from [3].

2.3.1.2 Induction interactions

The terms describing the induction energy, appear in the second order energy of the perturbation treatment for the intermolecular interaction. To understand the physical origin of the induction energy, the perturbation operator $\hat{H}'$ is written in terms of the multipole expansion,

$$\hat{H}' = Tq^A q^B + T_\alpha(q^A\hat{\mu}^B_\alpha - \hat{\mu}^A_\alpha q^B) - T_{\alpha\beta}\hat{\mu}^A_\alpha\hat{\mu}^B_\beta + \cdots . \tag{2.286}$$

After substitution of $\hat{H}'$ in equation 2.279, one integrates over the coordinates of molecule A, which gives

$$U^B_{\text{ind}} = -\sum_{n\neq 0}\frac{\langle 0|T_\alpha q^A\hat{\mu}^B_\alpha - T_{\alpha\beta}\mu^A_\alpha\hat{\mu}^B_\beta + \cdots|n\rangle\langle n|T_{\alpha'}q^A\hat{\mu}^B_{\alpha'} - T_{\alpha'\beta'}\mu^A_{\alpha'}\hat{\mu}^B_{\beta'} + \cdots|0\rangle}{E^B_n - E^B_0} \tag{2.287}$$

$$= -(q^A T_\alpha - \mu^A_\beta T_{\alpha\beta})\sum_{n\neq 0}\frac{\langle 0|\hat{\mu}^B_\alpha|n\rangle\langle n|\hat{\mu}^B_{\alpha'}|0\rangle}{E^B_n - E^B_0}(q^A T_{\alpha'} - \mu^A_{\beta'}T_{\alpha'\beta'}) \tag{2.288}$$

$$= -\frac{1}{2}F^A_\alpha(\boldsymbol{B})\alpha^B_{\alpha\alpha'}F^A_{\alpha'}(\boldsymbol{B}), \tag{2.289}$$

where, for simplicity, only dipoles have been considered. The sum over the states of B involving the dipole operator give the dipole-dipole polarizability of B, denoted by $\alpha^B_{\alpha\alpha'}$, and the constant terms in the round brackets involving the multipoles of A are the negative of the electric field due to A at $\boldsymbol{B}$. The electric field due to the dipole moment of molecule A interacts with the induced dipole $\delta\mu = \alpha F$ at B, which size depends on the dipole-dipole polarizability of B. The induction energy also depends on dipole-quadrupole and quadrupole-quadrupole, and higher multipole polarizabilities. Because the induced field is always the negative of the inducing field, the induction energy is always negative, *i.e.* the resulting force is attractive. Another important property of the induction energy is its non-additivity. Since the induced field depends on the total electric field acting on the molecule B, it is non-additive as the field

can totally change due to the presence of a second neighbour molecule C. Two identical dipolar neighbours can for example cancel each others fields or double the total dipole field, depending on their mutual orientation. [3]

2.3.1.3 Dispersion interactions

To investigate the dispersion energy, we first approximate the perturbation by dipole-dipole interactions only, *i.e.* $\hat{H}' = \hat{\mu}^A_\alpha T_{\alpha\beta} \hat{\mu}^B_\beta$ [3]. We substitute this simplified interaction in the equation 2.280 for the dispersion energy, obtained from perturbation theory and get,

$$U^{(6)}_{\text{disp}} = -\sum_{m\neq 0}\sum_{n\neq 0} \frac{\langle 00|\hat{\mu}^A_\alpha T_{\alpha\beta}\hat{\mu}^B_\beta|mn\rangle\langle mn|\hat{\mu}^A_\gamma T_{\gamma\delta}\hat{\mu}^B_\delta|00\rangle}{E^A_m - E^A_0 + E^B_n - E^B_0}. \tag{2.290}$$

This equation is not easily factorizable into terms depending only on A and only on B because of the denominator. Rearranging the terms gives,

$$\begin{aligned} U^{(6)}_{\text{disp}} = \; & -T_{\alpha\beta}T_{\gamma\delta}\sum_{m\neq 0}\sum_{n\neq 0}\frac{E^A_{m0}E^B_{n0}}{E^A_{m0}+E^B_{n0}} \\ & \times \frac{\langle 0_A|\hat{\mu}^A_\alpha|m_A\rangle\langle m_A|\hat{\mu}^A_\gamma|0_A\rangle}{E^A_{m0}} \frac{\langle 0_B|\hat{\mu}^B_\beta|n_B\rangle\langle n_B|\hat{\mu}^A_\delta|0_A\rangle}{E^B_{n0}}, \end{aligned} \tag{2.291}$$

where $E^A_{m0} = E^A_m - E^A_0$. From the product of the two T tensors of rank $n = 2$, and $T^{(n)}_{\alpha\beta\ldots} \propto R^{n-1}$, we find that the dipole-dipole dispersion interaction is proportional to R^{-6}. It decays faster with R than the respective electrostatic and induction interactions. From the matrix elements in equation 2.291, we may explain this part of the dispersion energy as arising due to the interaction of dipole moments on A and B, which are induced by the correlated motions of the electrons in A and B. Expressions for the frequency dependence of polarizabilities and therefore the dielectric function, involving transition dipole moments as in equation 2.291, where already known from dispersion theory. Hence, London coined the term 'dispersion forces' for $U^{(6)}_{\text{disp}}$. The transition dipole moments, in the form of oscillator strengths (f-values), and excitation energies both known from the electronic spectrum of the atoms, can be used to calculate $U^{(6)}_{\text{disp}}$ from experimental data. The double sum, could be replaced with the product of the polarizabilities of both molecules, if the term with mixed energies would be constant. Indeed, London [101] used the average energy approximation of Unsöld to arrive at the following expression for the dispersion energy,

$$U^{(6)}_{\text{disp}} \approx -\frac{U_A U_B}{4(U_A+U_B)} T_{\alpha\beta}T_{\gamma\delta}\alpha^A_{\alpha\gamma}\alpha^B_{\beta\delta}, \tag{2.292}$$

where U_A and U_B are average excitation energies for both molecules. For atoms the polarizabilities are spherically symmetric, which leads to

$$U^{(6)}_{\text{disp}} \approx -\frac{U_A U_B}{4(U_A+U_B)} \frac{\bar{\alpha}^A\bar{\alpha}^B}{R^6} = -\frac{C_6}{R^6}. \tag{2.293}$$

This shows, that we can expect stronger dispersion interactions for molecules and atoms with larger polarizabilities. An exact expression for $U^{(6)}_{\text{disp}}$ was given by Casimir and Polder [102],

$$U^{(6)}_{\text{disp}} = -\frac{1}{2\pi} T_{\alpha\beta} T_{\gamma\delta} \int_0^\infty \alpha^A_{\alpha\gamma}(i\nu) \alpha^B_{\beta\delta}(i\nu) d\nu, \tag{2.294}$$

involving the polarizability at the imaginary frequency $\omega = i\nu$,

$$\alpha^A_{\alpha\gamma}(i\nu) = \sum_{m'} \frac{\omega_m (\langle 0|\hat{\mu}_\alpha|m\rangle\langle m|\hat{\mu}_\gamma|0\rangle + \langle 0|\hat{\mu}_\beta|m\rangle\langle m|\hat{\mu}_\alpha|0\rangle)}{\omega_m^2 - (i\nu)^2}. \tag{2.295}$$

Frequency dependent polarizabilites occur in the determination of dipole moments induced by oscillating electric fields, $\mu_\alpha(t) = \alpha^A_{\alpha\gamma}(\omega) F_\beta e^{-i\omega t}$. Polarizabilites with imaginary frequencies formally describe the response to exponentially increasing electric fields ($F_\beta e^{\nu t}$). However, the dependence of $\alpha^A_{\alpha\gamma}(i\nu)$ on ν facilitates its *ab inito* calculation via equation 2.295: $\alpha^A_{\alpha\gamma}(i\nu)$ decreases monotonically from the value of the static polarizability at $\nu = 0$ to zero as $\nu \to \infty$. In practise, one evaluates $\alpha^A_{\alpha\gamma}(i\nu)$ and $\alpha^B_{\beta\delta}(i\nu)$ for a number of frequencies and determines the integral in equation 2.294, via numerical quadrature. The next non-zero contribution to the dispersion energy for atoms and centrosymmetric molecules as well as spherically averaged molecules involves the dipole-dipole polarizability on A and the quadrupole-quadrupole polarizability on B, and is proportional to R^{-8}. In general the dispersion energy can be expanded as,

$$U_{\text{disp}} = -\frac{C_6}{R^6} - \frac{C_8}{R^8} - \frac{C_{10}}{R^{10}} + \cdots, \tag{2.296}$$

where we used the convention $C_n > 0$. [3]

Drude model

Dispersion interactions are a quantum effect. To highlight this, and to gain a better understanding of the effect, London [2, 101] investigated it for a model system of two interacting Drude particles. A Drude particle consists of two equal charges of opposing signs, where the negative charge oscillates around the the static positive charge.

In the simplified version [6], presented here, each particle is represented by a 1D harmonic oscillator, where a negative charge -Q oscillates with a frequency ω_0 in z direction around a stationary positive charge +Q. The distance between the particles is R, and the internal coordinates of both particles are z_A and z_B. A sketch of the model is provided in figure 2.3. The mass of the negative charge is M and the force constant of the oscillators is k. Both particles are neutral and have no permanent dipole moment. Thus, there are no electrostatic and induction interactions present between the particles. However, both particles represent oscillating dipoles $\mu_a = Q z_A(t)$ and $\mu_b = Q z_B(t)$. The energy of the two harmonic oscillators in their ground states at infinite separation is given by

$$E(\infty) = E_A + E_B = \omega_0. \tag{2.297}$$

If the particles are brought closer together, their motion is correlated via the interaction of their oscillating dipole moments. The Hamiltonian for the two interacting Drude particles is

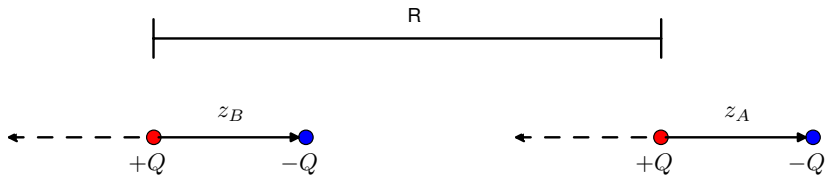


Figure 2.3: A simple model for dispersion: two 1D Drude particles.

given by,

$$\hat{H} = -\frac{1}{2M}\left(\frac{\partial^2}{\partial z_A^2} + \frac{\partial^2}{\partial z_B^2}\right) + \frac{1}{2}kz_A^2 + \frac{1}{2}kz_B^2 - \frac{2z_A z_B Q^2}{R^3}, \tag{2.298}$$

where the last term describes the dipole-dipole interaction between the two particles. Using the normal modes,

$$Z_1 = \frac{z_A + z_B}{\sqrt{2}} \tag{2.299}$$

$$Z_2 = \frac{z_A - z_B}{\sqrt{2}} \tag{2.300}$$

$$\tag{2.301}$$

the Hamiltonian is transformed to the uncoupled form

$$\hat{H} = -\frac{1}{2M}\left(\frac{\partial^2}{\partial Z_1^2} + \frac{\partial^2}{\partial Z_2^2}\right) + \frac{1}{2}k(1 - \frac{2Q^2}{R^3})Z_1^2 + \frac{1}{2}k(1 + \frac{2Q^2}{R^3})Z_2^2. \tag{2.302}$$

The normal mode frequencies are

$$\omega_1 = \omega_0\left(1 - \frac{2Q^2}{R^3 k}\right)^{1/2}, \tag{2.303}$$

$$\omega_2 = \omega_0\left(1 + \frac{2Q^2}{R^3 k}\right)^{1/2}. \tag{2.304}$$

$$\tag{2.305}$$

The energy of the interacting particles in the ground state is given by

$$\begin{aligned} E(R) &= \frac{1}{2}(\omega_1 + \omega_2) \\ &= \frac{1}{2}\omega_0\left(\sqrt{1 + \frac{2Q^2}{R^3 k}} + \sqrt{1 - \frac{2Q^2}{R^3 k}}\right) \\ &= \frac{1}{2}\omega_0\left(2 - \frac{1}{4}\frac{4Q^4}{R^6 k^2} + \cdots\right) \\ &= \omega_0 - \frac{1}{2}\frac{Q^4\omega_0}{R^6 k^2} + \cdots, \end{aligned} \tag{2.306}$$

where we have used the Taylor expansion of $\sqrt{1+x}$ at $x = 0$, which is valid in the long-range limit, where we assume the perturbation to be small. Keeping only the leading term, the interaction energy becomes,

$$U_{\text{disp}} = E(R) - E(\infty) \approx -\frac{1}{2}\frac{Q^4\omega_0}{R^6k^2}. \tag{2.307}$$

The polarizabiliy of a Drude particle is given by $\alpha = Q^2/k$, which, when substituted into the equation above, gives an expression for the dispersion energy familiar to equation 2.293 namely,

$$U^{(6)}_{\text{disp}} = -\frac{1}{2}\frac{\omega_0\alpha\alpha}{R^6}. \tag{2.308}$$

For classical harmonic oscillators the ground state energies are always zero, hence there would be no interaction between the Drude particles in the ground state. The dispersion interaction, thus arises because of the zero point motions of the quantum harmonic oscillators and the lowering of the ground state energy due to the correlation of their motions. Hence, the dispersion force is a true quantum effect, without a classical analogue. [3, 6]

2.3.2 Calculation of intermolecular interactions

Accurate ab initio potential energy surfaces for intermolecular interactions can be obtained by two methods. Namely, the supermolecular approach and symmetry adapted perturbation theory (SAPT).

2.3.2.1 Supermolecular approach

In the supermolecular approach one calculates the interaction energy $U = \Delta E$ as

$$\Delta E(R) = E^{\text{AB}}(R) - E^{\text{A}} - E^{\text{B}}, \tag{2.309}$$

where $E^{\text{AB}}(R)$ denotes the energy of the complex AB at the intermolecular distance R and E^{A} and E^{B} are the energies of the free monomers. Here we assume that the structures of the molecules A and B are unchanged in the dimer, thus we neglect deformation energies. This is called the rigid monomer approximation, which is usually valid for weak non-covalent interactions. The individual energies are calculated using size extensive methods like, MPn or CCSD(T), as described in sections 2.2.2.1 and 2.2.2.2.

Basis set superposition error

For the energy calculations finite atom centred basis sets are used, which leads to an inherent error present in supermolecular calculations using equation 2.309. This error is the basis set superposition error (BSSE) and can be explained in the following manner. If A and B are far apart, only the basis functions on A contribute to the description of the electronic structure of A, while the same holds for B. If A and B are brought closer together, the basis functions on A can also be used for B and vice versa. Thus, decreasing the distance R between the molecules, leads to an increase in the flexibility of the basis from the point of view of the monomers and thus an inconsistent lowering of the interaction energy. The size of the BSSE is large for small basis sets and small for larger basis sets near the complete basis set limit.

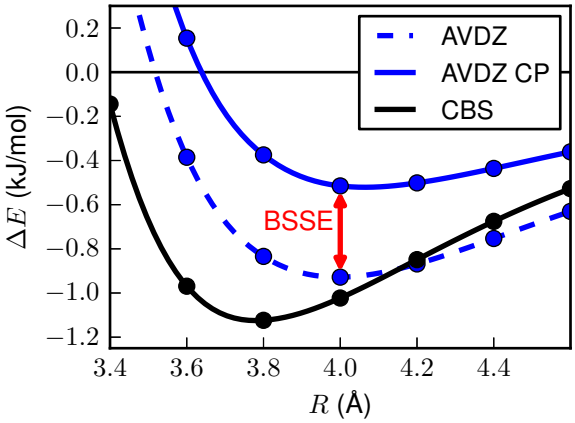


Figure 2.4: Interaction energy of the Ar_2 dimer at the CCSD(T) level of theory with the AVDZ basis set, with and without counterpoise correction (CP) to remove the BSSE and at the complete basis set (CBS) limit. The BSSE results in an overestimation of the binding interaction for the AVDZ basis set. For $R > 4.4$ Å the BSSE results even in an overbinding compared to the curve obtained in the basis set limit.

In order to eliminate the BSSE, the counterpoise method (CP) by Boys and Bernardi [103] is routinely used for the calculation of intermolecular interactions. In this method, three calculations are performed for each point of the interaction energy at R. One is for the complex giving $E^{\mathrm{AB}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B)$, where $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$ denote the positions of the nuclei of the monomers in the complex (although redundant, R is still given). One for each monomer, giving $E^{\mathrm{A}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B)$ and $E^{\mathrm{B}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B)$, where additionally all basis functions of the other monomer are included at $\boldsymbol{R}_B$ or $\boldsymbol{R}_A$, respectively. These extra basis functions in the monomer calculations are called ghost functions. Thus, the same basis is used for the complex as well as the monomers at each R, which leads to a consistent description of the interaction energy along the whole potential, given by

$$\Delta E^{\mathrm{CP}}(R) = E^{\mathrm{AB}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B) - E^{\mathrm{A}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B) - E^{\mathrm{B}}(R, \boldsymbol{R}_A, \boldsymbol{R}_B). \tag{2.310}$$

A detailed description of the BSSE problem and the CP method can be found in reference [104].

As an example for the BSSE, figure 2.4 shows the interaction energy of the Argon dimer at the CCSD(T) level of theory. Using a small basis set (AVDZ) without counterpoise correction leads to a curve that shows some over binding for $R > 4.4$ Å with respect to the curve in the complete basis set limit (CBS). The counterpoise corrected, and thus BSSE free, curve (AVDZ CP) is significantly less binding compared to the curve with BSSE, revealing the large magnitude of the error for this basis set. Furthermore, the CP corrected AVDZ curve lies at all R above the CBS curve. Counterpoise corrected interaction energies converge to the CBS interaction energy from above at all intermolecular distances with increasing basis set sizes, which is certainly not true for non-corrected interaction energies. The BSSE can be further separated into parts for the SCF energy and for the correlation energy. Generally, the BSSE is smaller for the SCF part than for the correlation part, which is due to the faster convergence to the CBS limit of the SCF energy. For the case of the Ar_2 dimer, the BSSE in the

correlation energy is almost twice as high as for the SCF energy at $R = 4.0$ Å. Using the larger AVQZ basis set for Ar_2 eliminates the BSSE at the SCF level (SCF-BSSE ≈ -0.02 kJ/mol), leaving the remaining BSSE of -0.15 kJ/mol at $R = 3.8$ Å attributed to the correlation energy. An alternative approach to reduce the correlation BSSE significantly, is to employ local correlation methods, as described in section 2.2.2.4. Local correlation methods reduce the inconsistent flexibility in the dimer basis, by restricting pair excitations to virtual orbitals, which are spatially close to the two localized MOs. Hence, tails of functions at distant atoms contribute less to the correlation energy, which is a major source for the BSSE [105].

2.3.2.2 Symmetry adapted perturbation theory

The RSPT treatment of intermolecular interactions is only valid in the long range, due to the neglect of the antisymmtery properties of the perturbed wavefunction in the zero order wavefunction. Hence, it cannot be used to calculate the correct intermolecular interactions over the whole range of intermolecular separations. Symmetry adapted perturbation theories (SAPT) overcome this problem by introducing the antisymmetrization operator $\hat{A}$ that ensures, that the Pauli principle is obeyed at all orders in the perturbation expansion [7]. In SAPT the interaction energy is calculated directly, hence it is free of BSSE. Another advantage is, that SAPT interaction energies can be interpreted in terms of different physical contributions, like the electrostatic, induction and dispersion interactions, which arise directly from the theory. A disadvantage of SAPT is, that it is usually restricted to interactions of dimers. Since the antisymmetrized zero-order wavefunction $\hat{A}\Psi_0 = \hat{A}\Psi_A\Psi_B$ is not an eigenfunction of the unperturbed Hamiltonian $\hat{H}_0$ it cannot be directly used in RSPT. This is because the Hamiltonian of the whole system is of higher symmetry, with respect to permutations of the electron indices, than the zero-order Hamiltonian. In the simplest SAPT formulation the symmetrized Rayleigh Schrödinger (SRS) theory [106] the wavefunction corrections are those of the RSPT, *i.e.* the polarization expansion, and the correct antisymmetry is only imposed in the expressions for the energy correction.

Double-perturbation theory

Since Ψ_0 is not known exactly, it has to be approximated by using HF determinants for the monomers and electron correlation at the monomer level is treated perturbatively using MPPT or CCPT for example. The interaction between the monomers is then the second perturbation, in this double perturbation theory approach [7]. The Hamilton operator is then given by,

$$\hat{H} = \hat{F}_A + \hat{F}_B + \hat{V} + \hat{W}_A + \hat{W}_B, \tag{2.311}$$

where $\hat{F}_A$ and $\hat{F}_B$ denote the Fock operators of the molecules A and B, $\hat{W}_A$ and $\hat{W}_B$ account for intramonomer correlation and $\hat{V}$ is the intermolecular interaction operator. Using $\hat{F} = \hat{F}_A + \hat{F}_B$ and $\hat{W} = \hat{W}_A + \hat{W}_B$ the perturbed SE becomes,

$$\left(\hat{F} + \lambda\hat{V} + \mu\hat{W}\right)\Psi_0(\lambda,\mu) = E(\lambda,\mu)\Psi_0(\lambda,\mu), \tag{2.312}$$

where the energy and the wave function are approximated by power series expansions around $\lambda = 0$ and $\mu = 0$. The energy corrections of the double SRS theory, are obtained by the power

series expansion of the function,

$$E_{\mathrm{SRS}}(\lambda,\mu) = \frac{\langle\Psi(0,\mu)|\lambda\hat{V}|\hat{A}\Psi(\lambda,\mu)\rangle}{\langle\Psi(0,\mu)|\hat{A}\Psi(\lambda,\mu)\rangle}. \tag{2.313}$$

The total double SAPT interaction energy is given by a double sum,

$$E_{\mathrm{int}} = \sum_{n=1}^{\infty}\sum_{l=0}^{\infty}(E_{\mathrm{pol}}^{(nl)} + E_{\mathrm{exch}}^{(nl)}), \tag{2.314}$$

over the standard RSPT polarization terms E_{pol}^{nl} and their corrections due to electron exchange E_{exch}^{nl}, where n and l denote the orders with respect to V and W. [7]

A great advantage of SAPT is that the occurring terms, at all orders, can be interpreted by physically meaningful quantities. In general, the main types of interactions can be listed [107] as:

$$E_{\mathrm{electrostatic}} = E_{\mathrm{elst}}^{(10)} + E_{\mathrm{elst,resp}}^{(12)} + E_{\mathrm{elst,resp'}}^{(13)}, \tag{2.315}$$

$$E_{\mathrm{exchange}} = E_{\mathrm{exch}}^{(10)} + E_{\mathrm{exch}}^{(11)} + E_{\mathrm{exch}}^{(12)}. \tag{2.316}$$

$$\begin{aligned} E_{\mathrm{induction}} = {} & E_{\mathrm{ind,resp}}^{(20)} + E_{\mathrm{exch\text{-}ind,resp}}^{(20)} \\ & + E_{\mathrm{ind}}^{(30)} + E_{\mathrm{exch\text{-}ind}}^{(30)} \\ & + {}^{t}E_{\mathrm{ind}}^{(22)} + {}^{t}E_{\mathrm{exch\text{-}ind}}^{(22)} + \delta E_{\mathrm{HF}}^{(3)}, \end{aligned} \tag{2.317}$$

$$\begin{aligned} E_{\mathrm{dipersion}} = {} & E_{\mathrm{disp}}^{(20)} + E_{\mathrm{disp}}^{(30)} + E_{\mathrm{disp}}^{(21)} \\ & E_{\mathrm{disp}}^{(22)} + E_{\mathrm{exch\text{-}disp}}^{(20)} + E_{\mathrm{exch\text{-}disp}}^{(30)}, \end{aligned} \tag{2.318}$$

where ${}^{t}E_{\mathrm{ind}}^{(22)}$ denotes the true induction correlation contributions obtained by eliminating terms from $E_{\mathrm{ind}}^{(22)}$, which are also present in $E_{\mathrm{ind,resp}}^{(20)}$ and likewise for ${}^{t}E_{\mathrm{exch\text{-}ind}}^{(22)}$ [7], the subscript 'resp' denotes orbital response, and $\delta E_{\mathrm{HF}}^{(3)}$ describes higher order induction components present in the HF iteration energy, but not in the truncated SAPT expansion. As can be seen in the above equations, SAPT introduces exchange as a main component to the interaction energy. The leading exchange contribution $E_{\mathrm{exch}}^{(10)}$ describes the 'closed shell repulsion' between the monomers at the HF level, which allows for the repulsive part of the interaction potential at small R. Further, the interactions are comprised of terms with $l = 0$ that describe the interactions of the unperturbed HF charge distributions and contributions $l > 0$, due to intramonomer correlation. As an example, the SAPT/AVQZ energies for the Ar_2 dimer are shown in figure 2.5. The binding interaction comes from the dispersion interaction, which is damped for $R < 4$ Å by the dispersion-exchange contribution. The failure of the RSPT for the electrostatic and induction interactions, which are zero in the long range ($R > 4.5$ Å), is manifested by binding induction and electrostatic interactions in the short range. The induction-exchange part, exactly cancels the binding induction interactions. The repulsive exchange interaction also completely corrects for the binding electrostatic interaction, leaving a repulsive interaction. The resulting interaction potential is in close agreement to that obtained using CCSD(T)/CBS given in figure 2.4. The SAPT analysis shows, that the van-der-Waals minimum of Ar_2 occurs due to underlying repulsive exchange on one hand and attractive dispersion interactions on the other.

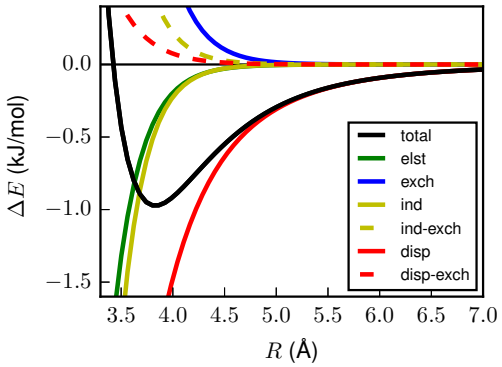


Figure 2.5: SAPT interaction energy of the Ar_2 dimer using the AVQZ basis set, with monomers treated at the HF level.

As SAPT uses the HF wavefunctions for the zero order description, it is possible to deduce which type of interactions are covered by the supermolecular approach at different levels of theory [108]. The HF interaction energy incorporates the following types of interactions,

$$\Delta E_{\mathrm{HF}} \approx E_{\mathrm{elst}}^{(10)} + E_{\mathrm{exch}}^{(10)} + E_{\mathrm{ind,resp}}^{(20)} + E_{\mathrm{exch\text{-}ind,resp}}^{(20)} + E_{\mathrm{ind,resp}}^{(30)} + E_{\mathrm{exch\text{-}ind,resp}}^{(30)} + \cdots . \quad (2.319)$$

The MP2 correlation interaction energy covers,

$$\Delta E_{\mathrm{MP2}}^{corr} \approx E_{\mathrm{disp}}^{(20)} + E_{\mathrm{exch\text{-}disp}}^{(20)} + E_{\mathrm{elst,resp}}^{(12)} + E_{\mathrm{exch}}^{(11)} + E_{\mathrm{exch}}^{(12)} + \cdots . \quad (2.320)$$

Hence, supermolecular HF interaction energies describe the main electrostatic, induction and exchange interactions. Dispersion is only described by the correlation energy contribution. The MP2 theory describes the dispersion interactions of two monomers with HF electron densities, *i.e.* $E_{\mathrm{disp}}^{(20)}$, and the respective exchange correction $E_{\mathrm{exch\text{-}disp}}^{(20)}$. The corrections to the dispersion interaction due to intramonomer correlation, as well as higher order terms $n > 2$, appear in higher order MPn and CC theories [108].

In the many body (MB)-SAPT approach presented so far, intramonomer correlation is handled by a double perturbation expansion starting from uncorrelated HF monomer wavefunctions. Another very popular approach is to combine DFT and SAPT in the so called DFT-SAPT theories. In DFT-SAPT, intramonomer correlation is handled by DFT, *i.e.* KS Slater determinants of the monomers are used for the zero-order wavefunctions. Furthermore, SAPT dispersion interactions, *i.e.* electron correlation between the DFT monomers, are described via response properties obtained from time dependent DFT. The formal scaling of DFT-SAPT is N^6 and can be further reduced to N^5 by density fitting, which is a significant reduction compared to MP-SAPT, which scales as N^7. A great advantage of DFT-SAPT is therefore, that it allows to treat larger systems than possible with MB-SAPT [109].

The following lessons, for the supermolecular approach, have been learned by comparison to SAPT. The supermolecular approach suffers from BSSE, which can be removed by the counterpoise scheme within a given basis set. Dispersion is described only by the correlation energy. The major part for the dispersion energy is obtained at the MP2 level of theory. In order to describe the total intermolecular interactions in a highly accurate manner, one has to use at least CCSD(T) theory.

3 Rotationally adiabatic pair interactions of H_2 with the dihalogen molecules F_2, Cl_2 and Br_2

3.1 Introduction

The strongly exothermic reaction of dihydrogen and dihalogen molecules often results in an explosion. Everyone knows the oxyhydrogen detonating gas reaction, and similar reactions also exist for the halogens as

$$H_2 + X_2 \xrightarrow{h\cdot\nu} 2HX \cdot \tag{3.1}$$

All these light induced radical chain reactions are highly exothermic due to the formation of the HX (X = F, Cl, Br, I) molecules. However, at liquid helium temperatures, the boiling point of ^{4}He is T_b = 4.2 K, molecular dihydrogen ($T_{melting}$ = 14.01 K) can be solidified and used for matrix isolation spectroscopy of dihalogen molecules. At these low temperatures, the above reaction is suppressed, but can be initiated in a controlled manner and studied spectroscopically [29, 30, 110].

Matrix-isolation spectroscopy allows the study of transient species, which decompose rapidly under normal conditions. Trapped in a matrix of rare gas atoms, transient molecular ions, radicals and complexes that define early stages of chemical reactions can be studied [111]. In comparison to the rare gas (Rg) matrices of Ar, Kr, Xe, the matrices of less polarizable species like He, Ne, and H_2 exhibit less interactions with the guest-molecules. In consequence, the respective spectra resemble those of the free molecules more closely [112]. The advantage of H_2 matrices is that the cage-effect is minimal in comparison to rare-gas matrices, which allows for the *in situ* generation of photo-dissociation fragments from molecular precursors [113].

Molecular hydrogen systems are of great interest, because quantum phenomena can be observed more easily, as it is the molecule with the lowest mass [112, 114]. For example, small clusters of hydrogen show superfluid behaviour, *i.e.* a fluid with zero viscosity, at temperatures below 1 K[115–118]. Solid hydrogen is an example for a translational quantum solid, where the zero point energy determines the crystal packing [119–121].

For the theoretical modelling of molecules trapped in matrices, knowledge of the interactions of the guest and host molecules is of key importance. The many-body interaction of the halogen molecule and the surrounding hydrogen molecules can be approximated with the pairwise sum of the H_2-H_2[122, 123] and H_2-X_2 pair potentials. The latter of which are the subject of this investigation. Specifically, we develop *ab initio* pair potentials, describing the van der Waals interactions of the hydrogen molecule, H_2, and the three halogen molecules F_2, Cl_2 and Br_2. Besides describing halogen molecules trapped in solid molecular hydrogen matrices, the potentials can be also used to describe single halogen molecules within clusters

of hydrogen molecules. From the many-body potentials rovibrational and librational spectra of halogen impurities in hydrogen clusters and solid hydrogen, respectively, can be calculated, and compared to experimental data. However, results from models that only take pairwise interactions into account should be taken with a grain of salt, as tree-body interactions are also important [112].

Another use for the pair-potentials is the description of the rovibrational states and spectra of H_2-X_2 van der Waals dimers. Van der Waals dimers of the type Rg-X_2 have been studied extensively and it will be interesting to compare the interactions of H_2 and Rg with the dihalogen molecules. H_2-I_2 is the only van der Waals dimer of the type H_2-X_2 for which binding energies have been measured and compared to calculated potentials [124, 125]. The results highlight the importance of quantum effects, *i.e.* large zero point energy, on the binding energy of the H_2-I_2 van der Waals dimer. It will be instructive to investigate the influence of quantum effects on the binding energy of dimers of H_2 with F_2, Cl_2 and Br_2. Simultaneously, we make predictions for the binding energies of the three dimers which have, to our knowledge, never been measured, or calculated.

Dihydrogen has two nuclear spin isomers, for which different manifolds of rotational quantum states are allowed, namely parahydrogen (pH$_2$) and orthohydrogen (oH$_2$). This affects their thermodynamic properties and their interactions with other molecules. Here, we use a technique called rotational adiabatization [27, 122] to capture the van der Waals interactions of both nuclear spin isomers with the halogen molecules. We then compare these rotationally adiabatic interactions not only among themselves, but also to interactions of rare gas atoms with the halogens. Further, the rotational adiabatic potentials are fitted to analytic model potentials, that facilitate their future use in modelling of cryogenic hydrogen halogen systems. Lastly, we calculate the bound states of the pH$_2$-X_2 van der Waals dimers.

3.1.1 Nuclear spin isomers of the dihydrogen molecule

Hydrogen nuclei (^{1}H) are fermions with a nuclear spin of $I = 1/2$. Hence, the molecular wave function of the hydrogen molecule is antisymmetric with respect to the interchange of the nuclei,

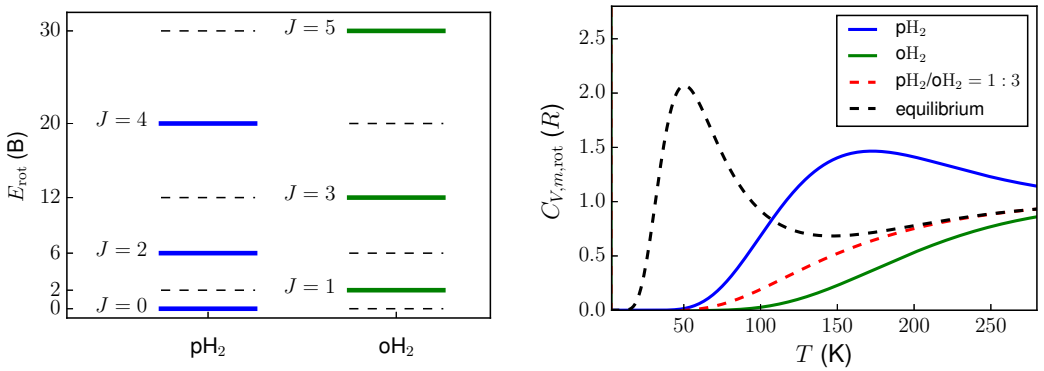
$$\Psi_{\mathrm{mol}}(\mathbf{R}_1, \Sigma_1, \mathbf{R}_2, \Sigma_2) = -\Psi_{\mathrm{mol}}(\mathbf{R}_2, \Sigma_2, \mathbf{R}_1, \Sigma_1), \tag{3.2}$$

where the $\mathbf{R}_1$ and $\mathbf{R}_2$ denote the spatial and Σ_1 and Σ_2 denote the nuclear spin coordinates. To a first approximation the total molecular wave function of H_2 can be written as a product,

$$\Psi_{\mathrm{mol}} = \Psi_{\mathrm{trans}}\Psi_{\mathrm{rot}}\Psi_{\mathrm{vib}}\Psi_{\mathrm{el}}\Psi_{\mathrm{n}}, \tag{3.3}$$

of the translational, rotational, vibrational, electronic and nuclear-spin wave functions. The translational and vibrational wave functions are symmetric under exchange of nuclei, as is the electronic wave function, if the hydrogen molecule is in its electronic ground state $^1\Sigma_g^+$. Hence, the product $\Psi_{\mathrm{rot}}\Psi_{\mathrm{n}}$ has to be antisymmetric under exchange of the nuclei. What are the possibilities for Ψ_{n}? From $I_1 = I_2 = 1/2$ total nuclear spin functions with $I_{\mathrm{mol}} = I \in \{0, 1\}$ can be obtained. The function with $I = 0$ is singly degenerate,

$$\Psi_{\mathrm{n}} = \Psi^{\mathrm{n}}_{I,m_I} = \Psi^{\mathrm{n}}_{0,0} = 1/\sqrt{2}\,(\alpha(\Sigma_1)\beta(\Sigma_2) - \beta(\Sigma_1)\alpha(\Sigma_2))\,, \tag{3.4}$$



(a) Rotational energy levels for parahydrogen (blue) and orthohydrogen (green) molecules. Forbidden rotational states are denoted by dashed lines.

(b) Component of the heat capacity due to rotational and nuclear spin degrees of freedom for pure parahydrogen (blue) and pure orthohydrogen (green) and for mixtures with a fixed 1:3 p/o ratio, as found in normal hydrogen, and for mixtures with ratios $(n_p/n_o)_T$, which are equilibrated for each temperature.

Figure 3.1: Para- and orthohydrogen differ in their manifolds of rotational states (a), which leads to different properties, here exemplified for heat capacities in (b). Adopted from reference [126].

and describes the anti parallel configuration $(\uparrow\downarrow - \downarrow\uparrow)$ of the nuclear spins with $I = 0, m_I = 0$. This function is clearly antisymmetric with respect to exchange of the two spin coordinates. Nuclear spin isomers with this function are called parahydrogen (pH$_2$). The nuclear spin state $I = 1$ is threefold degenerate, and is described by the following functions

$$\Psi^{\mathrm{n}}_{1,1} = \alpha(\Sigma_1)\alpha(\Sigma_2) \tag{3.5}$$

$$\Psi^{\mathrm{n}}_{1,0} = 1/\sqrt{2}\,(\alpha(\Sigma_1)\beta(\Sigma_2) + \beta(\Sigma_1)\alpha(\Sigma_2)) \tag{3.6}$$

$$\Psi^{\mathrm{n}}_{1,-1} = \beta(\Sigma_1)\beta(\Sigma_2), \tag{3.7}$$

for parallel nuclear spins $(\uparrow\uparrow, \downarrow\downarrow)$ and for the positive linear combination of antiparallel spins $(\uparrow\downarrow + \downarrow\uparrow)$. All of which are symmetric with respect to to exchange of the two spin coordinates. The nuclear spins isomers with $I = 1$ are called orthohydrogen (oH$_2$). The symmetry of the rotational wave function with respect to exchange of nuclei depends on the rotational quantum number J,

$$\Psi^{\mathrm{rot}}_{J,m_J}(\mathbf{R}_1, \mathbf{R}_2) = (-1)^J \Psi^{\mathrm{rot}}_{J,m_J}(\mathbf{R}_2, \mathbf{R}_1). \tag{3.8}$$

Depending on whether J is even or odd the rotational wave function is either symmetric or antisymmetric, respectively. Hence, the parahydrogen molecules with antisymmetric Ψ_{n} are restricted to symmetric Ψ_{rot} with $J = 0,\ 2,\ 4\ \cdots$, in order to fulfil the antisymmetry requirements of Ψ_{mol}. Likewise, orthohydrogen molecules with symmetric Ψ_{n} are restricted to rotational states with $J = 1,\ 3,\ 5,\ \cdots$, which are antisymmetric. The resulting rotational energy levels for parahydrogen and orthohydrogen are depicted in figure 3.1a. This leads to different rotational partition functions for both isomers and hence different contributions to the heat capacity, as is shown in figure 3.1b. The ratio 1:3 of para- to orthohydrogen at high temperatures is directly related to the degeneracy of the nuclear spin states. Pure parahydro-

gen can be obtained by cooling a para/orthohydrogen mixture in presence of a catalyst [126]. The nuclear spin isomers of hydrogen also show different electrostatic interactions with other molecules. Parahydrogen in the rotational ground state with $J = 0$ is spherically symmetric, and thus the quadruple moment of H_2 vanishes. The rotational ground state of orthohydrogen with $J = 1$ is only axially symmetric, thus a quadrupole moment is present and allows electrostatic interactions with other molecules. The aim of this work is to describe the different interactions of para- and orthohydrogen ab inito with dihalogen molecules, hence within the rotationally adiabatic model we have to account for their different rotational states.

3.1.2 Small molecules trapped in dihydrogen matrices

Since the 1980's solid molecular parahydrogen is used as a host matrix for matrix isolation high resolution IR spectroscopy [127]. Due to the weak interactions with the parahydrogen molecules the spectra of small guest molecules exhibit sharp lines, indicating that J and ν remain good quantum numbers [127]. Also the cage effect is greatly reduced in $p\mathrm{H}_2$ matrices, therefore photo reactions can be observed in situ. Dopant molecules induce IR activity in neighbouring $p\mathrm{H}_2$ molecules, which leads to additional adsorption bands in the spectrum directly related to the intermolecular interactions [128, 129]. Low concentrations of orthohydrogen present in the matrix interact via long-range quadrupole-quadrupole electrostatic interactions often stronger with the guest molecules compared to the parahydrogen molecules [112]. This can lead to the accumulation of $o\mathrm{H}_2$ near the dopant molecules [127]. The 'quantum diffusion' of $o\mathrm{H}_2$ through the $p\mathrm{H}_2$ matrix is dominated by a nuclear spin and rotational angular momentum exchange between neighbouring molecules. If the dopant molecule has a non-zero total nuclear spin its magnetic field can also induce an *o-p* conversion of the neighbouring molecules, which competes with the 'quantum diffusion' effect [127, 129]. A remarkable feature of $p\mathrm{H}_2$ matrices is the possibility to study photo reactions. By the solvation of traces of Cl_2 in solid parahydrogen at 2 K and subsequent photo dissociation of the Cl_2 molecules the reaction

$$\mathrm{Cl}(^2P_{3/2}) + \mathrm{H}_2(\nu{=}1\,) \longrightarrow \mathrm{HCl} + \mathrm{H} \tag{3.9}$$

could be observed [29, 30, 110]. It was found, that a vibrational excitation of the H_2 molecules to $\nu = 1$ is necessary for the reaction to proceed. The produced $\mathrm{Cl}(^2P_{3/2})$ atoms and the reaction products carry an excess of heat that is dissipated efficiently due to the high thermal conductivity of the $p\mathrm{H}_2$ matrix, which is another major advantage for the study of reactions at low temperatures. For the detailed quantum dynamical study of this and similar reactions, starting from X_2 in solid $p\mathrm{H}_2$, knowledge of the librational states [130, 131] of the X_2 guest and the response of the host molecules is necessary to predict an initial state of the system [27]. The quantum dynamics of similar hindered rotations of dopant molecules in rare gas matrices have been extensively studied in experiment and theory [132–137]. For this purpose, the pair interactions calculated in this work are one major ingredient to construct many-body potentials, as has been demonstrated for Cl_2 in solid $p\mathrm{H}_2$ in reference [27].

3.2 Pair potentials

In this work we closely follow the methodology of Accardi developed for H_2-Cl_2 [27], to describe also the dimers H_2-F_2 and H_2-Br_2. This will enable us to compare the interactions of

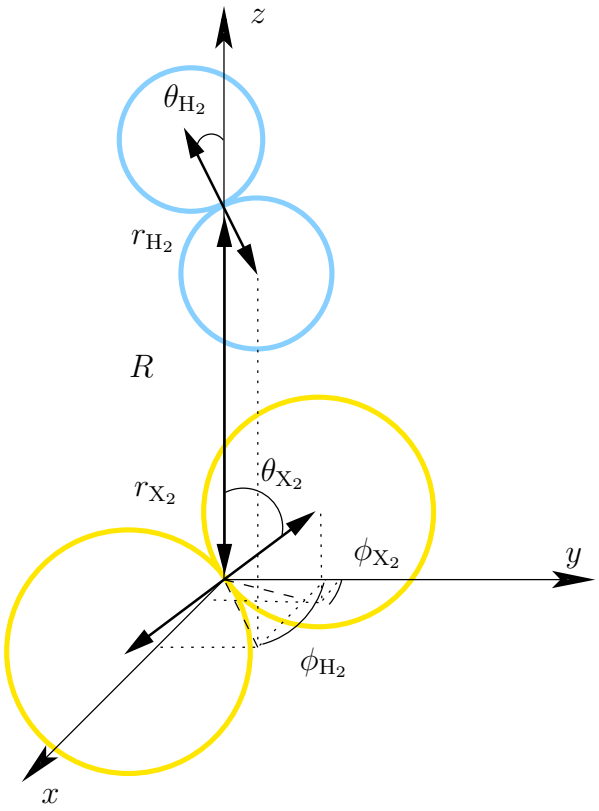


Figure 3.2: Internal Jacobi coordinates describing the structure and orientation of H_2-X_2 dimers. H_2 is depicted in blue and X_2 in yellow. This figure was published in reference [139].

the halogens to hydrogen and other small molecules. So far only few other works on the interactions of halogen molecules and hydrogen exist. The van der Waals interactions of H_2-I_2 have been studied experimentally [124, 125] and theoretically for selected high symmetry structures [124]. The pair potential of H_2-F_2 has been studied earlier by Pham Van at the CCSD(T) level of theory, for the calculation of second virial coefficients and simulation of vapor-liquid equilibria [138].

3.2.1 Methodology

The interaction energy ΔE of a halogen molecule X_2 with a hydrogen molecule H_2, where both are in their respective $^1\Sigma_g^+$ electronic ground states, is a function of six internal coordinates that describe the relative position of the molecules to each other. In order to describe the rotation of H_2 within the dimer efficiently, we have chosen a combination of spherical and Jacobi coordinates as depicted in figure 3.2. The monomers are characterized by their bond lengths r_{X_2} and r_{H_2}. The distance between the molecules, measured as the distance between their centers of mass (COM), is denoted by R. Their mutual orientation is given by their polar angles θ_{H_2} and θ_{X_2} and the difference between the azimuthal angles $(\phi_{H_2} - \phi_{X_2})$. The polar angle θ is measured as the angle between the molecular axis and the z-axis of the internal reference frame. Whereas the azimuthal angle ϕ is defined as the angle between the y-axis

and the projection of the molecular axis onto the xy-plane[1]. We can safely assume, that the interaction of the dimers is only weakly changed by the small deviations in the monomer bond lengths due to zero point motions. At the low temperatures considered here, all H_2 and X_2 molecules are in their vibrational ground states. Thus, to reduce the dimensionality of the problem we invoke the rigid monomer approximation by setting the monomer bond lengths to constant values close to their averages with respect to their vibrational ground states. Throughout the following values are applied $r_{H_2} = 75.7$ pm, $r_{F_2} = 141.2$ pm, $r_{Cl_2} = 203.3$ pm and $r_{Br_2} = 228.1$ pm. Within this approximation the interaction energy is a four-dimensional function $\Delta E(R, \theta_{X_2}, \theta_{H_2}, \phi_{H_2} - \phi_{X_2})$ which we calculate using the supermolecule approach as,

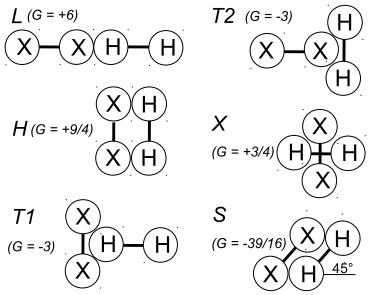
$$\Delta E = E(H_2X_2) - (E(X_2) + E(H_2)), \tag{3.10}$$

where $E(H_2X_2)$, $E(X_2)$ and $E(H_2)$ denote the total electronic energy of the dimer and the monomers. We also apply the counterpoise scheme [103] to correct for the BSSE, as discussed in section 2.3.2. The binding interactions of linear non-dipolar molecules are dominated by electrostatic quadrupole-quadrupole interactions and dispersion interactions. As already discussed in the analysis of SAPT interaction terms, the major part of the electrostatic interactions are described at the HF level, while the leading dispersion terms are effects of electron correlation at the MP2 level of theory. For a highly accurate description of the interaction energy CCSD(T) theory has to be applied. We have therefore calculated the interaction energy at the MP2, CCSD(T) and HF-SAPT levels of theory using the MOLPRO code [140–142]. For H, F and Cl, the all-electron augmented correlation consistent aug-cc-pVnZ basis sets (AVnZ) with cardinal numbers up to n = 6 (H, F) and n = 5 (Cl) [143], while for Br the scalar relativistic effective core potential ECP10MDF and the adapted aug-cc-pVnZ basis sets up to n = 5 have been used [144]. All quantum chemistry data for the H_2-Cl_2 dimer was provided by Accardi, as described in [27].

3.2.2 CCSD(T) results

We aim for a detailed description of the four-dimensional interaction energy hyper surface, on which we can perform the adiabatization with respect to the rotations of H_2. In practise this involves a large number of single point calculations at the MP2 level of theory, as using CCSD(T) for the whole surface is prohibited by our computational resources. To asses the errors introduced by using MP2 instead of CCSD(T), we have performed highly accurate calculations of the CCSD(T) interaction energy at the complete basis set (CBS) limit for six highly symmetrical H_2-X_2 structures, depicted in figure 3.3a. The selected structures fall into three groups. The first group collects the linear (L) and the second T-shaped (T2) structure, where in both cases $\theta_{X_2} = 0°$ holds, see also table 3.3b. In both cases the COM of the hydrogen molecule and the halogen atoms are arranged linearly. Hence, the hydrogen molecule sits on one end of the halogen molecule. The second group contains only the slipped parallel structure (S), which is an intermediate structure connecting the first and third groups of structures. The third group contains the parallel (H), first T-shaped (T1) and cross (X) type structures. Common for these structures is that $\theta_{X_2} = 90°$. In this group the COM of the hydrogen

[1] The usual convention is to define ϕ with respect to the x-axis of the reference frame. The different definition used here, however does not influence the results, as they are only dependent on the difference $(\phi_{H_2} - \phi_{X_2})$ between both angles.



(a) Six high symmetry structures of the H_2-X_2 dimers. The combination of angles for each structure is given in table 3.3b. The short names of the structures are denoted by letters of similar form, H for parallel, X for cross and T for T-shaped structures, and abbreviating letters L and S for linear and slipped parallel structures, respectively. This figure was published in reference [139].

structure	θ_{X_2}	ϕ_{X_2}	θ_{H_2}	ϕ_{H_2}
T2	0°	0°	90°	0°
L	0°	0°	0°	0°
S	45°	0°	45°	0°
H	90°	0°	90°	0°
T1	90°	0°	0°	0°
X	90°	0°	90°	90°

(b) Polar angles for the six structures of the H_2-X_2 dimers. For the definition of the angles with respect to the frame of reference see figure 3.2.

Figure 3.3: High symmetry structures of H_2-X_2 dimers.

molecule is positioned perpendicular to the halogen bond axis, directly above the COM of the halogen molecule.

3.2.2.1 The H_2-F_2 pair potential

The CCSD(T)/AV(56)Z potential energy curves for the H_2-F_2 dimer are shown in figure 3.4a. The overall strongest interaction in terms of absolute values of $\Delta E(R_e)$ is found for the T2 structure with a well depth of 1.283 kJ/mol at 0.347 nm. The weakest interaction with -0.303 kJ/mol and the largest R_e, with 0.392 nm is present for the L structure. Figures 3.4 b-e depict the convergence of the interaction energies for the T2 and L structures with respect to basis set size. For T2, the well depth increases from AVTZ to AV(56)Z by about 0.18 kJ/mol, which is 14% of the CBS value. For L, the respective change is 0.08 kJ/mol, or 25%. Using an AVQZ basis set significantly improves the description of the well depths, as the relative errors decrease to below 10%. Further small improvements are then obtained by the AV5Z and AV6Z basis sets. Well depth positions shift very slightly (< 0.05 nm) to smaller R upon increasing the basis set from AVTZ to AV(56)Z. Similar convergence behaviours were also found for the remaining structures. With a well depth of 0.673 kJ/mol at 0.363 nm, the minimum of structure S lies in between those of the T2 and L structures. Very similar well depths characterise the potentials of the third group, with 0.697, 0.835 and 0.658 kJ mol for H, T1 and X, respectively. Also the R_e values of 0.314 nm, 0.325 nm and 0.318 nm for H, T1 and X are very similar. The ordering in terms of well depths, characterised from strongest to weakest interaction, for H_2-F_2 is therefore T2 $\gg$ T1 $>$ H $\gtrsim$ S $\sim$ X $\gg$ L.

3.2.2.2 The H_2-Br_2 pair potential

Figure 3.5a shows the CCSD(T)/AV(Q5)Z pair potentials for the H_2-Br_2 dimer. Again, the T2 structure has the overall deepest well depth, with 3.182 kJ/mol at 0.4115 nm. The weakest interaction is also found for the L structure, where the well depth is 0.572 kJ/mol at 0.467 nm.

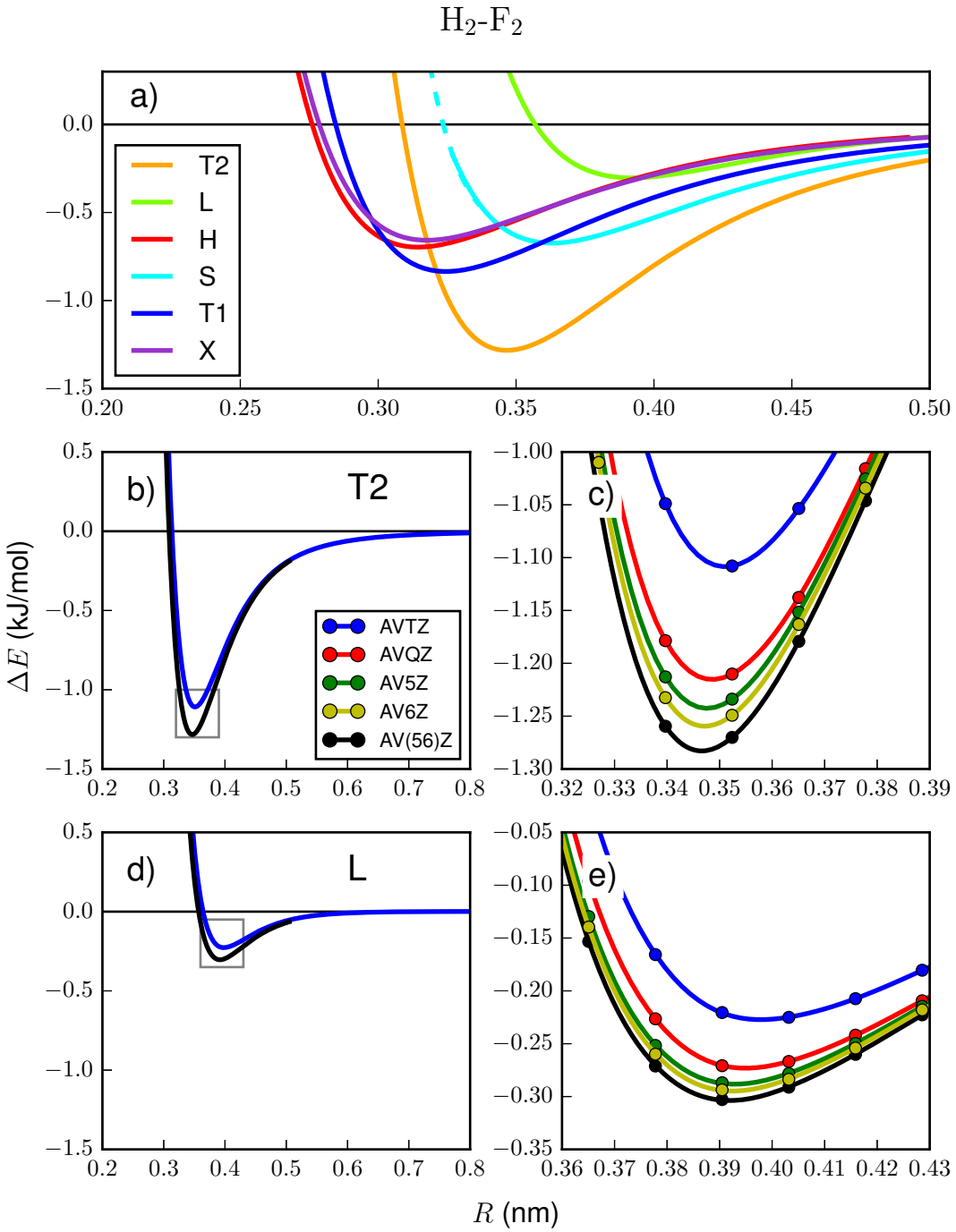


Figure 3.4: CCSD(T) interaction energy of the H_2-F_2 dimer for the six high symmetry structures shown in figure 3.3a. Panel a) shows the interaction energies of all structures obtained by extrapolation to the CBS limit as CCSD(T)/AV(56)Z. In b) and d) the CCSD(T) results obtained using the AVTZ basis for T2 and L show the long range behaviour of the interactions. Additionally, AV(56)Z results are superimposed around the minima. Panels c) and e) depict the convergence to the CBS limit near the minima for T2 and L. The area of the plots c) and e) is denoted by gray boxes in b) and d), respectively.

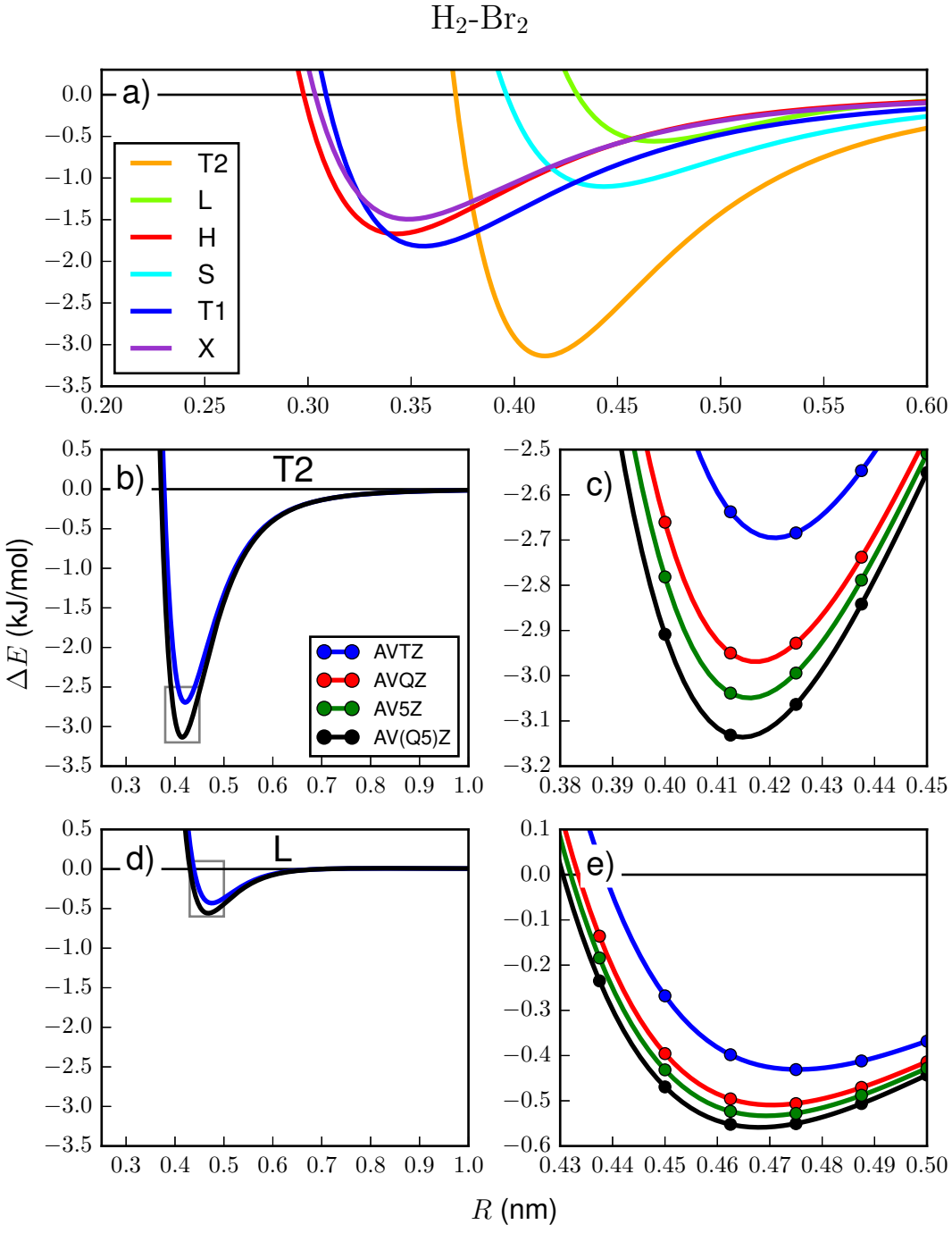


Figure 3.5: CCSD(T) interaction energy of the H_2-Br_2 dimer for the six high symmetry structures shown in figure 3.3a. Panel a) shows the interaction energies of all structures obtained by extrapolation to the CBS limit as CCSD(T)/AV(Q5)Z. In b) and d) the CCSD(T) results obtained using the AVTZ basis for T2 and L show the long range behaviour of the interactions. Additionally, AV(Q5)Z results are superimposed around the minima. Panels c) and e) depict the convergence to the CBS limit near the minima for T2 and L. The area of the plots c) and e) is denoted by gray boxes in b) and d), respectively.

Here, the AVQZ basis set also significantly improves the description of the pair potentials, compared to the AVTZ basis. The well depth of the S structure is with 1.038 kJ/mol close to that of the L structure and lies above those of the third group. Concerning the third group, the well depths are 1.704, 1.825 and 1.434 kJ/mol at 0.342, 0.356 and 0.351 nm for H, T1 and X. Hence, they are energetically more separated than those of the H_2-F_2 dimer. The larger size of the Br_2 molecule is evident from R_e values which are 0.03-0.07 nm larger compared to H_2-F_2. The ordering of the interactions for the different structures is T2 $\gg$ T1 $\gtrsim$ H > X $\gg$ S > L, where compared to H_2-F_2 the positions of the S and X structures are interchanged.

3.2.2.3 Analysis in terms of interaction contributions and comparison to the H_2-Cl_2 pair potential

The H_2-Cl_2 pair potentials for the orientations in 3.3a, were obtained by Accardi [27] at the CCSD(T)/AV(Q5)Z level of theory. The H_2-Cl_2 well depths follow the same pattern, as discussed for the two other halogens. Energetically they are intermediate between the two, while overall closer to those of H_2-Br_2. In order to better understand on the one hand, the interaction of H_2 with F_2, Cl_2 and Br_2, and on the other hand, the differences found for different structures for a given H_2-X_2 dimer, we have calculated individual interaction components at the HF-SAPT/AVQZ level of theory. The HF-SAPT calculations were performed only at the CCSD(T)/CBS minima R_e, for all structures including the H_2-Cl_2 dimer. The HF-SAPT/AVQZ results are presented in table 3.1. Focussing on the T2 structures of different halogens, we see that the most important binding contribution is the dispersion energy (E2disp+E2disp-exch), followed by electrostatic interactions (E1pol). Induction interactions (E2ind+E2ind-exch) are increasingly important only for Cl_2 and Br_2. However, the largest contribution for T2 structures is the repulsive exchange energy (E1exch). Size-wise the values of E1exch for T2 are comparable to the absolute values of the uncorrected dispersion interactions E2disp, *i.e.* without the dispersion exchange correction. The observed increase in the interaction of H_2 with F_2, Cl_2 and Br_2 can be attributed to the simultaneous increase of all types of interactions. Comparing the T2 dimers for F_2 and Br_2, we find that the electrostatic, exchange and dispersion interactions increase by more than a factor of two.

Next, we compare the different structures in terms of the individual interaction contributions. For the electrostatic interactions, we have also calculated the quadrupole-quadrupole (QQ) interactions, because the quadrupole moments are the lowest non-vanishing multipole-moments for linear neutral homonuclear molecules. We therefore expected these contributions to give an overall good approximation to the electrostatic energy. The quadrupole-quadrupole interactions are given by,

$$U_{Q_{X_2}Q_{H_2}}(R) = \frac{Q_{X_2}Q_{H_2}}{4\pi\epsilon_0 R^5} \times G(\theta_{X_2}, \theta_{H_2}, \phi_{H_2} - \phi_{X_2}), \tag{3.11}$$

see also equation 2.285. The components of the traceless quadrupole tensors along the bond axis are $Q_{H_2} = 0.4252\ ea_0^2$ [27], $Q_{F_2} = 0.7635\ ea_0^2$, $Q_{Cl_2} = 2.5427\ ea_0^2$[27], $Q_{Br_2} = 3.4936\ ea_0^2$, calculated at the CCSD(T)/AV6Z, MP2/AV6Z, MP2/AV6Z[27] and MP2/AV5Z levels of the-

ory, respectively. The geometrical factor G is given by

$$G(\theta_{X_2}, \theta_{H_2}, \phi_{H_2} - \phi_{X_2}) = \tfrac{3}{4}\{1 - 5\cos^2\theta_{X_2} - 5\cos^2\theta_{H_2} - 15\cos^2\theta_{X_2}\ \cos^2\theta_{H_2} + \\ + 2[4\cos\theta_{X_2}\ \cos\theta_{H_2} - \sin\theta_{X_2}\ \sin\theta_{H_2}\ \cos(\phi_{H_2} - \phi_{X_2})]^2\}, \quad (3.12)$$

where the respective G values for the six structures are shown in figure 3.3a. For plots of the QQ potentials $U_{Q_{X_2}Q_{H_2}}(R)$, see figure 3.6a, we define the minimal contact distance in terms of a hard sphere model. In this model atoms are represented by hard spheres with van der Waals radii of 110 pm, 147 pm, 175 pm and 185 pm for the H, F, Cl and Br atoms, respectively [145]. The values of $U_{Q_{X_2}Q_{H_2}}(R_e)$ at the minima of the CCSD(T)/CBS curves are given in table 3.1. HF-SAPT gives binding electrostatic interactions for all structures, with the exception of the L structure. The QQ interactions are binding for T2, S, T1 and repulsive for L, H and X, due to the signs of the G factor. The best agreement between E1pol and QQ at the minima is found for the T1 structure and the L structure of the dimer with F_2. For all other structures the differences between both models suggest, that interactions between multipole moments of higher than quadrupolar order become important near the minima. The electrostatic interactions are significantly higher for the T2 structure, than for the other structures. For S, T1 and X the values of E1pol are similar, while for H the second highest electrostatic interactions are present in the case of the dimers with Cl_2 and Br_2. The exchange interaction E1exch is highest for T2, followed by H, it is similar for T1 and X, and also for L and S. The induction interactions (E2ind+E2ind-exch) are weak in comparison to the other interactions. They are basically absent in the interaction of H_2 with F_2, but increase for the interaction with Cl_2 and Br_2. The induction interaction is more important for T2 and L structures, while less important for the other structures. Higher than second order induction terms do have a sizeable contribution for the T2 structure, as can bee seen from the δHF contribution. Dispersion interactions (E2disp+E2disp-exch) are the overall largest binding interactions. None of the structures would be bound without dispersion interactions, as the binding electrostatic interactions are too weak in comparison to the exchange repulsion interactions. The ordering in terms of dispersion interactions is T2 >> H > T1 >$\approx$ X > L > S. Concerning the total HF-SAPT/AVQZ energy we find some differences with respect to the CCSD(T)/CBS values. Especially the absolute well depth of the T2 (not for H_2-F_2) and S structures is overestimated by HF-SAPT. For H_2-F_2 the ordering of the SAPT well depths is the same as with CCSD(T). For H_2-Cl_2 and H_2-Br_2 the order of HF-SAPT the well depths for the third group of structures is changed, in that the well depth of H is predicted to be larger than that of T1. These discrepancies, can be explained by the HF description of the monomers in HF-SAPT and by the fact, that we compare finite AVQZ basis to extrapolated CBS results. A better agreement between the SAPT and CCSD(T) results could be obtained with MP2-SAPT or DFT-SAPT. However, it is clear that the SAPT analysis provides a much better picture of the interactions, compared to the simple QQ model. In conclusion, the T2 structure shows the largest interaction energy of all structures, because the largest repulsive exchange interactions are counter balanced by the largest dispersion, electrostatic and induction interactions, where for the latter two, higher order terms, estimated by δHF, are contributing significantly to the interaction energy. The L and S structures have the smallest interaction energy, because, although the repulsive exchange interactions are smallest, the dispersion interactions are also small, plus the electrostatic interaction of the L structure is repulsive. The third group of

structures, H, T1 and X, has interaction energies, intermediate between those of T2 and S. The reason for this is a fine balance of similar electrostatic but larger exchange and dispersion interactions, in comparison to those of the S structure.

3.2.2.4 MP2/AVQZ pair potentials

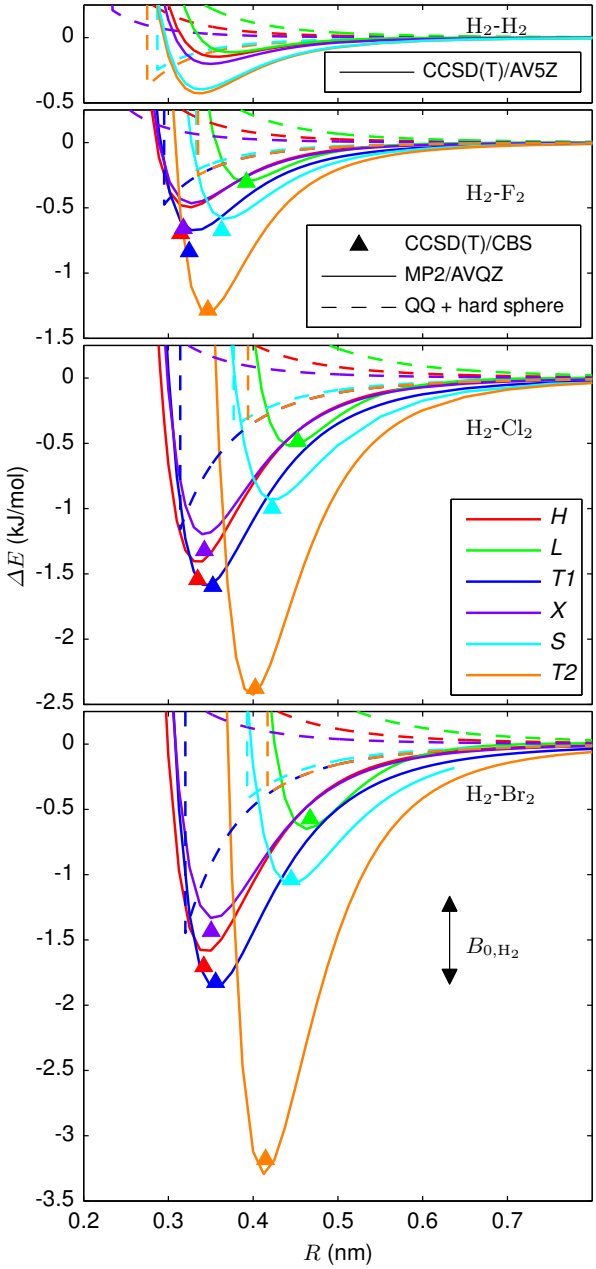
In order to describe the rotation of H_2 in the presence of X_2, the potential energy hyper surface $\Delta E(R, \theta_{X_2}, \theta_{H_2}, \phi_{H_2} - \phi_{X_2})$ has to be calculated for a large number of orientations and intermolecular distances. Hence, we will use the computationally less demanding MP2 method to describe the full surface and use the expensive CCSD(T) method only for the parts where symmetry can be exploited. Therefore, we need to access the accuracy of MP2 for the pair potentials. MP2 only describes dispersion interactions of second order, thus it is less exact than CCSD(T). For dispersion bound systems using MP2 often results in an over estimation of well depths, which opens the opportunity to exploit error cancellation effects of the overestimation of dispersion interactions by the method on one hand and underestimation of dispersion interactions due to the basis set incompleteness error on the other hand. The MP2/AVQZ pair potentials are shown in figure 3.6a, where the results for H_2-Cl_2 and H_2-H_2 where taken from reference [27]. The relative errors of MP2 well depths for AVnZ basis sets of increasing size are shown in figure 3.6b. Using the AVTZ basis, results in a severe underestimation of the well depths, as shown for H_2-F_2 and H_2-Br_2. For MP2/AVQZ the well depths are much improved, with mean absolute relative deviations of 15%, 6%, and 7% for the dimers with F_2, Cl_2 and Br_2. Using the AVQZ basis, the MP2 well depths of the T2, T1 and S structures of H_2-Cl_2 and H_2-Br_2 and those of T2 and L for H_2-F_2 are in excellent agreement to the CCSD(T) reference values. Absolute errors for these structures range from 0.002 kJ/mol (T1, H_2-Cl_2) to 0.01 kJ/mol (S, H_2-F_2). The respective remaining well depths, H, X, (T1), are underestimated by 0.13 kJ/mol-0.22 kJ/mol, but the energetic ordering remains the same as with CCSD(T)/CBS. The only change in the ordering occurs for the H, X, and S structures of H_2-F_2, where the CCSD(T)/CBS well depths lie in a range of just 0.04 kJ/mol. Here the CCSD(T)/CBS ordering H > S > X, is changed to S > H > X with MP2/AVQZ. The absolute errors for H_2-F_2 are similar to those of the other dimers, but the overall lower magnitude of the interaction energy results in larger relative errors and larger errors for the anisotropy of the potential. The AV5Z basis set does not lead to an improvement over the AVQZ results, as it gives on one hand slightly lower errors for S, H and X, while on the other hand the errors for L ans T2 increase. Overall, MP2/AVQZ results in a very good description of the pair potentials at significantly reduced cost, in comparison to CCSD(T)/CBS.

H_2-F_2	T2	L	S	H	T1	X
R_e (nm) / ΔE (kJ/mol)	0.347	0.392	0.363	0.314	0.325	0.318
E1pol	−1.256	0.240	−0.364	−0.352	−0.323	−0.272
E1exch	2.787	0.714	0.790	1.122	1.089	0.968
E2ind+E2ind-exch	−0.023	−0.029	−0.003	0.008	−0.008	0.008
E2disp+E2disp-exch	−2.233	−1.207	−1.062	−1.487	−1.464	−1.334
δHF	−0.442	−0.115	−0.069	−0.053	−0.061	−0.043
SAPT/AVQZ	−1.167	−0.397	−0.708	−0.762	−0.767	−0.673
CCSD(T)/CBS	−1.283	−0.303	−0.673	−0.697	−0.835	−0.658
MP2/AVQZ	−1.310	−0.299	−0.585	−0.498	−0.675	−0.465
QQ	−0.212	0.229	−0.137	0.259	−0.294	0.082

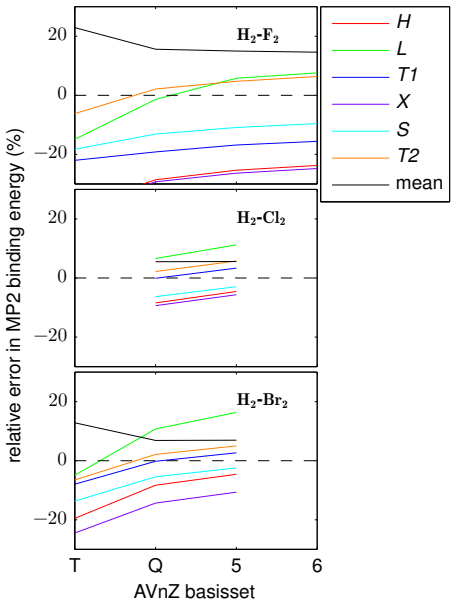
H_2-Cl_2	T2	L	S	H	T1	X
R_e (nm) / ΔE (kJ/mol)	0.40	0.45	0.42	0.33	0.35	0.34
E1pol	−2.690	0.866	−0.692	−1.024	−0.569	−0.568
E1exch	5.072	1.170	1.171	3.050	2.047	2.113
E2ind+E2ind-exch	−0.413	−0.174	−0.051	−0.062	−0.084	−0.031
E2disp+E2disp-exch	−4.072	−2.061	−1.659	−3.474	−2.878	−2.784
δHF	−0.725	−0.177	−0.104	−0.195	−0.161	−0.108
SAPT/AVQZ	−2.828	−0.377	−1.335	−1.704	−1.645	−1.378
CCSD(T)/CBS*[27]	−2.376	−0.486	−0.996	−1.543	−1.595	−1.321
MP2/AVQZ*[27]	−2.428	−0.518	−0.934	−1.413	−1.593	−1.197
QQ*[27]	−0.345	0.383	−0.220	0.677	−0.673	0.194

H_2-Br_2	T2	L	S	H	T1	X
R_e (nm) / ΔE (kJ/mol)	0.415	0.467	0.445	0.342	0.356	0.351
E1pol	−3.737	1.062	−0.704	−1.092	−0.725	−0.634
E1exch	6.177	1.364	1.043	3.108	2.551	2.281
E2ind+E2ind-exch	−1.206	−0.362	−0.085	−0.158	−0.183	−0.094
E2disp+E2disp-exch	−5.194	−2.530	−1.683	−3.777	−3.464	−3.108
δHF	0.292	0.005	−0.017	−0.022	−0.068	0.013
SAPT/AVQZ	−3.668	−0.461	−1.446	−1.940	−1.890	−1.543
CCSD(T)/CBS	−3.182	−0.572	−1.038	−1.704	−1.825	−1.434
MP2/AVQZ	−3.203	−0.619	−1.043	−1.533	−1.815	−1.280
QQ	−0.396	0.436	−0.226	0.781	−0.853	0.230

Table 3.1: Interaction energies of the H_2-X_2 dimers at the minima R_e of the CCSD(T)/CBS pair potentials for the six structures given in table 3.3a. To understand the nature of the interactions, the table further shows SAPT/AVQZ energy components and electrostatic quadrupole quadrupole (QQ) interactions. *CCSD(T) results for H_2-Cl_2 are taken from Accardi [27].



(a) Interaction energies ΔE between H_2 and H_2, F_2, Cl_2 and Br_2 for the static orientations shown in figure 3.3a. *Ab initio* results (full lines) for H_2-H_2 are calculated at the CCSD(T)/AV5Z level of theory. The interactions of H_2 with the halogens, denoted by full lines, are of MP2/(H,F,Cl: AVQZ, Br: ECP10MDF AVQZ) quality. Triangles at the maximal CCSD(T)/CBS binding interactions are included as reference points to judge the accuracy of the MP2 results. Dashed lines show hard sphere quadrupole–quadrupole model interactions. Additionally, the magnitude rotational constant of H_2 is depicted by a double arrow. This figure was published in reference [139].



(b) Basis set dependence of relative deviations of MP2 well depths D_e with respect to CCSD(T)/CBS values for the structures shown in figure 3.3a. This figure was published in reference [139].

Figure 3.6

3.3 Rotationally adiabatic potential energy surfaces

So far, we have described the interaction of H_2 with F_2, Cl_2 and Br_2 for static structures, *i.e.* regardless of the nuclear spin states of the H_2 molecule. In order to describe the same interactions in terms of the nuclear spin isomers pH$_2$ and oH$_2$, we have to adiabatize the H_2-X_2 potentials with respect to the rotational degrees of freedom of H_2. The rotational adiabatization (RA) is justified by the fact that the rotational constant of H_2 ($B_0 = 0.6812$ kJ/mol) is 62, 244, and 724 times larger than the rotational constants of $^{19}F_2$ ($B_0 = 0.0106$ kJ/mol), $^{35}Cl_2$ ($B_0 = 0.0028$ kJ/mol) and $^{79}Br_2$ ($B_0 = 9.82 \cdot 10^{-4}$ kJ/mol), respectively. This implies, that the lighter H_2 molecule rotates much faster in comparison to the heavier halogen molecules. In the isotropic averaging technique, the influence of the intermolecular interactions on the rotational states is neglected and free rotor states are used to obtain respective averages over the interaction potential. Comparing the rotational constant of H_2 in the vibrational ground state $B_0 = 0.6812$ kJ/mol with the pair potentials ΔE in figure 3.6a, we see, however, that the well depths are up to 2, 3.5, and 4.7 times larger than B_0. Hence, especially near the minima of the potentials, we expect the rotation of H_2 to be notably perturbed. In this case, it is more accurate to consider the perturbation of rotational states explicitly by using the 'adiabatic hindered rotor' approximation [26]. For this we solve the rotational Schrödinger equation,

$$\begin{aligned}\hat{\mathbf{H}}_{\mathrm{rot}}(\theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2}; R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2}) \Psi_n(\theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2}; R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2}) = \\ W_n(R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2}) \Psi_n(\theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2}; R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2}),\end{aligned} \tag{3.13}$$

in the variables θ_{H_2} and ϕ_{H_2}, which are separated from the parameters, indicated by the semicolon, R, θ_{X_2} and ϕ_{X_2} by the different time and energy scales. The Hamiltonian

$$\hat{\mathbf{H}}_{\mathrm{rot}}(\theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2}; R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2}) = B_0 \ \hat{J}^2_{\mathrm{H}_2} + \Delta E(R, \theta_{\mathrm{X}_2}, \theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2} - \phi_{\mathrm{X}_2}), \tag{3.14}$$

describes the rotation of rigid H_2, with a rotational constant $B_0 = 0.6812$ kJ/mol and angular momentum operator $\hat{J}_{\mathrm{H}_2}$, perturbed by the potential ΔE due to the interaction with X_2. The solutions to equation 3.13 are rotational wave functions Ψ_n and the rotational energies $W_n(R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2})$.

To solve equation 3.13, we expand the wave function in the basis of the free rigid-rotor eigenfunctions, *i.e.* the spherical harmonics $Y_{JM}(\theta_{\mathrm{H}_2}, \phi_{\mathrm{H}_2})$. The eigenvalues $W_n(R, \theta_{\mathrm{X}_2}, \phi_{\mathrm{X}_2})$ and the respective eigenvectors with the expansion coefficients for Ψ_n, are obtained by diagonalization of the matrix representation of the Hamiltonian, *i.e.* eq. 3.13, in the truncated basis of free pH$_2$ states $Y_{J,M}$ with J=$\{0, 2, 4\}$ and separately in the truncated basis of the free oH$_2$ states $Y_{J,M}$ with J=$\{1, 3, 5\}$. Because the interaction potential ΔE vanishes for large R, the state with $n = 0$ is asymptotically correlated with the free rotor state $Y_{J,M} = Y_{0,0}$, *i.e.* the pH$_2$ ground state. Likewise, the states with n equal to 1, 2 and 3 are asymptotically correlated to the free rotor states $Y_{1,-1}$, $Y_{1,0}$ and $Y_{1,-1}$, which represent the triply degenerate rotational ground state of oH$_2$ with the energy $2B_0$. The difficult part in the set up of the

matrix representation is the evaluation of the potential energy matrix elements given by

$$\begin{aligned}\mathbf{\Delta E}_{(J'M'JM)}(R,\theta_{\mathrm{X}_2},\phi_{\mathrm{X}_2}) = \frac{1}{4\pi}\int_1^{-1} d\cos\theta_{\mathrm{H}_2}\int_0^{2\pi} d\phi_{\mathrm{H}_2} \qquad (3.15)\\ \Delta E(R,\theta_{\mathrm{X}_2},\theta_{\mathrm{H}_2},\phi_{\mathrm{H}_2}-\phi_{\mathrm{X}_2})\times\\ Y^*_{J'M'}(\theta_{\mathrm{H}_2},\phi_{\mathrm{H}_2})Y_{JM}(\theta_{\mathrm{H}_2},\phi_{\mathrm{H}_2})\\ \approx \frac{1}{4\pi}\sum_{k=1}^{8} w_k \sum_{j=1}^{16} w_j \times \qquad (3.16)\\ \Delta E(R,\theta_{\mathrm{X}_2},\theta_{\mathrm{H}_2,\mathrm{k}},\phi_{\mathrm{H}_2,j}-\phi_{\mathrm{X}_2,\mathrm{const.}})\times\\ Y^*_{J'M'}(\theta_{\mathrm{H}_2,k},\phi_{\mathrm{H}_2,j})Y_{JM}(\theta_{\mathrm{H}_2,k},\phi_{\mathrm{H}_2,j}).\end{aligned}$$

Here we have employed the Gaussian quadrature (GQ) technique to numerically compute the double integral over the product of the pair potential and two spherical harmonics. This involves the weights w_k and w_j and discrete values of ΔE at 8×16 Gauss-Legendre grid-points $(\theta_{\mathrm{H}_2,k},\ \phi_{\mathrm{H}_2,j})$, with ϕ_{X_2} held constant. The size of the GQ grid was tested by Accardi for H_2-Cl_2, using basis functions up to $J_{max} = 5$, who found that it is more than sufficient to obtain well converged solutions to the first eight eigenstates [27]. The interaction potential of the H_2-Br_2 dimer is not much stronger, compared to the H_2-Cl_2 interaction, while the H_2-F_2 interaction is weaker. Thus, we can use the same 8×16 GQ grid for these systems as well. Hence, each point of the rotational adiabatic potential is based on 128 MP2/AVQZ values for $\Delta E(R,\theta_{\mathrm{X}_2},\theta_{\mathrm{H}_2,\mathrm{k}},\phi_{\mathrm{H}_2,j}-\phi_{\mathrm{X}_2})$, or 384 single point calculations due to the CP correction. We performed linear scans over θ_{X_2}, ranging from 0° to 90°, with an increment of 5°. For each value of θ_{X_2}, a varying number of monomer separations R were evaluated. The smallest increment used in the R direction was 10 pm. Collectively the ΔE contain interaction energies for 50176, 87552 and 41600 individual structures for H_2-F_2, H_2-Cl_2 and H_2-Br_2, respectively. For $\theta_{\mathrm{X}_2} = 0°$, ΔE is of cylindrical symmetry, thus independent of ϕ_{H_2}, which reduces the number of non-equivalent Gauss–Legendre points from 128 to 8. For $\theta_{\mathrm{X}_2} = 90°$ the potential has C_{2v} symmetry, hence it is sufficient to evaluate one quarter of the sphere, which means 32 points in our case. Exploiting these symmetries, we also performed CCSD(T)/AVQZ level of theory scans along R for $\theta_{\mathrm{X}_2} = 0°$ and $\theta_{\mathrm{X}_2} = 90°$.

At large separations R, $W_0(R,\theta_{\mathrm{X}_2})$ converges to zero for free $p\mathrm{H}_2$, while $W_{1-3}(R,\theta_{\mathrm{X}_2})$ converge to the rotational energy $2B_0$ of free $o\mathrm{H}_2$. The interaction energies ΔW for the complexes are thus given by $\Delta W_0(R,\theta_{\mathrm{X}_2}) = W_0(R,\theta_{\mathrm{X}_2})$ for $p\mathrm{H}_2$ and $\Delta W_{1-3}(R,\theta_{\mathrm{X}_2}) = W_{1-3}(R,\theta_{\mathrm{X}_2}) - 2B_0$ for $o\mathrm{H}_2$.

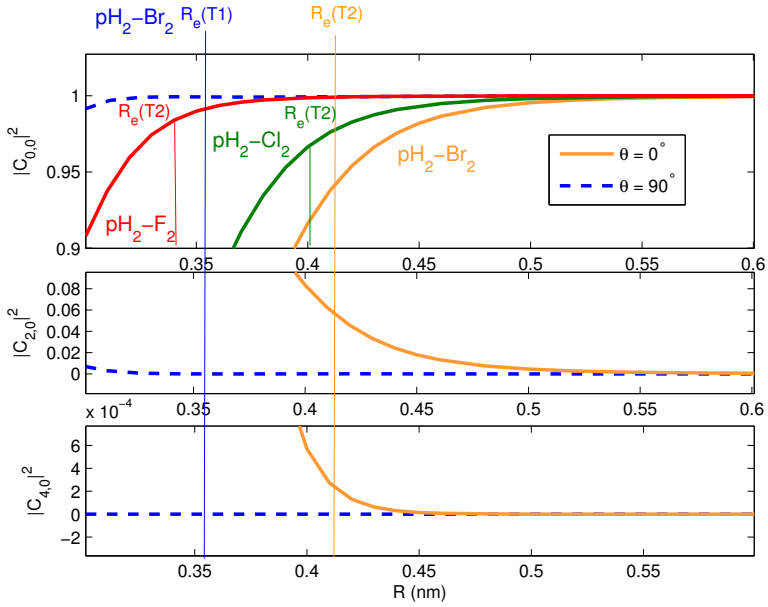


Figure 3.7: Change in the absolute squares of the expansion coefficients for $Y_{0,0}$, $Y_{2,0}$ and $Y_{4,0}$ of the pH_2 ground state Ψ_0 of pH$_2$-Br$_2$ with R for $\theta = 0°$ and $\theta = 90°$. Additionally for pH$_2$-F$_2$ and pH$_2$-Cl$_2$ the absolute squared expansion coefficients for $Y_{0,0}$ are given for $\theta = 0°$, to showcase the decrease in the perturbation for the lighter halogens.

3.3.1 Nuclear spin effect

At this point, we can discuss the RA interaction potentials $\Delta W_0(R, \theta_{X_2})$ for the nuclear spin isomers pH$_2$ and oH$_2$, and investigate the effect of the potential ΔE on the rotational wave function of pH$_2$, and on the probability densities $|\Psi_n|^2$. Since, the perturbation due to the intermolecular interaction is most pronounced for H$_2$-Br$_2$ we will focus our analysis on this system. To highlight this, the R dependence of the absolute squares of the expansion coefficients for the pH$_2$ ground state Ψ_0 for $\theta = 0°$ and $\theta = 90°$ are given in figure 3.7. Only the functions $Y_{0,0}$, $Y_{2,0}$ and $Y_{4,0}$ contribute to Ψ_0 for $\theta = 0°$, as shown for pH$_2$-Br$_2$. Even at the minimum of T2, Ψ_0 is almost equivalent to the free pH$_2$ ground state $Y_{0,0}$, as is evident by the value of $|C_{0,0}|^2(0.415\text{ nm}) = 0.95$. The very small contribution of $Y_{4,0}$ vanishes for $R > 0.45$ nm. For $R > 0.55$ nm also the contribution of $Y_{2,0}$ vanishes, so that Ψ_0 is equivalent to $Y_{0,0}$. Since the interactions of Cl$_2$ and F$_2$ are weaker and of shorter range, the $|C_{0,0}|^2$ values for Ψ_0 at $\theta = 0°$ are consequently smaller at the R_e for T2 and are effectively 1 for $R > 0.4$ nm, and $R > 0.5$ nm. In contrast to $\theta = 0°$, for $\theta = 90°$ in the pH$_2$-Br$_2$ system, Ψ_0 is effectively equivalent to $Y_{0,0}$ at the R_e for T1.

Let us now begin to examine the nuclear spin effect for the pH$_2$-Br$_2$ system, for two points (R, θ_{Br_2}) close to the global and local minima of the RA potentials ΔW_n. The specific point $\mathcal{L}$ near the global RA minima is given by $R = 420$ pm and $\theta_{Br_2} = 0°$, hence it represents a collinear orientation of the Br$_2$ bond axis and the COM of H$_2$. Hence, the 4D pair potential $\Delta E(\mathcal{L}) = \Delta E(420\text{ pm}, 0°, \theta_{H_2}, \phi_{H_2} - 45°)$, that perturbs the rotation of pH$_2$ and oH$_2$, only depends on θ_{H_2}, see figure 3.8. The minimum of the potential is given by the static T2

structure, $\theta_{H_2} = 90°$, and the maximum is given by L, $\theta_{H_2} = 0°$ and $\theta_{H_2} = 180°$, with intermediate energies in between. The rotational density $|\Psi_0|^2$ of the pH$_2$ ground state ($n = 0$) is given in figure 3.8.

From the analysis of the expansion coefficients we know, that the functions $Y_{2,0}$ and $Y_{4,0}$ are coupled to $Y_{0,0}$. Due to the interaction, the density is not exactly spherical, but slightly disturbed along θ_{H_2}. Probability density is added in the T2 region and lost in the L region. In other words, pH$_2$ is more aligned to the energetically favourable T2 structure and less to the L structure. The resulting RA interaction energy for pH$_2$ with Br$_2$ at $\mathcal{L}$ is $\Delta W_0 = -2.3$ kJ/mol. Hence, it is much lower in magnitude compared to the global minimum of ΔE with -3.2 kJ/mol. This can be explained by the quantum mechanical averaging of the energies of static structures, between the extremes of the strongly binding (T2) and weakly binding (L) structure due to the zero point rotations of the pH$_2$ molecule. Further, it can be shown that the quadrupole-quadrupole interactions vanish for $Y_{0,0}$ [27].

Next, we turn to the interactions of oH$_2$ with Br$_2$ for the same point, by inspecting the solutions for $n = 1,\ 2,\ 3$, given in figure 3.8. The interaction energies for $n = 1,\ 2$ are degenerate with $\Delta W_{1,2} = -2.64$ kJ/mol. Thus the ground state of oH$_2$ interacts stronger with Br$_2$ than does oH$_2$ in the first excited state ($\Delta W_3 = -1.34$ kJ/mol). Inspection of the respective densities shows, that $|\Psi_1|^2$ and $|\Psi_2|^2$ are degenerate in θ_{H_2} but not in ϕ_{H_2}, since $\Delta E(\mathcal{L})$ is independent of ϕ_{H_2} both solutions are degenerate in energy. Moreover, both oH$_2$ densities are aligned to the T2 structure, thereby avoiding the unfavourable L structure, which leads to a stronger interaction with Br$_2$. Likewise, $|\Psi_3|^2$ is aligned to the L structure, resulting in a weaker RA interaction energy.

The second point, denoted by $\mathcal{T}$, for which we repeat this analysis, is close to the local minima of the RA potentials. This point is given by $R = 354$ pm and $\theta_{Br_2} = 90°$, and includes the structures, where the Br$_2$ bond axis and the COM of H$_2$ form a T-shape. Figure 3.8 shows the potential $\Delta E(\mathcal{T}) = \Delta E(354\text{ pm}, 90°, \theta_{H_2}, \phi_{H_2} - 45°)$, and rotational probability distributions and RA interaction energies for $n = 0,\ 1,\ 2,\ 3$. At this point ΔE is weaker and more isotropic, as it is described by the energetically similar T1, H and X structures. The probability distribution of the pH$_2$ ground state $|\Psi_0|^2$ is effectively spherical, since its alignment to the T1 structure would imply a node in the wave function, only possible by a mixing of higher angular momentum states, that would require kinetic energies in the order of $6B_0$. The rotational states of oH$_2$ lose their degeneracy for the T-shaped orientations near Br$_2$, as $\Delta E(\mathcal{T})$ depends on both θ_{H_2} and ϕ_{H_2}. Again, the ordering of the oH$_2$ interaction energies can be explained by the alignment of $|\Psi_n|^2$ to more interacting structures, *i.e.* $|\Psi_1|^2$ aligns to T1, $|\Psi_2|^2$ to H and $|\Psi_3|^2$ to X. In summary, part of the rotational wave functions of oH$_2$ show alignment to structures with larger binding energies, while for the effectively spherical pH$_2$ distribution less binding structures enter the rotational averaging of the interaction energy with an increased weight. This leads to a stronger interaction for the rotational ground state of oH$_2$ with X$_2$, compared to the rotational ground state of pH$_2$.

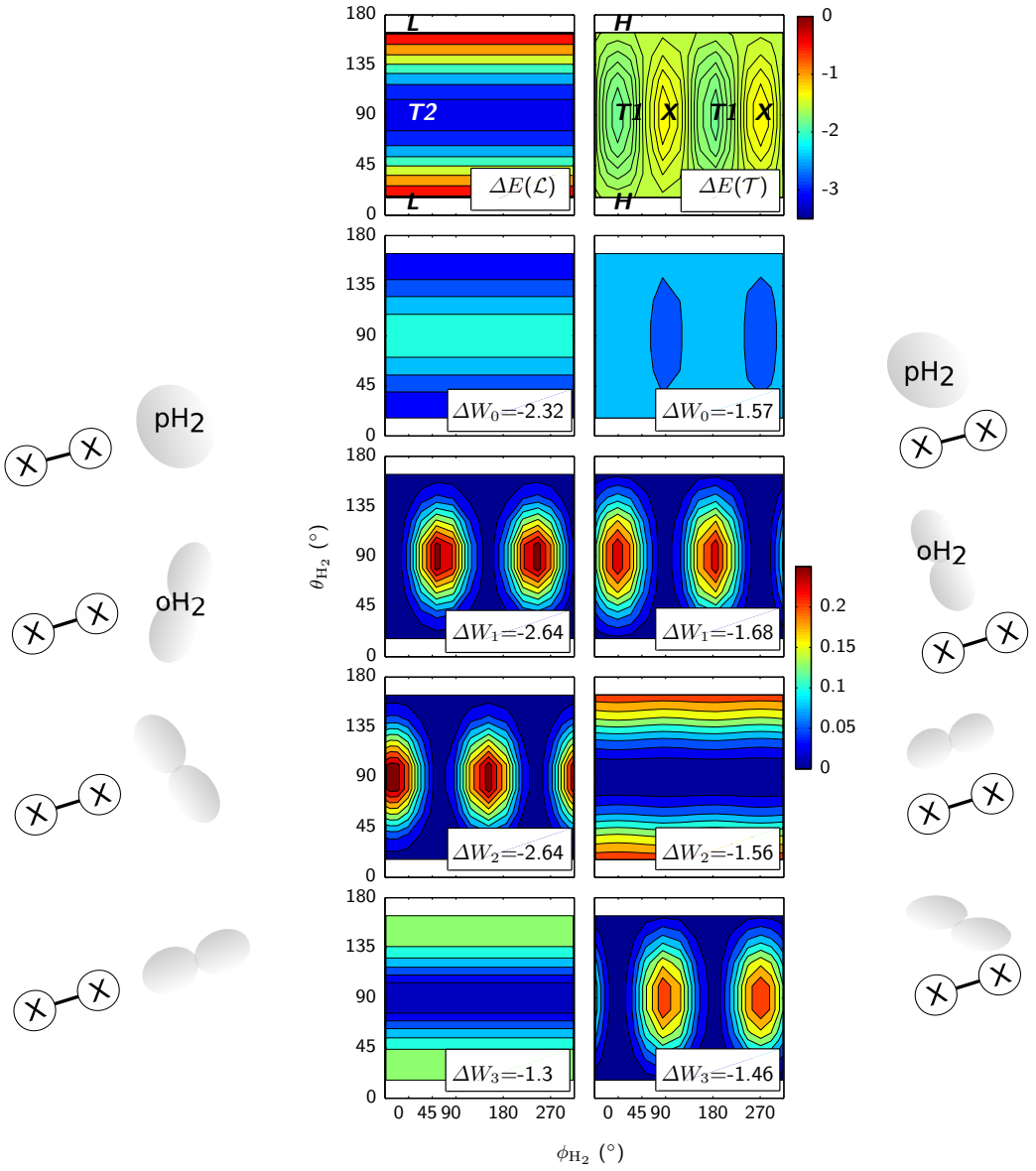


Figure 3.8: Potentials for H_2–Br_2 at the MP2/AVQZ level of theory (1st row, in kJ/mol) and the rotational densities for pH_2 (2nd row) and oH_2 (3rd-5th row), in the linear minimum (left panel, $\theta_{Br_2} = 0°$, $R = 420$ pm) and the T-shape minimum (right panel, $\theta_{Br_2} = 90°$, $R = 354$ pm) of the RA potential of H_2-Br_2. The position of high symmetry structures is indicated by T1, T2, ... see Fig. 3.3a. The central figure was published in reference [139]. Sketches of the rotational densities are given to the left and right of the respective density plots, to highlight the mutual orientation of the pH_2 and Br_2 molecules. Note that the size of the hydrogen molecule has been greatly exaggerated in these sketches.

3.3.2 Potential energy surfaces for the interaction of $p/o\mathrm{H}_2$ with X_2

In this part, we compare the two-dimensional potential energy surfaces $\Delta W_n(R, \theta_{\mathrm{X}_2})$ for the interaction with X_2 of $p\mathrm{H}_2$ and $o\mathrm{H}_2$ in the rotational ground states with $n = 0$ and $n = 1$, respectively. Remember, that ΔW_n were obtained by rotational adiabatization of the four-dimensional MP2/AVQZ pair potentials. In general, the RA potentials, given figure 3.9, share similar features. Each potential has two minima, the global minimum at $\theta_{\mathrm{X}_2} = 0°$ and the local minimum at $\theta_{\mathrm{X}_2} = 90°$. As discussed previously, the global minimum corresponds to the linear $p\mathrm{H}_2$-X_2 structure, involving an averaging over the static T2 and L structures. Whereas the local minimum corresponds to a T-shaped $p\mathrm{H}_2$-X_2 structure, where the averaging involves the static T1, H and X structures. Both, minima are connected via a saddle point near $\theta_{\mathrm{X}_2} = 45°$. The respective D_e and R_e values are given in table 3.2. The global minima range from -0.93 kJ/mol for $p\mathrm{H}_2$-F_2 to -2.64 kJ/mol for $o\mathrm{H}_2$-Br_2. The local minima range from -0.55 kJ/mol for $p\mathrm{H}_2$-F_2 to -1.68 kJ/mol for $o\mathrm{H}_2$-Br_2. The ranges for the saddle points go from -0.41 kJ/mol to -0.95 kJ/mol. For a given halogen molecule the interaction with $o\mathrm{H}_2$ ($n = 1$) is always stronger than for $p\mathrm{H}_2$ ($n = 0$). The para-ortho splittings at the global minima are 0.15, 0.26 and 0.31 kJ/mol for F_2, Cl_2 and Br_2, respectively. At the local minima the para-ortho splittings are smaller in comparison, with 0.06, 0.10 and 0.11 kJ/mol, due to the lower anisotropy of the underlying pair potentials. At the saddle points the para-ortho splittings are with 0.08, 0.10 and 0.17 kJ/mol sightly larger in comparison to the values at local minima, but still lower when compared to the splittings at the global minima. Inspection of the minimum energy paths along θ_{X_2}, given in figure 3.10, shows that the para-ortho splittings along these paths have a maximum at $\theta_{\mathrm{X}_2} = 0°$ and are basically constant from $\theta_{\mathrm{X}_2} = 5°$ onwards. Figure 3.10 also highlights the fact, that the RA potentials for Cl_2 and Br_2 are more similar to each other and have a much higher anisotropy, in comparison to the respective potential for F_2.

For $\theta_{\mathrm{X}_2} = 0°$ and $\theta_{\mathrm{X}_2} = 90°$ CCSD(T)/AVQZ RA potentials have been obtained to counter check the MP2/AVQZ results. The respective D_e and R_e values are given in table 3.2. Comparing the CCSD(T) to the MP2 values for D_e shows that RA interactions at $\theta_{\mathrm{X}_2} = 0°$ are overestimated by MP2 by 0.1 to 0.2 kJ/mol. For $\theta_{\mathrm{X}_2} = 90°$, MP2 underestimates D_e by 0.1 kJ/mol for the $p/o\mathrm{H}_2$-F_2 complex and overestimates D_e by about 0.05 kJ/mol for $p/o\mathrm{H}_2$-Cl_2 and $p/o\mathrm{H}_2$-Br_2. The biggest relative error in the MP2 results is the overestimation of the already low energetic anisotropy of the $p/o\mathrm{H}_2$-F_2 potentials. In view that CCSD(T)/AVQZ underestimates the interaction energy compared to CCSD(T)/CBS, the MP2 results can be considered to be reliable.

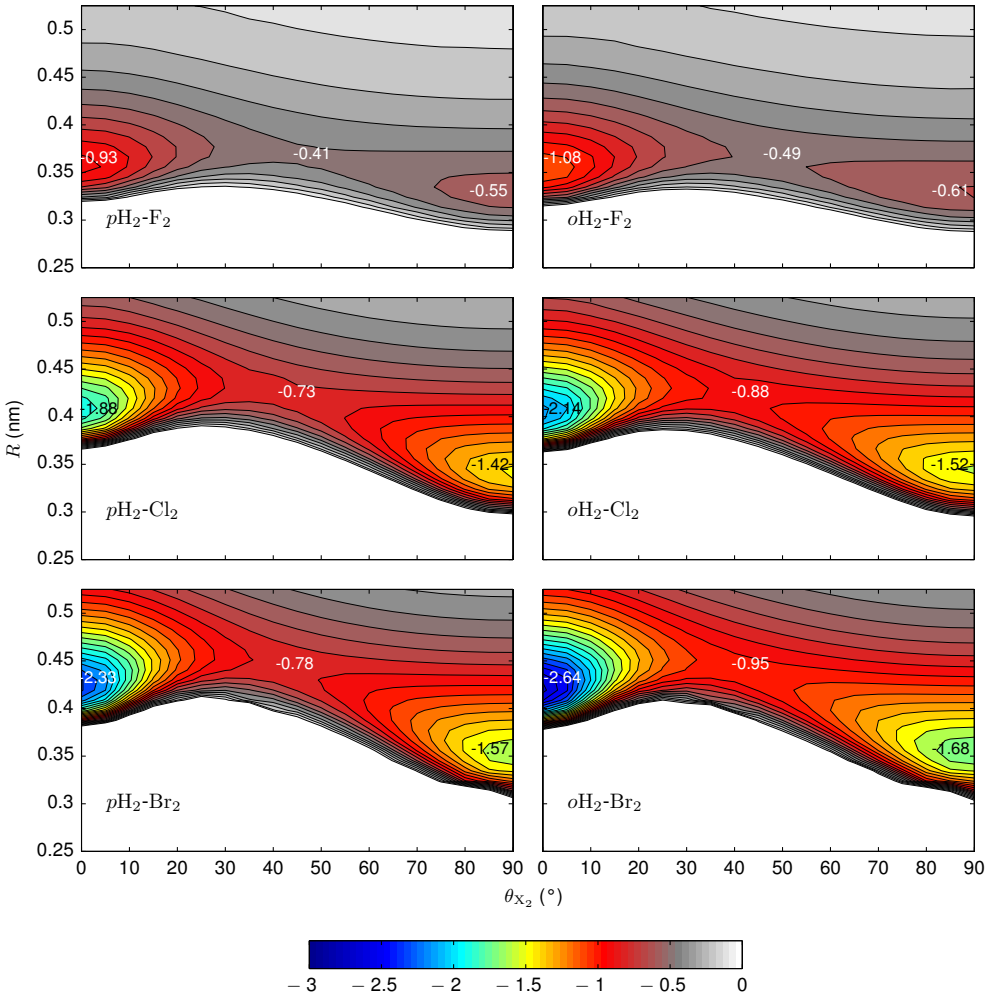


Figure 3.9: Rotationally adiabatic MP2/AVQZ potentials for the interaction ΔW of para- and ortho-hydrogen, in their rotational ground states, with fluorine, chlorine and bromine molecules. Contour levels for ΔW, ranging from -3 to 0 kJ/mol, are shown at 0.1 kJ/mol increments. Figure was published in reference [139].

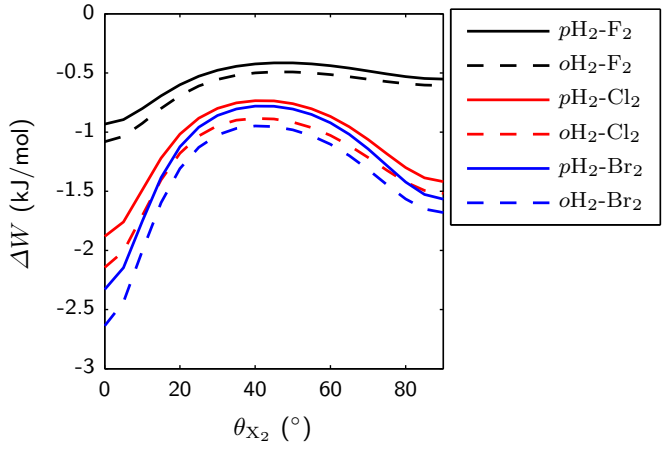


Figure 3.10: Minimum energy paths along θ_{X_2} for the RA potentials shown in figure 3.9. Figure was published in reference [139].

3.3.3 Fit to analytic function

Analytic expressions of potentials facilitate the future use in simulations. The rotationally adiabatic MP2 pair potentials ΔW_n were fitted to the Hartree–Fock–dispersion (HFD) function [146].

$$\begin{aligned}\Delta W^{\mathrm{HFD}}(R,\theta_{\mathrm{X}_2}) = a(\theta_{\mathrm{X}_2})\exp[-(b(\theta_{\mathrm{X}_2})R + c(\theta_{\mathrm{X}_2})R^2] \\ -S(R)\left(\frac{C_6(\theta_{\mathrm{X}_2})}{R^6} + \frac{C_8(\theta_{\mathrm{X}_2})}{R^8} + \frac{C_{10}(\theta_{\mathrm{X}_2})}{R^{10}}\right),\end{aligned} \tag{3.17}$$

where the first term represents the, in total repulsive, sum of the Pauli-repulsion, electrostatic and induction interactions, whereas the second term models the attractive dispersion interaction. The dispersion terms decay algebraically, hence a switching function

$$S(R) = \begin{cases} \exp[-(1.28R_e(0^\circ)/R - 1)^2], & \text{if } R < 1.28R_e(0^\circ) \\ 1, & \text{if } R \geq 1.28R_e(0^\circ), \end{cases} \tag{3.18}$$

which truncates the dispersion energy toward short internuclear distances, becomes necessary. The shape of the switching function and the parameter 1.28 originates from a fit to the potential of the $^3\Sigma_u^+$ state of H_2, for witch both terms in equation 3.17 without $S(R)$ are known [146]. In order to decrease the number of parameters of the original HFD function, equation 3.17, we performed fits of 1D cuts for for fixed θ_{X_2} using different parameter sets. We found, that we could omit the R^2 dependence in the repulsion exponent and the R^{-10} dependence of the original dispersion term. Neglect of the R^{-8} dependence also resulted in good 1D fits. However, for the 2D fits, the R^{-8} dependence was necessary in order to increase the flexibility of the HFD term. Thus, the R^{-8} terms was kept. To further reduce the number of parameters, we decided to neglect the angular dependence of the switching function S, by making it only dependent on the position $R_e(0^\circ)$ of the global minimum of the $p\mathrm{H}_2$-X_2 potential. This leaves four angular dependent parameters a, b, C_6 and C_8. The modified HFD function used in this work thus reads,

$$\Delta W_n^{\mathrm{modHFD}}(R,\theta_{\mathrm{X}_2}) = a(\theta_{\mathrm{X}_2})\exp[-b(\theta_{\mathrm{X}_2})R] - S(R)\left(\frac{C_6(\theta_{\mathrm{X}_2})}{R^6} + \frac{C_8(\theta_{\mathrm{X}_2})}{R^8}\right). \tag{3.19}$$

The angular dependence of these parameters is represented by linear expansions in the first six even Legendre polynomials (LP) P_{2k} given by,

$$X(\theta_{\mathrm{X}_2}) = \sum_{k=0}^{5} X_{2k}P_{2k}(\cos\theta_{\mathrm{X}_2}), \tag{3.20}$$

where X_0 to X_{10} are the six corresponding expansion coefficients. Hence, each modified HFD potential depends on 25 parameters, of which one ($R_e(0^\circ)$) is constant and determined prior to the fitting procedure. The use of the HFD function and the expansion in even-ordered Legendre polynomials is motivated by the fit of the $p\mathrm{H}_2$-Cl_2 potential by Accardi [27]. Additionally, the HFD function has been designed explicitly to describe the similar weak rage gas interactions [146]. Here, we employ a HFD function, which in comparison to the one in reference [27] has less physically motivated parameters, but adds angular flexibility in the re-

maining parameters due to larger LP expansions. An flexibility that is clearly necessary in view of the angular anisotropy of the RA potentials, see figures 3.10 and 3.9.

In order to simultaneously optimize the remaining 24 parameters, we used to following strategy, which was adapted from [27]. First, we obtained a, b, C_6 and C_8 values for 1D cuts along θ_{X_2}, by minimizing the root mean square deviation (RMSD) of the fitfunction $\Delta W_n^{\mathrm{modHDF}}$ with respect to the tabulated *ab inito* RA potential ΔW_n for a fixed θ_{X_2}. For each parameter, the values for $\theta_{X_2} = 0°$, 15°, 45°, 60°, 75° and 90°, were used individually to generate an initial guess for each of the six expansion coefficients. The resulting initial angular dependence for each parameter is given in figures 3.11-3.13, by dotted lines. Then, fits of the 2D RA potentials were performed by nonlinear optimization of the expansion coefficients, under the constraint of positive a, b, C_6 and C_8 values, by minimization of the RMSD of $\Delta W_n^{\mathrm{modHDF}}$ with respect to ΔW_n. For the nonlinear optimization the 'fminsearch' routine of Matlab (version R2011a) was used. The resulting angular dependence of the parameters a, b, C_6 and C_8 is shown in figures 3.11-3.13. The repulsion scaling parameter a behaves similar to a step function, being maximal for small θ_{X_2} angles and small at larger angles. Parameter b has nearly constant values around 2 $1/a_0$. The angular dependence of C_6 shows a greater variety and is often coupled to that of C_8. For small θ_{X_2} where C_8 becomes more important, C_6 has lower values in many cases.

The functions $\Delta W_n^{\mathrm{modHDF}}$ give smooth and accurate representations of the RA potentials as shown in the contour plots for $\Delta W_n^{\mathrm{modHDF}}$ also given in the figures 3.11-3.13. The contour plots also show all RA potentials that are asymptotically connected to the free rotor ground state of $o\mathrm{H}_2$.

Additionally, 1D cuts of the final $\Delta W_n^{\mathrm{modHDF}}$ potential and the original data points are given for $\theta_{X_2} = 0°$, 45° and 90°. This highlights the overall excellent quality of the obtained fits. Indeed, the global minima are reproduced with maximal deviations of 2%. The fit parameters for $\Delta W_{n=0,1,2,4}^{\mathrm{modHDF}}$ were made available in the supplementary material of reference [139].

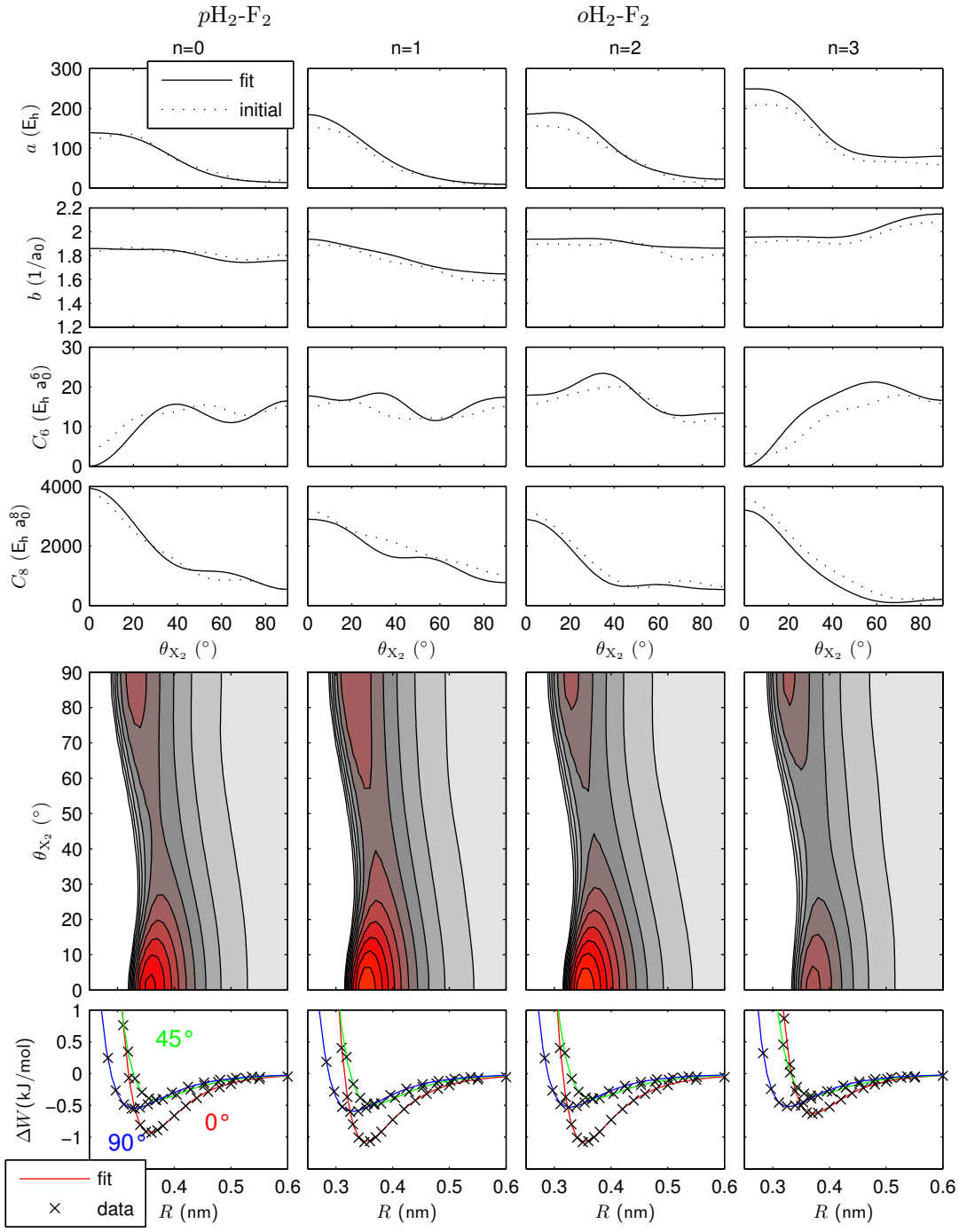


Figure 3.11: H_2-F_2: The figure shows initial and final fit parameters, contour plots and cuts of the RA potentials $\Delta W^{\text{modHDF}}_{n=0,1,2,4}$ overlaid with original data points of $\Delta W_{n=0,1,2,4}$. Contour lines and color scale as in figure 3.9.

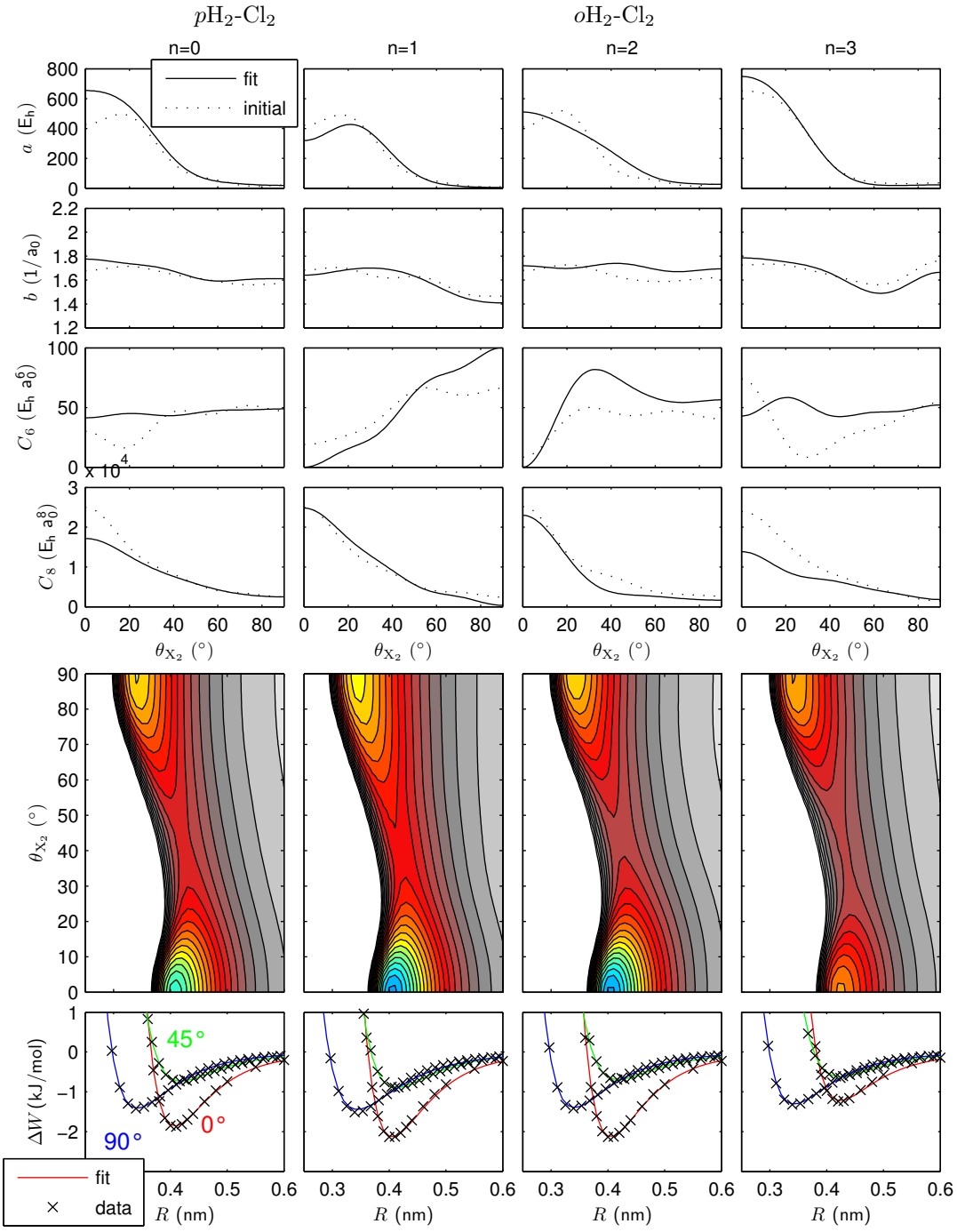


Figure 3.12: H_2-Cl_2: The figure shows initial and final fit parameters, contour plots and cuts of the RA potentials $\Delta W^{\text{modHDF}}_{n=0,1,2,4}$ overlaid with original data points of $\Delta W_{n=0,1,2,4}$. Contour lines and color scale as in figure 3.9.

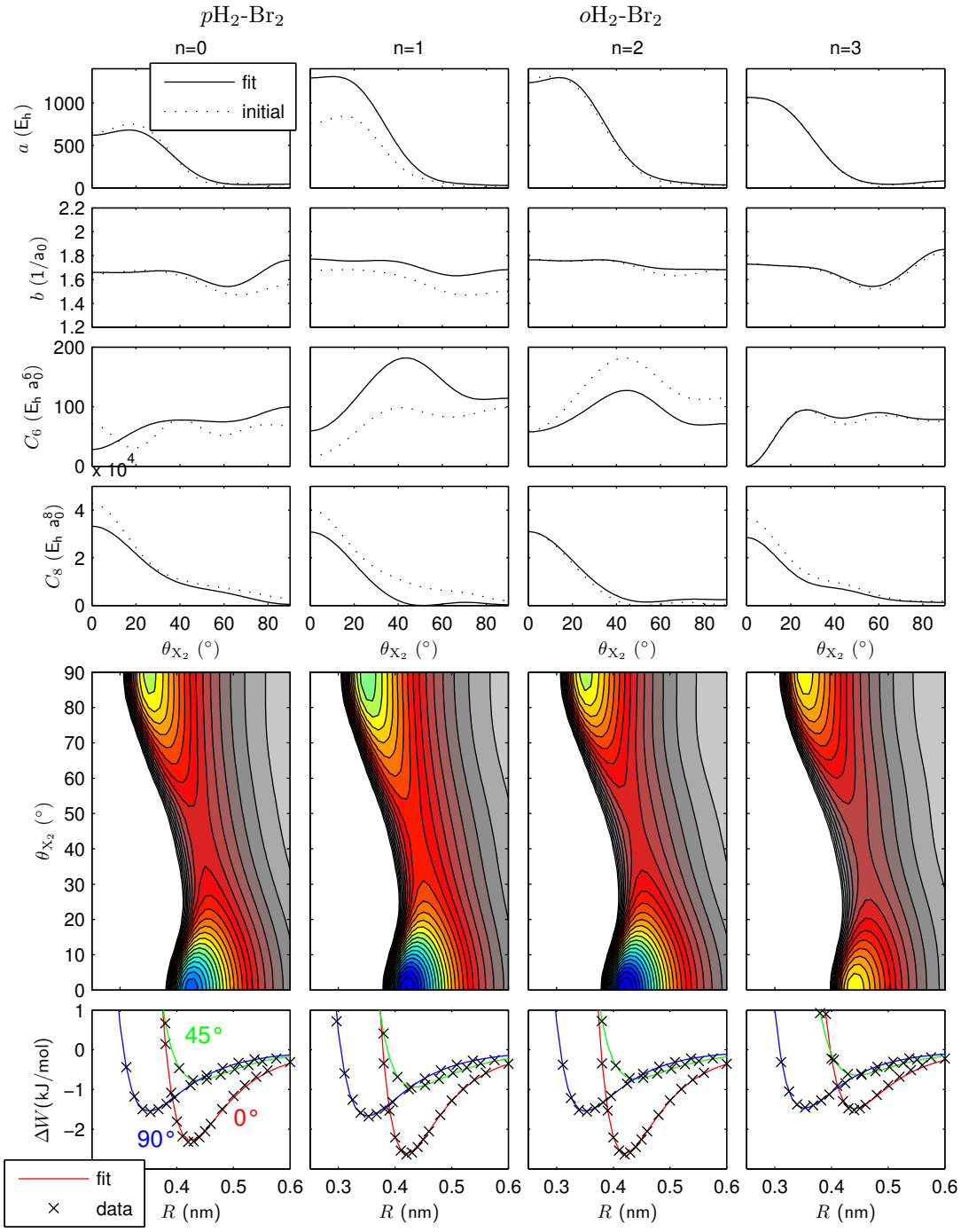


Figure 3.13: H_2-Br_2: The figure shows initial and final fit parameters, contour plots and cuts of the RA potentials $\Delta W^{\text{modHDF}}_{n=0,1,2,4}$ overlaid with original data points of $\Delta W_{n=0,1,2,4}$. Contour lines and color scale as in figure 3.9.

3.3.4 Comparison of interactions

Now, that we have determined the lowest RA pair interactions of $p/o\mathrm{H_2}$-$\mathrm{X_2}$ ($n = 0$, $n = 1$), we can compare them to the interactions present in similar systems. First, we want to take into account the interaction of rage gas (Rg) atoms (He, Ne and Ar) with dihalogens, *i.e.* the Rg-$\mathrm{X_2}$ pair potentials for $\mathrm{X_2} = \{\mathrm{F_2}, \mathrm{Cl_2}, \mathrm{Br_2}\}$. Chan *et al.* calculated the well depths of the Rg-$\mathrm{F_2}$ potentials at the CCSD(T)/(aug-cc-pVQZ+bf) level of theory, where the atom centred basis set was augmented by bond functions (bf), *i.e.* functions located in-between the rare gas atom and the $\mathrm{F_2}$ molecule [147]. The minima of the respective pair potentials are found for linear and T-shaped structures and are of similar depth, with the global minima assigned to the linear structures. In terms of D_e and R_e the $p/o\mathrm{H_2}$-$\mathrm{F_2}$ potential is most similar to that of Ne-$\mathrm{F_2}$. Comparing the well depths of the linear structures, we find that $p/o\mathrm{H_2}$ fits in-between Ne and Ar. The full ordering of the D_e values for the linear structures is therefore: He < $p\mathrm{H_2}$ < $o\mathrm{H_2}$ < Ne < Ar. For the T-shaped structures, the well depth for $p/o\mathrm{H_2}$ is again similar to Ne but effectively smaller, which may be related to the fact that Ne can approach $\mathrm{F_2}$ closer $R_e(\mathrm{Ne}) = 308$ pm vs. $R_e(p\mathrm{H_2}) = 322$ pm. The following ordering of T-shape well depths results: He < Ne < $p\mathrm{H_2}$ < $o\mathrm{H_2}$ < Ar, for the interaction with $\mathrm{F_2}$, see also figure 3.14. In summary, the interaction of $\mathrm{F_2}$ with $p\mathrm{H_2}$ is twice as large in comparison to He, it is very similar to Ne and weaker than that with Ar.

Changing to the next halogen $\mathrm{Cl_2}$, we compare our results to the He-$\mathrm{Cl_2}$ and Ne-$\mathrm{Cl_2}$ MP4 and Ar-$\mathrm{Cl_2}$ CCSD(T) well depths of Williams *et al.* [148], who used large well tempered basis sets by Huzinaga and Klobukowski augmented by polarization and bond functions [148]. For $\mathrm{Cl_2}$ the interaction with $p\mathrm{H_2}$ is stronger than with He and Ne, for both linear and T-Shaped structures. The ordering of absolute D_e values is thus, He < Ne < $p\mathrm{H_2}$ < $o\mathrm{H_2}$ < Ar, see also figure 3.14. Interestingly, the R_e values of $p\mathrm{H_2}$-$\mathrm{Cl_2}$ are slightly smaller than for He-$\mathrm{Cl_2}$. Again, the well depths of the interactions with rare gas atoms are almost identical for both structures, while for $p\mathrm{H_2}$ the highest interaction is clearly present for the linear structure. The potentials for the Rg-$\mathrm{Br_2}$ complexes have been investigated by Prosmiti et al. [149] at the CCSD(T) level of theory. For Br they used the ECP from the Stuttgart group and the related SDD basis set, augmented with (sp) diffusion and (3df) polarization functions. For He the aug-cc-pV5Z and for Ne and Ar the respective aug-cc-pVQZ basis set where employed. Additionally, (3s3p2d2f1g) bond functions were placed in the center of the van der Waals bond. When we compare the well depths for the linear and T-shape complexes of $p/o\mathrm{H_2}$-$\mathrm{Br_2}$ to those of Rg-$\mathrm{Br_2}$, we find the same ordering as previously for $\mathrm{Cl_2}$, see also figure 3.14. For the linear structure, the CCSD(T) values for R_e are also considerably smaller in the case of $p\mathrm{H_2}$-$\mathrm{Br_2}$ (428 pm) in comparison to He-$\mathrm{Br_2}$ (442 pm).

Next, we compare the interaction of $p\mathrm{H_2}$ with dihalogens to that of $p\mathrm{H_2}$ with other small molecules, $p\mathrm{H_2}$ [26, 150], CO [26, 151], $\mathrm{CO_2}$[26, 152], $\mathrm{N_2O}$ [153], HCN [154, 155], OCS [156, 157] and $\mathrm{H_2O}$ [158, 159]. The minima of the respective potential energy surfaces given in the literature are compiled in table 3.2. For the linear heteronuclear diatomic or triatomic molecules without a center of inversion, three minima for the interaction with $p\mathrm{H_2}$ are present. Two linear minima and one distorted T-shape minimum. Only for the $p\mathrm{H_2}$-HCN complex a linear structure corresponds to the global minimum. The global minima of all other complexes corresponds to a T-shaped ($\mathrm{CO_2}$) or distorted T-shaped structure. Concerning the linear minima, $p\mathrm{H_2}$ interacts stronger with the atom of higher atomic number, *i.e.* for $p\mathrm{H_2}$-CO and

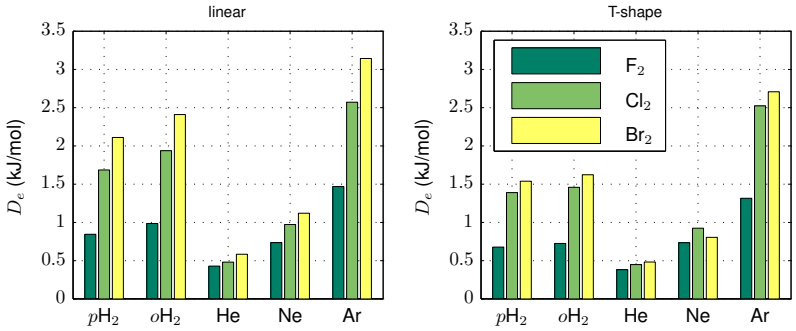


Figure 3.14: Comparison of well depths for the interaction of halogens with para- and orthohydrogen (CCSD(T)/AVQZ) from this work, and with rare gas atoms, from references [147–149]. The values are also given in table 3.2. Details on the levels of theory (CCSD(T) except for He-Cl_2, Ne-Cl_2 MP4) are given in subsection 3.3.4.

pH$_2$-N$_2$O the linear well near O is deeper compared to that near C and N. For HCN and OCS the linear well near N and S is deeper, compared to the well near H and O. The interaction between two pH$_2$ molecules is just -0.295 kJ/mol, whereas the highest interaction energy in this set is present for pH$_2$-OCS with -1.729 kJ/mol. Comparing the interactions with pH$_2$ including also our RA MP2/AVQZ results for the halogens, the ordering from weak to strong reads: $p\mathrm{H_2} < \mathrm{CO} < \mathrm{F_2} < \mathrm{HCN} < \mathrm{H_2O} < \mathrm{CO_2} < \mathrm{N_2O} < \mathrm{OCS} < \mathrm{Cl_2} < \mathrm{Br_2}$. This ordering follows the well known trend that the attractive interactions with closed shell systems with more electrons are stronger due to an increase in the dipole polarizabilites and in turn the attractive dispersion interactions.

			Linear ($\theta_{X_2} = 0°$)		T–shape ($\theta_{X_2} = 90°$)	
Complex	Ref.	Method	D_e	R_e	D_e	R_e
pH$_2$-F$_2$		CCSD(T)	0.845	361	0.678	322
		MP2	0.932	357	0.552	328
oH$_2$-F$_2$		CCSD(T)	0.986	356	0.725	325
		MP2	1.081	353	0.606	330
He-F$_2$	[147]	CCSD(T)	0.429	347	0.382	300
Ne-F$_2$	[147]	CCSD(T)	0.736	359	0.735	308
Ar-F$_2$	[147]	CCSD(T)	1.469	388	1.316	344
pH$_2$-Cl$_2$		CCSD(T)	1.686	412	1.390	343
		MP2	1.882	408	1.419	342
oH$_2$-Cl$_2$		CCSD(T)	1.938	409	1.459	345
		MP2	2.144	404	1.518	343
He-Cl$_2$	[148]	MP4	0.481	420	0.449	345
Ne-Cl$_2$	[148]	MP4	0.973	427	0.925	350
Ar-Cl$_2$	[148]	CCSD(T)	2.572	448	2.524	374
pH$_2$-Br$_2$		CCSD(T)	2.111	428	1.539	353
		MP2	2.330	424	1.567	352
oH$_2$-Br$_2$		CCSD(T)	2.410	425	1.623	356
		MP2	2.638	421	1.682	354
He-Br$_2$	[149]	CCSD(T)	0.584	442	0.482	358
Ne-Br$_2$	[149]	CCSD(T)	1.120	449	0.805	360
Ar-Br$_2$	[149]	CCSD(T)	3.143	463	2.708	380

			Linear				T–shape	
Complex	Ref.	Method	D_e	R_e	D_e	R_e	D_e	R_e
pH$_2$-pH$_2$	[26, 150]	CCSD(T)	0.295	346				
pH$_2$-CO	[26, 151]	CCSD(T)	$\theta_1 = 0°$(C)		$\theta_1 = 180°$(O)		$\theta_1 = 85.8°$	
			0.443	440	0.514	400	0.610	360
pH$_2$-CO$_2$	[26, 152]	CCSD(T)			$\theta_1 = 180°$		$\theta_1 = 90.0°$	
					0.667	444	1.340	319
pH$_2$-N$_2$O	[153]	CCSD(T)	$\theta_1 = 0°$ (N)		$\theta_1 = 180°$ (O)		$\theta_1 = 92.58°$	
			0.60	460	0.821	429	1.691	308
oH$_2$-N$_2$O	[153]		$\theta_1 = 0°$ (N)		$\theta_1 = 180°$		$\theta_1 = 92.88°$	
			0.86	460	1.02	430	2.09	302
pH$_2$-HCN	[154, 155]	CCSD(T)-F12a	$\theta = 0°$ (H)		$\theta = 180°$(N)		$\theta \approx 60°$	
		spherical average	≈ 0.66	410	0.948	430	0.750	407
pH$_2$-OCS	[156, 157]	MP4	$\theta = 0°$ (S)		$\theta = 180°$ (O)		$\theta = 105°$	
		spherical average	1.093	452	0.828	492	1.729	335
pH$_2$-H$_2$O	[158, 159]	CCSD(T),			$\theta = 110.00°$, $\chi = 0°$			
		CCSD(T)-R12			1.167	336		

Table 3.2: *Ab initio* well depths D_e in kJ/mol and equilibrium distances R_e in pm, for the p/oH$_2$-X$_2$ and Rg-X$_2$ complexes with X = F, Cl, Br; Rg = He, Ne, Ar as well as small molecules from the literature. This table was published in reference [139].

3.4 Bound state calculations for $p\mathrm{H_2}$-$\mathrm{X_2}$ van der Waals dimers

For the comparison with spectroscopic data, the well depths D_e are of limited use, as only the dissociation energy D_0 of the vibrational ground state can be measured. Since H_2 is a very light molecule, we expect the zero point energy (ZPE), *i.e.* the difference between D_0 and D_e to be significant.

While no experimental data on H_2-F_2, H_2-Cl_2 and H_2-Br_2 is (to our knowledge) available, Darr *et al.* and Kenny *et al.* reported measurements on H_2-I_2 [124, 125]. The MP2/AVTZ+bf well depths for the T2, L, and T1 structures are 3.11 kJ/mol, 1.71 kJ/mol and 1.95 kJ/mol [124]. The ZPE contribution can be estimated by comparison of the calculated well depths to the measured ground state binding energy D_0, which is 1.23 kJ/mol for the linear $p\mathrm{H_2}(J{=}0)$-$\mathrm{I_2}$ complex [124]. The linear $o\mathrm{H_2}(J{=}1)$-$\mathrm{I_2}$ dimer has a binding energy of 1.42 kJ/mol [124]. It is stronger bound than the $p\mathrm{H_2}(J{=}0)$-$\mathrm{I_2}$ dimer. The binding energy of the T-shaped $o\mathrm{H_2}(J{=}1)$-$\mathrm{I_2}$ dimer has been measured to lie between 1.09 and 1.12 kJ/mol [125]. The estimated ZPE contribution to the binidng energy of the linear $p\mathrm{H_2}(J{=}0)$-$\mathrm{I_2}$ dimer relative to the T2 well depth is thus 60%. For the present three dimers we expect at least similar ZPEs.

In this chapter, we determine binding energies D_0 and whole bound state energy spectra for the $p\mathrm{H_2}$ species of the three van der Waals dimers in their rotational ground states. For this, we solve the time independent Schrödinger equation for the nuclear motion on the RA MP2/AVQZ potentials numerically, as described in the following paragraph.

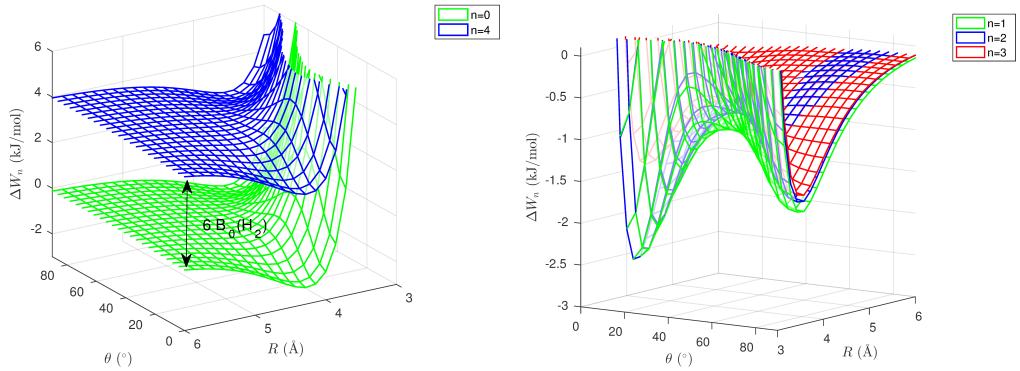
Before we proceed, we briefly discuss why the Born-Oppenheimer approximation is justified for the RA $p\mathrm{H_2}$-$\mathrm{X_2}$ potentials, but beaks down for those of $o\mathrm{H_2}$-$\mathrm{X_2}$. The lowest RA potential for $p\mathrm{H_2}$-$\mathrm{X_2}$ is separated by about $6B_0(\mathrm{H_2}) = 4.1$ kJ/mol from the potential of the next higher $p\mathrm{H_2}(J{=}2)$ states, as indicated exemplary for the $p\mathrm{H_2}$-$\mathrm{Br_2}$ complex in figure 3.15a. Therefore, we expect that the NACT between the bound states of the relevant $p\mathrm{H_2}$-$\mathrm{X_2}$ RA potentials ΔW_0 and the five RA potentials of the $p\mathrm{H_2}(J{=}2)$-$\mathrm{X_2}$ species are small. Thus, we assume that the BOA holds for the $p\mathrm{H_2}(n{=}0)$-$\mathrm{X_2}$ RA potential. The three RA potentials that are asymptotically correlated with the $J{=}1$ ground state of $o\mathrm{H_2}$, *i.e.* ΔW_1, ΔW_2 and ΔW_3, are very close in energy, as is exemplary shown in figure 3.15b for $o\mathrm{H_2}$-$\mathrm{Br_2}$. Hence, we expect that the bound states of these three potentials are highly coupled. This coupling complicates the bound state calculations for $o\mathrm{H_2}$-$\mathrm{X_2}$ significantly. It would be necessary to calculate the NACTs and then solve the coupled equations on the three RA surfaces simultaneously. In this case, it is probably easier to avoid the RA and calculate the lowest bound states of the $o\mathrm{H_2}$-$\mathrm{X_2}$ dimers directly from the 4D potentials. This has been achieved for example for H_2-CO in ref. [151].

The bound states of the $p\mathrm{H_2}$-$\mathrm{X_2}$ van der Waals complexes can be calculated by solving the rovibrational Schrödinger equation with the Hamiltonian,

$$\hat{H}_{\nu=\nu'=0} = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R} + \frac{\hat{l}^2}{2\mu R^2} + \frac{\hat{j}^2_{\mathrm{X_2}}}{2\mu_{\mathrm{X_2}} r^2_{\mathrm{X_2},\nu=0}} \tag{3.21}$$

$$+\frac{\hat{j}^2_{\mathrm{H_2}}}{2\mu_{\mathrm{H_2}} r^2_{\mathrm{H_2},\nu'=0}} + \Delta E(R, \theta_{\mathrm{X_2}}, \theta_{\mathrm{H_2}}, \phi_{\mathrm{H_2}} - \phi_{\mathrm{X_2}}) \tag{3.22}$$

$$+E_{\mathrm{vib},\mathrm{X_2},\nu=0} + E_{\mathrm{vib},\mathrm{H_2},\nu'=0}, \tag{3.23}$$



(a) pH$_2$-Br$_2$: The ground state RA potential ΔW_0 is separated by about $6B_0(\mathrm{H_2}) = 4.1$ kJ/mol from the RA potentials associated with the next higher pH$_2$(J=2) state. Of these only ΔW_4 is depicted schematically by shifting of ΔW_0. NACTs are estimated to be small.

(b) oH$_2$-Br$_2$: The first three RA potentials ΔW_1, ΔW_2 and ΔW_3 are close in energy. NACTs are estimated to be large.

Figure 3.15: RA pair potentials for o/pH$_2$-Br$_2$: estimation of NACTs.

where $\mu = \frac{m_{\mathrm{H_2}} m_{\mathrm{X_2}}}{m_{\mathrm{H_2}} + m_{\mathrm{X_2}}}$ is the reduced mass of the complex, $\mu_{\mathrm{X_2}} = \frac{m_{\mathrm{X}}}{2}$ and $\mu_{\mathrm{H_2}} = \frac{m_{\mathrm{H}}}{2}$ denote the reduced mass of the dihalogen and dihydrogen, $r_{\mathrm{X_2},\nu=0}$ and $r_{\mathrm{H_2},\nu=0}$ are the average bond lengths of the dihalogen and dihydrogen in their respective vibrational ground states with energies $E_{\mathrm{vib,X_2},\nu=0}$ and $E_{\mathrm{vib,H_2},\nu'=0}$, respectively. The first term describes the kinetic energy of the motion along R, the second term describes the rotation of the complex, whereas the third and and the fourth terms describe the rotations of the dihalogen and the dihydrogen molecule. All motions are subject to the interaction potential given by ΔE. The total angular momentum operator is $\hat{J}_{\mathrm{tot}} = \hat{l} + \hat{j}_{\mathrm{X_2}} + \hat{j}_{\mathrm{H_2}}$. After the rotational adiabatization with respect to the fast rotation of dihydrogen, the Hamiltonian becomes dependent on the quantum number n for the rotational state of dihydrogen, where $n = 0$ is asymptotically related to $j_{\mathrm{H_2}} = 0$ (pH$_2$). The rotationally adiabatized Hamiltonian then reads,

$$\hat{H}_{n=0,\nu=\nu'=0} = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R} + \frac{\hat{l}^2}{2\mu R^2} + \frac{\hat{j}^2_{\mathrm{X_2}}}{2\mu_{\mathrm{X_2}} r^2_{\mathrm{X_2},\nu=0}} + W_{0,\nu=0,\nu'=0}(R,\theta_{\mathrm{X_2}}) \tag{3.24}$$

$$+E_{\mathrm{vib,X_2},\nu=0} + E_{\mathrm{vib,H_2},\nu'=0}. \tag{3.25}$$

Since we are interested only in energies relative to the individual dissociation limits, we omit E_{vib} and use $\Delta W_0 = W_0$. The Hamiltonian then only contains terms for the rotation and vibration of the complex as a whole and the rotation of the dihalogen,

$$\hat{H}_{n=0,\nu=0,\nu'=0} = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R} + \frac{\hat{l}^2}{2\mu R^2} + \frac{\hat{j}^2_{\mathrm{X_2}}}{2\mu_{\mathrm{X_2}} r^2_{\mathrm{X_2},\nu=0}} + \Delta W_{0,\nu=\nu'=0}(R,\theta_{\mathrm{X_2}}). \tag{3.26}$$

For the calculation of the bound states, we fix the total angular momentum to $J_{\mathrm{tot}} = 0$. In this case, we can neglect the rotation of the complex and only consider the vibration along R

X	$\mu = \mu_{H_2X_2}$	μ_{X_2}	$r_{X_2,\nu=0}$
^{19}F	3489.5886	17315.9881	2.6777
^{35}Cl	3571.7725	31872.1647	3.8418
^{79}Br	3628.3836	71929.6758	4.3105

Table 3.3: Reduced masses and constant dihalogen bond lengths in a.u. as used in the calculation of the bound state eigenvalue spectra.

and the rotation of the dihalogen simultaneously, further simplifying the Hamiltonian to

$$\hat{H}_{J_{\rm tot}=0,n,\nu=\nu'=0} = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R} + \frac{\hat{j}^2_{X_2}}{2\mu_{X_2} r^2_{X_2,\nu=0}} + \Delta W_{0,\nu=\nu'=0}(R,\theta_{X_2}). \tag{3.27}$$

The respective Schrödinger equation was solved numerically on a two-dimensional discrete variable representation (DVR) grid in R and θ using the Wavepacket code [160, 161]. For the R dependence we used 64 evenly spaced DVR points between 2.5 and 14 Å and plane waves for the basis expansion. The basis functions for the angular degree of freedom are the even and odd spherical harmonics with $m = 0$ up to $l = 64$ evaluated at 65 discrete Gauss-Legendre grid points. Due to time constrains, we did not evaluate the basis expansions individually for even and odd functions, as this feature was not yet implemented in the Wavepacket code. Consequently, some states which are localized in the linear well are two fold degenerate and we could not resolve the respective tunnelling splitting. The reduced masses and the dihalogen equilibrium bond lengths used in the bound state calculations are given in table 3.3.

Modes The shape of the RA potentials with two minima for the linear and T-shaped complex separated by an isomerization barrier restricts the internal motions of the pH$_2$-X$_2$ van der Waals dimers. From the perspective of a fixed X$_2$ the relative motions of the pH$_2$ molecule may be described by stretching in R and two bending modes associated with the linear and T-shape wells, as depicted in figure 3.16. The nodal structure of the bound state wave functions may be translated into respective auxiliary quantum numbers ν_R, ω_L and ω_T, with nodal surfaces perpendicular to the mode vectors. For bound states with energies below the isomerization barrier and those which are clearly localized within one of the two wells, this scheme my apply. For bound states, where L-bending and T-bending modes are simultaneously excited, the assignment in terms of (ν_R, ω_L, ω_T) may be ambiguous and thus difficult.

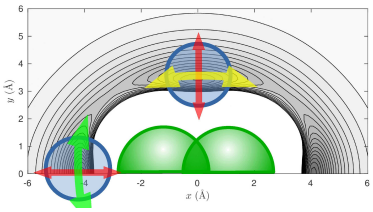


Figure 3.16: Schematic depiction of the L and T-shape forms of the pH$_2$-Cl$_2$ dimer super imposed on the RA-potential in Cartesian coordinates. Arrows denote the R-stretching mode (red), the L-bending mode (green) and the T-bending bode (yellow).

3.4.1 Bound states of the pH$_2$-F$_2$ van der Waals dimer.

The resulting bound states of the $p\mathrm{H}_2(n=0,\nu'=0)$-$\mathrm{F}_2$, $(\nu=0)$ dimer based on the respective RA potential are given in figure 3.17, which shows density plots superimposed on the RA potential. Further, table 5.1 (appendix) gives the bound state energies and the quanta of bending and stretching modes. In order to compare the progression of the bound state energy levels of all three investigated $p\mathrm{H}_2$ dimers on the same scale, we refer to figure 3.20.

The present RA potential gives rise to five bound states, which all lie above the isomerization barrier of -0.41 kJ/mol. The ground state has an energy of -0.209 kJ/mol and is nearly degenerate with the second state at -0.207 kJ/mol. Both states are mainly localized in the linear well, with considerable density tails over the isomerization barrier. The third bound state has an energy of -0.169 kJ/mol and is mainly localized in the T-shape well. Some density of this state is also present in the linear well. The fourth and fifth state are even more delocalized and can be characterised by one and two additional quanta in the T-bending mode. Thus, the $p\mathrm{H}_2$-F_2 dimer is weakly bound and dissociates upon excitation of the stretching mode along R. The relative ZPE contribution to the $p\mathrm{H}_2$-F_2 van der Waals bond amounts to 85%, with respect to the well depth of the 4D MP2/AVQZ PES (T2 minimum: 1.31 kJ/mol).

The RA potential obtained from the 4D MP2/AVQZ PES predicts $p\mathrm{H}_2$-F_2 to be linear. The energetic difference to the T-shaped dimer is just 0.04 kJ/mol, which raises the question whether the prediction of a linear structure is reliable. Concerning the T-shape, MP2/AVQZ underestimates the well depth by 0.13 kJ/mol with respect to CCSD(T)/AVQZ, see table 3.2. Likewise, the depth of the linear well is overestimated by MP2/AVQZ in comparison to CCSD(T)/AVQZ. It seems possible, that the RA potential at the CCSD(T) level may give rise to a T-shaped ground state for the $p\mathrm{H}_2$-F_2 dimer. We can therefore make no reliable prediction with respect to the shape of the $p\mathrm{H}_2$-F_2 dimer, other than that it is highly flexible. Quantum effects due to nuclear motion clearly dominate the structure and the binding energy of the $p\mathrm{H}_2$-F_2 van der Waals dimer.

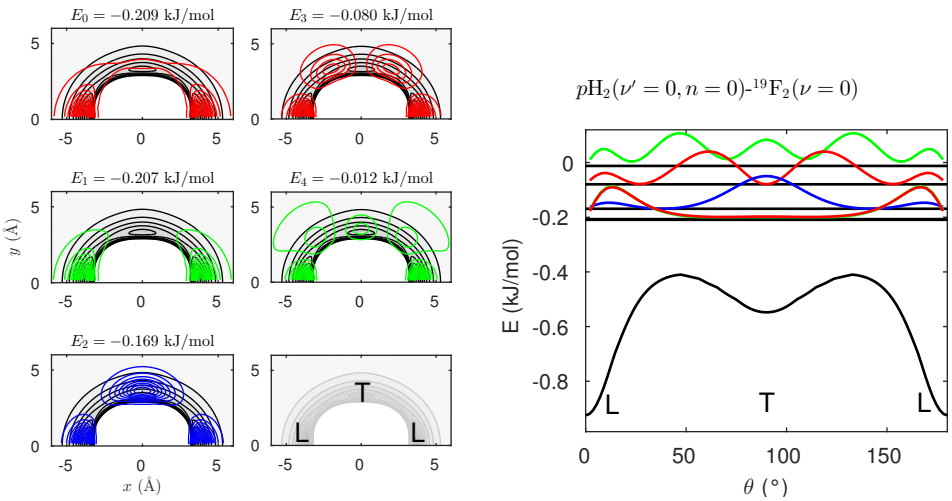


Figure 3.17: Bound states densities for the $p\mathrm{H}_2(n=0,\nu'=0)$-$\mathrm{F}_2$, $(\nu=0)$ with $J_{\mathrm{tot}}=0$ van der Waals dimer, overlaid on the RA potential. The contour plots of the densities show that the ground state is mainly localized in the linear well. However, some density is also present in the T-shaped well.

3.4.2 Bound states of the $p\mathrm{H_2}$-$\mathrm{Cl_2}$ van der Waals dimer.

The RA potential of the $p\mathrm{H_2}(n = 0, \nu' = 0)$-$\mathrm{Cl_2}, (\nu = 0)$ van der Waals dimer with $J_{\mathrm{tot}} = 0$ gives rise to 21 bound states, given in table 5.2 (appendix), of which the lowest two form a nearly degenerate pair. This twofold degenerate ground state has an energy of -0.893 kJ/mol and is localized in the linear well of the potential, see figures 3.18 and 3.20, resulting in a respective ZPE contribution of 63%, with respect to the well depth of the 4D MP2 PES, see table 3.4. The ground state and the first exited state lie below the isomerization barrier of -0.73 kJ/mol, where the latter is localized in the T-shape well and has an energy of -0.798 kJ/mol. The next higher states (3, 4, 7), which are also localized in the T-shape well, show a progression in the T bending quantum number. With up to two quanta in this bending mode, the densities of the respective states, are completely localized in the T-shape well. Beginning with state number 7, with three T-bending quanta, functions localized in the linear well mix in as well. The mixing becomes more pronounced for state 8 and higher, for which no clear localization could be assigned, these states are labelled in table 5.2 (see appendix) with (L,T). State 5 and 6 from a nearly degenerate pair, with one quantum in the linear bending mode, see also figure 3.18. The potential also supports one quantum in the R-stretching mode, giving rise to six states, of which two (13, 17) are localized in the T-shape well, three in the linear well (14, 15, 20) and one delocalized over both wells (19). Besides the singly excited R-stretching mode only one additional bending quantum in either the L or T-bending gives rise to the bound states 17 and 20. Interesting is, that the first bound state excited in the R-stretching mode is localized in the T-shape well, and not the deeper linear well. The bound state structure of the $p\mathrm{H_2}$-$\mathrm{Cl_2}$ dimer shows clearly, that motion along the T-bending mode is less hindered than the motion along the L-bending mode. For example, the energy, that is needed to excite the L-bending mode by one quantum, exceeds the energy of threefold excitation of the T-bending mode, see also figure 3.20.

The energetic difference between the linear ground state and the lowest T-shape bound state is 0.09 kJ/mol. Since the difference between the linear and T-shape well depths at the MP2/AVQZ level is with 0.46 kJ/mol, in good agreement to the CCSD(T)/AVQZ value of 0.30 kJ/mol, see table 3.4, it is likely that $p\mathrm{H_2}$-$\mathrm{Cl_2}$ is indeed linear.

3.4.3 Bound states of the $p\mathrm{H_2}$-$\mathrm{Br_2}$ van der Waals dimer.

The bound state spectrum of the $p\mathrm{H_2}(n = 0, \nu' = 0)$-$\mathrm{Br_2}, (\nu = 0)$, $J_{\mathrm{tot}} = 0$ dimer consists of 43 bound states of which 10 from degenerate or nearly degenerate pairs. The large number of bound states is not surprising, considering that $p\mathrm{H_2}$-$\mathrm{Br_2}$ interacts more strongly than $p\mathrm{H_2}$-$\mathrm{Cl_2}$ and $p\mathrm{H_2}$-$\mathrm{F_2}$ and that the rotational constant of $\mathrm{Br_2}$ is smaller than that of the lighter halogens. The energies and locations and associated quanta in the bending and stretching modes of the bound states are given in table 5.3 (appendix). Density plots of selected bound states are provided in figure 3.19. The ground state is located in the linear well and has an energy of -1.3402 kJ/mol, the respective relative ZPE contribution is thus 57%, with respect to the well depth of the 4D MP2 PES. The first excited bound state has an energy of -0.961 kJ/mol and is localized in the T-shape well, see figure 3.19. In total four energy levels, of which two are twofold degenerate, lie below the isomerization barrier of -0.78 kJ/mol, with up to one quantum in the L and T-bending modes. Beginning with the sixfold excitation of the T-bending, and three fold excitation of the L-bending mode the bound states with $\nu_R = 0$

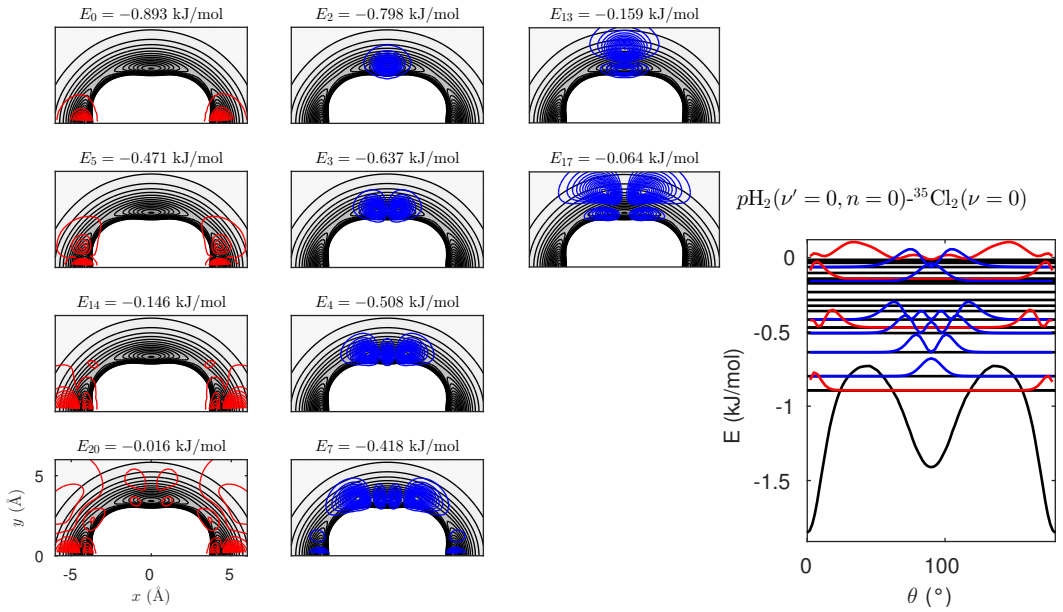


Figure 3.18: Selected bound states densities of the $p\mathrm{H}_2(n = 0, \nu' = 0)$-$\mathrm{Cl}_2, (\nu = 0)$ van der Waals dimer with $J_{\mathrm{tot}} = 0$, overlaid on the RA potential.

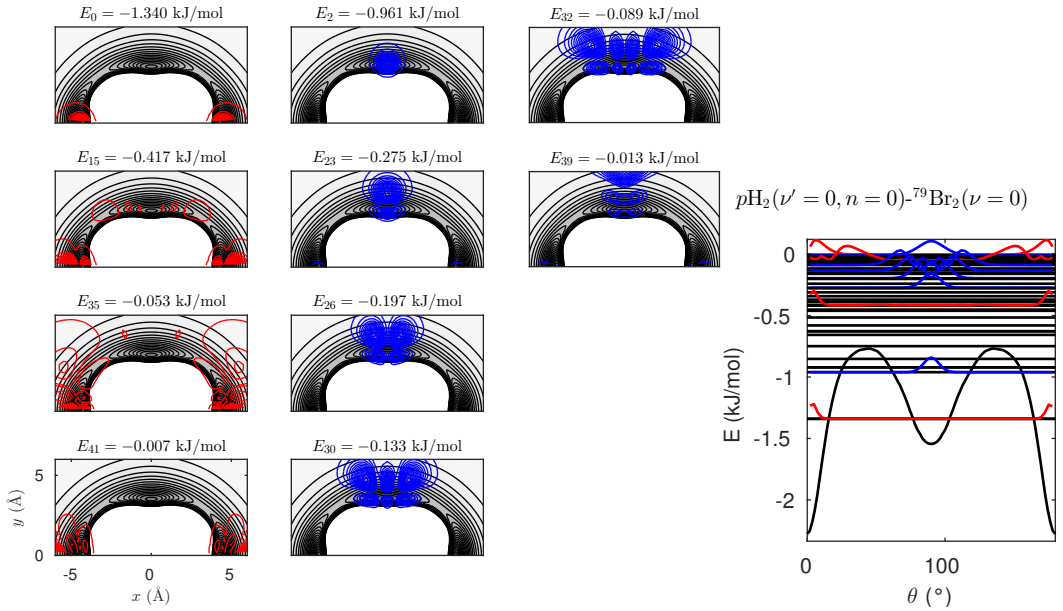


Figure 3.19: Selected bound states densities of the $p\mathrm{H}_2(n = 0, \nu' = 0)$-$\mathrm{Br}_2, (\nu = 0)$ van der Waals dimer with $J_{\mathrm{tot}} = 0$, overlaid on the RA potential.

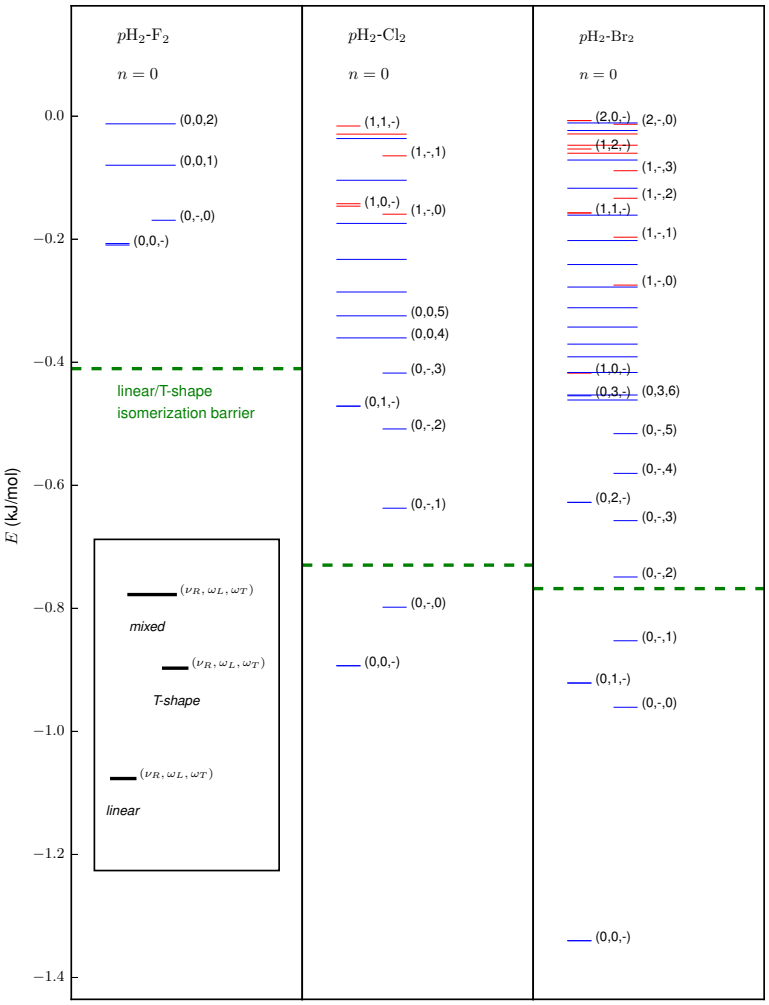


Figure 3.20: Energy level diagram for the bound states for the $p\mathrm{H}_2(n = 0, \nu' = 0)$-$\mathrm{X}_2, (\nu = 0)$ $J_{\mathrm{tot}} = 0$ van der Waals dimers, based on the RA MP2/AVQZ pair potentials. The height of the respective isomerization barrier is given by green dashed lines. The labels ν_R, ω_L and ω_T denote quanta for the R-stretching mode, the L-bending mode and the T-bending mode, respectively. The energies of bound states with $\nu_R > 0$ are denoted by red color.

become highly mixed over both wells, representing almost free rotation of $p\mathrm{H}_2$ from the point of view of a static Br_2. Accordingly, the determination of the exact number of quanta for the bending modes becomes increasingly difficult. The R-stretching mode can be excited by up to two quanta, see also the selected bound state densities in figure 3.19. Due to the deeper well depths in comparison to $p\mathrm{H}_2$-Cl_2, the first bound state with an excited R-stretching mode is localized in the linear well, see also figure 3.20. With respect to $p\mathrm{H}_2$-Cl_2 this situation is reversed in $p\mathrm{H}_2$-Br_2. Note also the increased energetic separation between the first states with $\nu_R = 1$, in comparison to H_2-Cl_2. States with one quantum in the R-stretching mode support also more than 2 and 4 additional quanta in the L and T-bending modes. The states with a twofold excited R-stretching mode lie so close to the dissociation barrier, that excitation of the bending modes would lead to a dissociation of the complex. If these bound states could be found also for RA potentials at the CCSD(T) level of theory is questionable.

The predicted linear structure of $p\mathrm{H}_2$-Br_2 is assumed to be reasonable, based on the energetic difference of 0.38 kJ/mol between the lowest linear and T-shape bound states, see also table 3.4. And because MP2/AVQZ overestimates the difference between the linear and T-shape wells of the RA PES by only 0.2 kJ/mol, with respect to CCSD(T)/AVQZ.

Dimer	Structure	$D_{e,el}$	$D_{e,RA}$	D_0	$1-\frac{D_{e,RA}}{D_{e,el}}$	$1-\frac{D_0}{D_{e,el}}$
pH$_2$-F$_2$	linear	1.31	0.93	0.21	0.30	0.85
	T-shape	0.68	0.55	0.17		
MP2/AVQZ	$\Delta(L-T)$	0.64	0.38	0.04		
CCSD(T)/AVQZ			0.17			
pH$_2$-Cl$_2$	linear	2.43	1.88	0.89	0.23	0.63
	T-shape	1.59	1.42	0.80		
MP2/AVQZ	$\Delta(L-T)$	0.84	0.46	0.09		
CCSD(T)/AVQZ			0.30			
pH$_2$-Br$_2$	linear	3.20	2.33	1.34	0.29	0.57
	T-shape	1.82	1.57	0.96		
MP2/AVQZ	$\Delta(L-T)$	1.39	0.76	0.38		
CCSD(T)/AVQZ			0.57			

Table 3.4: Comparison of well depths (kJ/mol): $D_{e,el}$ (linear = T2, T-shape =T1) 4D MP2/AVQZ PES and $D_{e,RA}$ for 2D RA PES, dissociation energies D_0 (kJ/mol) of the lowest linear and T-shape bound states of the RA potentials, the energy difference $\Delta(L-T)$ between linear and T-shaped dimers and the relative ZPE contribution with respect to $D_{e,el}$, for the rotation of H_2 and the vibration of the van der Waals dimer.

3.5 Summary and conclusions

In most conditions, dihydrogen and dihalogen molecules react violently. However, the reaction is suppressed at ultra low temperatures. Van der Waals interactions then give rise to the formation of dimers H_2-X_2 with interesting bound state spectra, which depend explicitly on the nuclear spin isomer of the dihydrogen molecule present in the dimer. In this chapter, we described the development of rotationally adiabatic pair potentials for the p/oH$_2$-X$_2$ dimers with $X_2 = F_2, Cl_2, Br_2$.

We also investigated the underlying intermolecular interactions. The comparison of the SAPT analysis with the quadrupole-quadrupole interactions shows, that the latter interactions are not the reason for the much stronger binding interaction of the T2 structures. According to the SAPT results, higher-order electrostatic, induction and dispersion interactions are more dominant for the T2 structures, making them the static global minimum structures of the H_2-X_2 dimers. A more sophisticated electrostatic model, for example with distributed multipoles [3], would allow a more realistic description of the electrostatic interactions near the van der Waals minimum of the T2 structure, compared to the quadrupole-quadrupole interaction model. Such a model would then also predict the global minimum for the T2 structure, solely based on electrostatics. The electrostatic preference for the collinear structures is well known in the concept of the σ-hole halogen bond [5]. In this regard, the investigated H_2-X_2 interactions for the T2 structures can be classified as weak, dispersion-dominated σ-hole bond [5].

The rotational adiabatization of the four-dimensional pair potentials, yielded effective two dimensional potential energy surfaces for p/oH$_2$-X$_2$. We restricted our analysis to the pH$_2$-X$_2$ ground state surfaces and the three surfaces (per X$_2$) which are asymptotically correlated with the three fold degenerate rotational ground state of oH$_2$. Two oH$_2$-X$_2$ surfaces show stronger

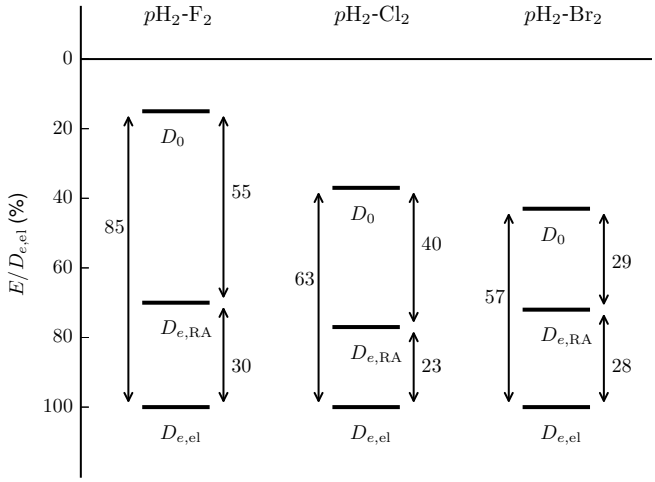


Figure 3.21: RA linear well depth and binding energy relative to the global minimum well depth of the 4D MP2/AVTZ potential.

binding interactions than pH$_2$-X$_2$, while one shows less binding. This can be explained by the alignment of the rotational states to static structures of different energy, and the loss of quadrupole-quadrupole interactions for pH$_2$-X$_2$.

The global minimum of the pH$_2$-X$_2$ and of two lowest oH$_2$-X$_2$ surfaces belong to linear p/oH$_2$-X$_2$ dimers, while the only other local minimum corresponds to a T-shaped dimer. For these surfaces, the global minimum is significantly lower in comparison to the second minimum, hence the lowest bound state for the pH$_2$-X$_2$ dimers in the rotational ground state $J_{\mathrm{tot}} = 0$ is localized in the linear well. However, for pH$_2$-F$_2$ it seems possible, that a T-shape ground state could be favoured at the CCSD(T) level of theory. The prediction of linear structures for pH$_2$-Cl$_2$ and pH$_2$-Br$_2$ on the other hand is more reliable.

Because H$_2$ is very light, a large ZPE gives rise to dissociation energies (D_0) that are considerably lower than the well depths (D_e). In other words, consideration of nuclear quantum effects is absolutely necessary in order to determine binding energies of van der Waals dimers involving H$_2$. For example, the T2 global minimum well depth of the static MP2/AVQZ 4D pair potential of H$_2$-F$_2$ is 1.310 kJ/mol (1.283 kJ/mol) at the MP2/AVQZ (CCSD(T/CBS) level of theory. The rotational adiabatization decreases the MP2 well depth by 30% to 0.931 kJ/mol for the linear pH$_2$-F$_2$ isomer. The dissociation energy D_0 with respect to the vibrational ground state of the pH$_2$-F$_2$ dimer with $J_{\mathrm{tot}} = 0$ is 0.209 kJ/mol. Thus, the vibrational ZPE (after the rotational adiabatization) is 0.722 kJ/mol, or 78% of the well depth of the RA potential or 55% of that of the 4D pair potential. In total, quantum effects dominate the dissociation energy of the pH$_2$-F$_2$ van der Waals dimer, which is 85% lower than the electronic interaction energy at the global minimum. As explained for pH$_2$-F$_2$, the differences between the binding energy and the well depths relative to $D_{e,\mathrm{el}}$ are in %: 85; 30; 55 ($(D_{e,\mathrm{el}} - D_0)$; $(D_{e,\mathrm{el}} - D_{e,\mathrm{RA}})$; $(D_{e,\mathrm{RA}} - D_0)$, see also figure 3.21), while for H$_2$-Cl$_2$ they are 63; 23; 40, and for H$_2$-Br$_2$ we find 57; 28; 29, always with respect to pH$_2(n = 0, \nu' = 0)$-X$_2$, $(\nu = 0)$, $J_{\mathrm{tot}} = 0$ system. A comparison of the RA well depth and binding energy relative to the well depth of the global minimum of the 4D potential is shown in figure 3.21.

Dimer	Structure	$R_{e,RA}$	R_0	$\Delta(R_0 - R_{e,RA})$
pH$_2$-F$_2$	L	3.57	4.11	0.54
	T	3.30	3.92	0.62
pH$_2$-Cl$_2$	L	4.08	4.37	0.30
	T	3.41	3.71	0.30
pH$_2$-Br$_2$	L	4.25	4.48	0.23
	T	3.51	3.78	0.27

Table 3.5: Position of the linear and T-shape minima $R_{e,RA}$ (Å) of the RA potentials and the expectation values R_0 (Å) for the lowest bound state localized in the respective well are given for the pH$_2$-X$_2$ van der Waals dimers. $\Delta(R_0 - R_{e,RA})$ denotes the difference between the expectation value of the van der Waals bond length and the RA value.

Due to the weaker interactions the low-lying bound states of the pH$_2$-F$_2$ complex are less localized in comparison to those of pH$_2$-Cl$_2$ and pH$_2$-Br$_2$. In consequence, the expectation values $\langle R \rangle$ for the lowest state in both wells are shifted with respect to the positions of the potential minima R_e by 0.54 and 0.62 Å in the case of H$_2$-F$_2$, compared to 0.30 and 0.30 Å and 0.23 and 0.27 Å for H$_2$-Cl$_2$ and H$_2$-Br$_2$, see table 3.5. Relative to R_e, the pH$_2$-F$_2$ bond (R_0) is stretched by 15%, in comparison to 5% for pH$_2$-Br$_2$, by the zero point motion.

The bound states calculation of the dimers with pH$_2$ can be performed on a single RA PES, due to the energetic separation between pH$_2$(J=0) and pH$_2$(J=2). The respective bound state energy levels are shown in figure 3.20. A bound states calculation for dimers with oH$_2$ on the three RA PES would demand their non-adiabatic coupling. Therefore, they were not considered here. It is however clear, that the rotational energy spectrum of the oH$_2$-X$_2$ dimers must be much richer in comparison to the ones with pH$_2$, as the three fold degeneracy of oH$_2$ is lifted upon interaction with the dihalogen.

The calculated binidng energies of the pH$_2$-Br$_2$ dimers are similar to the measured values for pH$_2$-I$_2$. For the linear complexes of pH$_2$ the RA MP2/AVQZ values are 1.34 kJ/mol with Br$_2$, compared to the measured value of 1.23 kJ/mol with I$_2$ [124]. As our model is based on MP2 theory and on many adiabatic separations, either measurements of the binding energy of pH$_2$-Br$_2$, or calculations of CCSD(T)/CBS quality for both systems, could truly resolve whether pH$_2$-Br$_2$ is stronger bound than pH$_2$-I$_2$ as indicated by our results. The measured binding energies of the linear (T-shape) He-Br$_2$ complex, 0.203 kJ/mol (0.199 kJ/mol) [162] and He-I$_2$ complex 0.195 kJ/mol (0.199 kJ/mol) [163] also confirm the similarity of the van der Waals interactions of both heavy dihalogens with weakly polarizable species.

Most of the results regarding the development of the RA pair potentials, have been published by us in ref. [139]. All data points of the 4D MP2/AVQZ pair potentials and of the RA potentials, as well as the parameters of the analytic expressions of the RA potentials are available in the supplementary material of ref. [139]. Additionally this thesis also contains results, which were obtained after publication of ref. [139]. These results include the SAPT analysis, details on the obtained fits of the RA PES and the calculation of the bound states of the van der Waals dimers.

3.6 Outlook

The present results might stimulate future research in many directions. Firstly, it remains an open question, whether the bound state energies, and the RA PES calculated with our model, can be experimentally confirmed. For this purpose, experiments similar to those already performed for $p/o\mathrm{H_2}$-$\mathrm{I_2}$ [124, 125] should be extended to the lighter halogens.

On the theory side, the *ab initio* RA potentials for $p/o\mathrm{H_2}$-$\mathrm{I_2}$ need to be calculated. Secondly, the rovibronic spectra for the $p/o\mathrm{H_2}$-$\mathrm{X_2}$ dimers for $J_{\mathrm{tot}} \geq 0$, should be calculated and compared to future experimental results. Thirdly, the presented pair potentials can be used to build model potentials for $\mathrm{F_2}$, $\mathrm{Cl_2}$ and $\mathrm{Br_2}$ surrounded by two or more $p/o\mathrm{H_2}$ molecules. These clusters are highly interesting model systems for multidimensional quantum dynamics studies [164–168]. Fourthly, the librations of dihalogens in solid $p\mathrm{H_2}$ can be studied, see for example [27]. The obtained states give information of the orientation of the dihalogen in the $p\mathrm{H_2}$ matrix and can be used as initial states for time-resolved simulations of photochemical reactions [29, 30, 110]. Starting from the 4D potentials, available in the supplementary materials of ref. [139], the RA can be repeated for $\mathrm{D_2}$, under the approximation $r_{0,\mathrm{D_2}} = r_{0,\mathrm{H_2}}$, see also [27]. Yielding RA potentials for $o/p\mathrm{D_2}$-$\mathrm{X_2}$, for which the present analysis can be repeated and for which the same future research outlook applies.

The present potentials also leave room for future improvements in accuracy. In that regard possible routes are, reduction of the total number of points to necessary regions, exploitation of explicitly correlated MP2/CBS+ΔCCSD(T)/[small basis] schemes in combination with bond functions and core electron correlation. Concerning the intermolecular interactions, one may also calculate the whole 4D SAPT PES and then obtain the RA interaction energy components for $p/o\mathrm{H_2}$-$\mathrm{X_2}$ dimers. Another route for future investigations is the extension of the $\mathrm{H_2}$-$\mathrm{X_2}$ potentials beyond the fixed monomer approximation and the inclusion of excited electronic states. This would enable interesting studies of vibrational pre-dissociation [28] and full *ab initio* simulation of the experimental spectra in ref. [124]. Moreover, highly accurate 6D CCSD(T)/CBS potentials for the $\mathrm{H_2}$-$\mathrm{F_2}$ dimer and respective bound state calculations, including coupling to the rotation of the whole dimer, should give a definitive prediction of the shape of the $p\mathrm{H_2}$-$\mathrm{F_2}$ dimer. Another question for future investigations of the $\mathrm{H_2}$-$\mathrm{F_2}$ van der Waals dimer regards the nuclear spin effect with respect to $p/o\mathrm{F_2}$.

4 Stability of perfluoroalkylalkane hairpin conformers

4.1 Introduction

n-Alkanes (C_nH_{2n+2}) are the most fundamental class of hydrocarbon compounds. Their structural motive, the alkyl chain, occurs in fatty acids which are one building block of lipids. The latter fulfil multiple important biological roles, such as: energy storage, major component of cell membranes and precursors for hormones [169]. Alkanes are the major constituents of natural gas and fossil fuels, which are important chemical energy materials and are further used to produce polymers [170]. Alkanes can also be synthesized from carbonmonoxide and dihydrogen in the presence of a catalyst (Co, Ni) by the Fisher-Tropsch process [171]. Polyethylene is the most common plastic and consist of long interconnected alkyl chains. The physical properties of *n*-alkanes are directly related to their chain length n. At standard ambient temperature (25 °C) and pressure (100 kPa), *n*-alkanes with $n < 5$ are gaseous, liquid for $5 \leq n < 18$ and solid for $n > 18$ [172]. Alkanes are commonly used as non-polar solvents. They can form arbitrary mixtures among themselves but are insoluble in water, *i.e.* they are lipophilic and hydrophobic. In general, alkyl groups are not very reactive, the most common reactions involving alkanes as a reactant are the reaction with O_2 (combustion), radical halogenation and catalitic dehydrogenation [173].

Perfluoro-*n*-alkanes (C_nF_{2n+2}) are the perfluorinated analogues of the *n*-alkanes, *i.e.* all hydrogen atoms are substituted with fluorine atoms. Perfluoroalkanes are obtained by fluorination of alkanes, either by F_2 directly, more common are fluorinating agents (CoF_3) or electrochemical fluorination in HF [12]. In contrast to alkyl chains, perfluoroalkyl chains do not occur naturally and have no biological role, as most fluorine is bound in fluorspar (CaF_2), which has a very low solubility in water [12]. Perfluorinated and partially fluorinated alkanes find use in a diverse range of applications, for example as solvents, polymers (Polytetrafluoroethylene), lubricants, refrigerants, fire extinguishing agents, liquid crystals, contrast agents, inhalation anesthetics, blood substitutes and respiratory fluids [12].

These applications are directly related to their unique properties. The C-F bond has a binding energy of 484 kJ/mol, which is considerably higher than that of the C-H bond (410 kJ/mol), the C-C bond (347 kJ/mol) and other C-halogen bonds (C-Cl: 323 kJ/mol), it is in fact the strongest single bond to carbon [12]. Moreover the van der Waals radius of F is with 147 pm, larger than the one of H (120 pm), which effectively protects the carbon atoms from nucleophilic attacks. Therfore, perfluoroalkanes are highly inert and non toxic, which makes them ideal for applications, where these properties are desired as for example in lubricants and clinical use. [12]

The inertness also causes problems due to the accumulation of fluorocarbons in the biosphere, mainly due to them being green house gases. Because chlorofluorocarbons (CF_3Cl), manly used

as refrigerants, deplete the ozone layer, which is important to filter out harsh UV radiation, hydrofluorocarbons, for example (CF_3CFH_2), are used as substitutes due to their lower ozone depleting potential. They still have a higher global warming potential (GWP), compared to CO_2 but are released in considerably lower quantities into the atmosphere. However, they are replaced also by hydrofluoroolefins, which are in comparison less stable and thus have lower lifetimes in the atmosphere. The GWP is a measure of the heat trapped by a compound in the atmosphere, it is given in relative units to CO_2, which has a GWP of 1. The GWP of CF_4 is 5700, and the atmospheric lifetime is 50 000 years. In comparison CF_3CFH_2 has a GWP of 1600, and an atmospheric lifetime of only 13.6 years. The compound HFO-1234yf (2,3,3,3-tetrafluoro-1-propene) has an atmospheric lifetime of less than a year and a GWP of 4, i.e. just four times higher than that of CO_2. Thus, partially fluorinated compounds, play an important role as environmentally friendly alternatives to traditional chlorofluorocarbon refrigerants, and blowing agents. [12]

Perfluorocarbon liquids, for example perfluoro-*n*-octyl bromide, can dissolve high amounts of O_2 and CO_2, which is attributed to the formation of "cavities" within in the liquid. This property makes them useful as oxygen carrying components in blood substitutes, which can be used during surgeries, and for liquid ventilation/breathing applications. The inert perfluorocarbons seldom interfere with biological mechanisms and leave the body unaltered via the lungs. [12]

Another striking property of perfluoroalkanes is their solvation behaviour with alkanes. Alkanes and perfluoroalkane molecules are bound by similarly weak van der Waals interactions. Evidence for the similar intermolecular interactions of *n*-alkanes and perfluoro-*n*-alkanes is given by the similar experimental molar heats of vaporization for pure liquids of molecules of equal chain length for $n > 1$, see figure 4.1. The additional fact that perfluoroalkanes occupy significantly larger molar volumes, also given in figure 4.1, leads to the observed lower heat of vaporization per volume of perfluoroalkanes, compared to alkanes, see figure 4.1. The only exception from this behaviour is perfluoromethane, due to stronger average binding interactions in comparison to methane [19]. Because of the similar magnitude and nature of the interactions, one would expect that liquid alkanes and perfluoroalkanes should readily mix. However, they actually form separate phases below a certain critical temperature [14, 16]. Alkanes are therefore classified as hydrophobic, lipophilic and fluorophobic and perfluoroalkanes are hydrophobic, lipophobic and fluorophilic. As perfluoroalkanes have a higher density than alkanes and water, the perfluoroalkane phase is located below the water phase, on top of which is the alkane phase, in a triphasic mixture. This effect is exploited in liquid-liquid biphasic synthesis, where for example a fluorophilic catalyst, is initially separated from the non-fluoros organic phase, in which the reactants are solvated. Upon heating, both phases mix and the reaction can proceed. After reaction the mixture is cooled, the products and the catalyst can be recovered from the separated phases [12, 22]. Therefore perfluoroalkane solvents, and the design of fluorophilic catalysts by introduction of perfluoroalkyl groups, play an important role in "Green Chemistry", where reactions and techniques are developed that reduce the consumption of resources, and minimize the use and generation of hazardous substances [15]. Moreover, perfluoroalkyl-perfluoroalkyl interactions in water, are exploited for the design of self assembled supermolecular architectures of perfluoroalkyl-tagged polyglycerol dendrimers, dendrons and perfluoroalkylated linear polyglycerols [23, 24]. The structures of the resulting assemblies

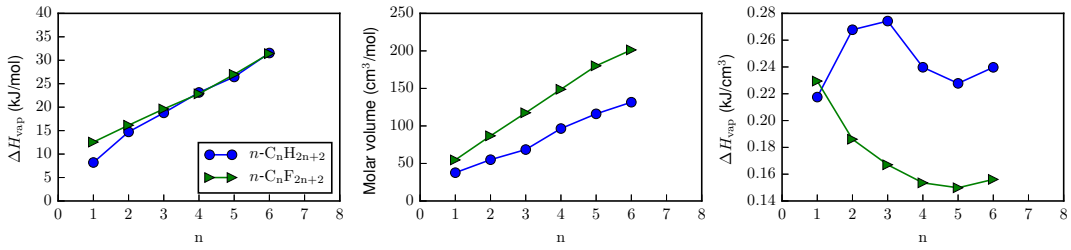


Figure 4.1: From left to right: Experimental heats of vaporization per mol indicate similar average intermolecular interactions for n-alkanes and perfluoro-n-alkanes with $n > 1$. Molar volumes of perfluoroalkanes are considerably larger compared to those of alkanes. The difference in molar volumes and the similar interactions give rise to much smaller heats of vaporization per volume for perfluoroalkanes in comparison to alkanes [19]. Adapted from a table given in ref. [19], see also references in [17] and [174].

		CH_4/CH_4		CH_4/CF_4		CF_4/CF_4	
Method	Ref.	R_e (Å)	D_e (kJ/mol)	R_e	D_e	R_e	D_e
MP2/aug(df)-6-311G*	[19]					4.0	2.89
MP2/aug(df,pd)-6-311G*	[19]	3.8	1.84				
MP2/CBS	[19]					4.0	2.72
CCSD(T)/aug(df)-6-311G*	[19]					4.0	3.26
MP2/CBS (Helgaker)	[175]	3.67	2.05	3.77	2.93	4.01	2.76
CCSD(T)/aug(df,pd)-6-311G(d,p)	[175]		2.13		2.89		3.26
CCSD(T)/aug-cc-pVTZ	[175]				3.10		
CCSD(T)/CBS	[176]	3.63	2.13				
CCSD(T)/CBS est. (D_3, C_3)	[18]		2.1		3.8		3.8
			$(C_2H_6)_2$				$(C_2F_6)_2$
MP2/aug(df)-6-311G*	[19]					4.8	4.27
MP2/aug(df,pg)-6-311G**	[19]	4.0	3.77				
			$(C_3H_8)_2$				$(C_3F_8)_2$
MP2/aug(df,pg)-6-311G**	[20]	3.8	7.74			4.8	6.07

Table 4.1: Intermolecular separation R_e (Å) and well depths D_e (kJ/mol) of the global minima of alkane and perfluoroalkane dimers from the literature. Note that the values of ref. [18] are obtained for dimers with D_3 (C_3) symmetry, whereas the other results are reported for dimers with D_{3d} (C_{3v}) symmetry.

such as microbubbles and microspheres can be tuned and are potential candidates as oxygen and drug carriers [23, 24].

4.1.1 Ab initio studies on small alkane and perfluoroalkane dimers

From a theoretical perspective, the special role of interactions involving perfluoroalkyl groups and alkyl groups in water and other complex media, including the unusual mixing behaviour of alkanes and perfluoroalkanes has not yet been fully understood. One problem is the large complexity of such systems, which can only be handled by a multiscale approach [18]. It is clear, that the first step to an understanding of such interactions requires to study the homo and hetero dimers of perfluoromethane and methane, *i.e.* $(CH_4)_2$, $(CF_4)_2$ and CH_4-CF_4.

Comprehensive MP2 and CCSD(T) studies on the intermolecular interactions of CF_4, C_2F_6 and C_3F_8 dimers were conducted by Tsuzuki *et al.* [19, 20]. Comparing the interaction energies

of eight orientations, they found the CF_4 dimer with face to face orientation and D_{3d} symmetry to be the global minimum structure, which is the same as for the CH_4 dimer. For $(C_2F_6)_2$, seven orientation were considered and the deepest minimum was found for the structure with C_{2h} symmetry and parallel C-C bonds. They conducted an analysis of interaction components using distributed multipoles and concluded that dispersion is the main attractive interaction. Twelve orientations were studied for $(C_3F_8)_2$ at the MP2 level of theory, here the global minimum structure was found to be of C_{2h} symmetry with parallel chains, which is the same as for the C_3H_8 dimer. Tsuzuki *et al.* also compared the interaction energies of the perfluoroalkane and alkane dimers at the respective global minima. While the CF_4 and C_2F_6 dimers interact stronger than the respective alkanes, C_3F_8 dimers interact less in comparison to the C_3H_8 dimer, see also table 4.1. This is explained with the increasing intermolecular separation of the larger perfluoroalkanes compared to the alkanes, and thus smaller dispersion interactions. The orientation dependence of the interaction energy of $(C_3F_8)_2$ is considerably lower than that of $(C_3H_8)_2$, and orientationally averaged interaction energies are very similar [20].

Biller and Mecozzi studied the CH_4-CF_4 dimer and found the largest interaction for a structure with face to face orientation, similar to those of the CF_4 and CH_4 dimers [175]. At the CCSD(T)/aug(df,pd)-6-311G(d,p) level of theory, the global minimum of the interaction energy of the mixed dimer is intermediate between CF_4 and CH_4, see table 4.1.

Mahlanen *et al.* used the basis set developed by Tsuzuki and the MP2 method to compare CX_4 dimers, with $X = H, F, Cl, Br$ [21]. A number of authors, for example Tsusuki *et al.* [177], Takatani *et al.* [178] and Li *et al.* [176], have reported CCSD(T)/CBS interaction energies for the CH_4 dimer. The best estimate for the CCSD(T)/CBS global minimum of the CH_4 dimer is -2.13 kJ/mol at $R_e = 3.64$ Å [176]. Li *et al.* also note, that the CCSD(T)/CBS value can be estimated using a small basis set correction to the MP2/CBS value by

$$E^{CBS}_{\mathrm{CCSD(T)}} \approx E^{\mathrm{CBS}}_{\mathrm{MP2}} + \left(E^{\text{small basis}}_{\mathrm{CCSD(T)}} - E^{\text{small basis}}_{\mathrm{MP2}} \right), \tag{4.1}$$

to a very good accuracy ($\pm$ 0.04 kJ/mol) using the aug-pVDZ basis for the CCSD(T) correction [176].

The most recent study on CH_4/CH_4, CF_4/CH_4 and CF_4/CF_4 was performed by Chattoraj *et al.* [18]. They used the approximation formula stated above to obtain CCSD(T)/CBS results for all three dimers. From this reference data they developed the B3LYP-D3(CF4)/TZVP model with an adjusted D3 correction. The authors performed Monte Carlo (MC) simulations of pure liquid methane, using an interpolated potential based on the DFT-D3 model for the CH_4/CH_4 dimer. They remark that similar MC simulations for CF_4/CH_4 and CF_4/CF_4 are challenging due to the much rougher energy landscape of mixtures involving CF_4, which calls for a significantly higher number of discretization points. They could however obtain temperature dependent effective 1D potentials, and deduce, that the entropy near the minima of CF_4/CF_4 is very small, making CF_4/CF_4 interactions unfavourable at high temperatures. They reason that demixing, and thus increased CF_4/CF_4 interactions, is favoured at low temperatures, because entropy then plays a less important role [18].

In table 4.1 we have collected the well depths and intermolecular separations for alkane and perfluoroalkane dimers from the references discussed above. The interactions of similar alkanes and perfluoroalkanes are of similar magnitude. While the CH_4/CH_4 MP2/CBS interaction energy differs from the CCSD(T)/CBS result by only 0.1 kJ/mol, the interaction energy of

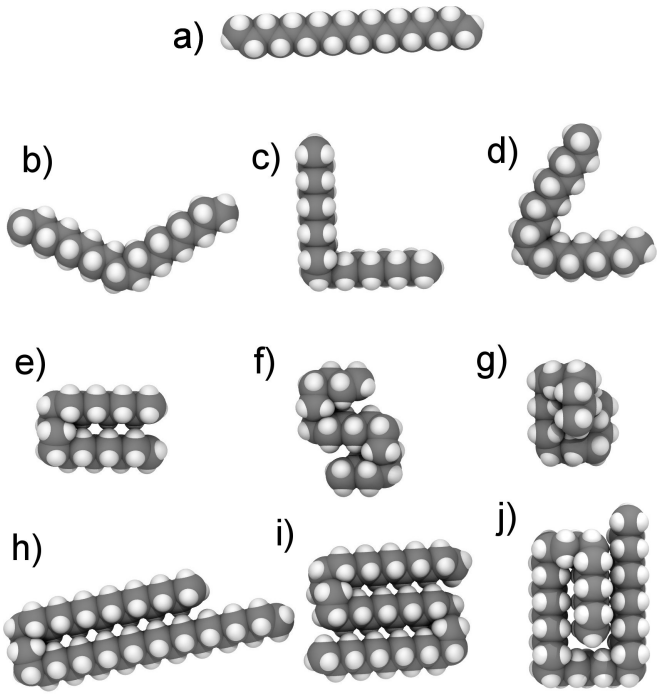


Figure 4.2: Selected conformers of long alkane chain molecules. n-$C_{18}H_{38}$: a) linear: all-trans C-backbone dihedrals, b) V-shape: one gauche dihedral, c) L-shape: two adjacent gauche dihedrals, d) V-shape: two adjacent gauche dihedrals with opposite sign (g_-g_+), e) hairpin: $ggtgg$ sequence in center, f) double hairpin: two $ggtgg$ sequences, g) double hairpin bundle. n-$C_{36}H_{74}$: h) asymmetric hairpin, i) double hairpin, j) broken paperclip: two $ggtgg$ and one $ggttttg_-g_-$ sequence.

dimers involving CF_4 changes by up to 1 kJ/mol when electron correlation beyond MP2/CBS is considered. Thus, inclusion of higher order correlation effects through CCSD(T) theory is a requirement for the accurate description of perfluoroalkyl interactions. Furthermore, perfluoroalkyl interactions are also more demanding in terms of the basis set size in comparison to alkyl interactions. The increased number of valence electrons in perfluoroalkanes in comparison to respective alkanes, *i.e.* 7 per F-atom vs. 1 per H-atom, increases the computational effort additionally. In conclusion, perfluoroalkyl interactions, of even small molecules, pose a substantial challenge for accurate *ab intio* wave function based calculations. Today, the best strategy is to evaluate the MP2/CBS energy and to use small basis corrections to estimate the effect of CCSD(T) electron correlation contributions beyond MP2 [18, 176].

4.1.2 n-Alkane folding

The conformational space of n-alkane molecules increases drastically as a function of the chain length. Preferred C-backbone dihedral angels are the well known trans $t = 180°$, gauche $g_\pm = \pm 60°$ and cross $x_\pm = \pm 95°$ [179]. Beginning with n-butane gauche minima are present and cross minima appear for n-pentane and longer chains. The gauche conformer of n-butane is 2.5 kJ/mol higher in energy than the trans conformer. The trans-gauche barrier is 13.8 kJ/mol and the cis barrier connecting the two gauche minima is 18.7 kJ/mol [179]. Therefore, small n-alkanes ($n = 4 - 8$) do prefer linear all-trans conformations in the gas phase. However, long

n-alkanes must have folded conformers, that are lower in energy than the linear all trans conformer, because they are stabilized by intramolecular van der Waals dispersion interactions [31]. Depending on the temperature, a fraction of molecules will be in the all trans conformation, see figure 4.2 a). Another large fraction will have one gauche angle, see figure 4.2 b), or or more two gauche angles distributed along the C-backbone. Two adjacent gauche angles (gg, g_-g_+) give rise to L and V-shaped conformers, figure 4.2 c) and d), which do not benefit from increased intramolecular interactions. A four twist turn corresponding to a sequence of $ggtgg$ in the central eight carbons gives rise to the hairpin conformer [31], shown in figure 4.2 e). In the hairpin conformer, the two all-trans ends of the n-alkane chain are parallel and close to each other, which maximizes the stabilizing effect of the intermolecular van der Waals interactions [31]. Hairpins do not have to be symmetric, see 4.2 f), however corresponding symmetric hairpins will be lower in energy. With two four twist turns, even more folded conformers are possible, for example a S-shaped double hairpin, figure 4.2 f) and i), or a broken paperclip, figure 4.2 g) and j), these might become increasingly important for very long $n > 30$ n-alkanes. It is clear, that numerous more folded conformers are possible.

The question, from which chain length n onwards hairpins and other folded conformers are energetically more favourable compared to the linear all-trans conformation, has been studied by a number of authors. The more general form of this question is: what is the last globally stable linear n-alkane? In 1997 Goodman [31] was the first to raise this question and to investigate hairpin and other folded conformers of n-alkanes. He found, that the chain length of the shortest n-alkane with a non all-trans minimum conformation depends strongly on the force field or semi-empirical method. For the isolated molecules he found shortest stable hairpins with 18, 25 and 26 carbons for the MM2, MM3 and AMBER force fields and 12 and 60 carbons for the PM3 and AM1 methods, [31]. Goodman also investigated solvent effects using continuum models and found that in water a smaller hairpin was preferred, *i.e.* 22 (AMBER/water) vs. 25. In the less polar solvent chloroform the predicted chain length for the first stable hairpin raised to 142. He noted, that in hexane and other apolar solvents, hairpins are thus possibly highly unfavoured, due to the competing intermolecular interactions [31].

In 2006 Thomas *et al.* reported the first stable hairpin as $C_{22}H_{46}$ with the OPLS-AA force field. They also noted, that gauche energies are overestimated by OPLS-AA and thus estimated the more likely range for the change from all-trans to the hairpin conformer to occur between n = 16-18. They also investigated doubly folded conformers such as the double hairpin bundle and broken paperclip, which are similar to the conformers g) and j) in figure 4.2, and found them to be more stable than the linear conformer for $n = 30$ but less stable than than the corresponding hairpin [180].

It took 16 years until the question posed by Goodman could be answered experimentally. In 2013 Lüttschwager *et al.* reported on the Raman spectroscopy of supersonic jet expansions of linear alkanes with n = 13-20, and on theoretical results based on a combination of local correlation methods LMP2-F12, LCCSD(T) and DFT in the form of B3LYP-D3 [32]. From their experimental and theoretical results follows, that the smallest stable hairpin at (100-150 K) is possibly octadecane ($n = 18$) but definitively nonadecane ($n = 19$) [32]. They also note, that further experiments at even lower conformational temperatures could settle the matter experimentally, since the resolution of their experiments is about 2 kJ/mol [32]. In a following paper, Lüttschwager *et al.* extended their analysis [33].

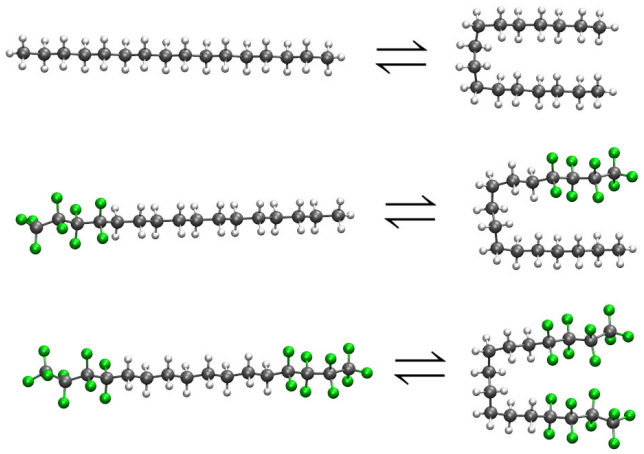


Figure 4.3: Linear (extended) and hairpin (folded) conformers of alkanes, perfluoroalkylalkanes of the type $C_mF_{2m+1} - C_8H_{16}$-C_mH_{2m+1}, and diperfluoroalkyloctanes $C_mF_{2m+1} - C_8H_{16}$-C_mF_{2m+1} chain molecules. Examples given for the total chain length of $n = 16$. Note the flexible central octane unit, which enables the folding. The different hairpins are stabilized by either alkyl-alkyl, perfluoroalkyl-alkyl or perfluoroalkyl-perfluoroalkyl interactions.

The experimental evidence for n-alkane folding, renewed the interest in theoretical investigations on alkane hairpins, with a focus on electron correlation methods. Byrd *et al.* reported CCSD(T)/CBS energies for hairpins up to $n = 14$ [34]. Longer chains, up to $n = 18$, were evaluated using the lower scaling frozen natural orbital (FNO) method. In this method, the number of virtual orbitals necessary for an accurate description of the perturbative triples (T) correlation contribution is greatly reduced, by using natural orbitals constructed from the MP2 density matrix up to a certain occupation threshold. Additionally, ZPE and temperature effects were derived from MP2/cc-pVTZ harmonic vibrational frequencies. From the calculated enthalpy difference between the linear and the hairpin conformer, they predict the hairpin to be more stable for $n \geq 16$, in agreement with the experiment.

Liakos and Neese, reported highly accurate results on all n-alkanes from $n = 6$ up to $n = 19$, using an CCSD(T)/CBS extrapolation sheme based on the domain based pair natural orbital coupled cluster method DLPNO-CCSD(T) [35]. They calculated also the free enthalpy difference and accessed the errors of their method, and give with $n = 18$ at 100 K, the latest prediction for the first stable hairpin, in accordance with the results by Byrd *et al.* and Lüttschwager *et al.*. Pastorczak *et al.* performed intramolecular SAPT calculations on the alkane hairpins [181]. Their results highlight the balance between repulsive electrostatic and exchange interactions and binding London dispersion interactions, the latter of which become increasingly important for the hairpins of longer chains.

4.1.3 Perfluoroalkylalkane hairpins

In this work, we extend the question of the chain length dependence on the stability of linear vs. hairpin conformers of chain molecules to partially fluorinated n-alkanes. Additionally to n-alkanes (C_nH_{2n+2}) we study perfluoroalkylalkane (C_mF_{2m+1}-C_8H_{16}-C_mH_{2m+1}) and diperfluoroalkyloctane (C_mF_{2m+1}-C_8H_{16}-C_mF_{2m+1}) chain molecules. Concerning the window of chain

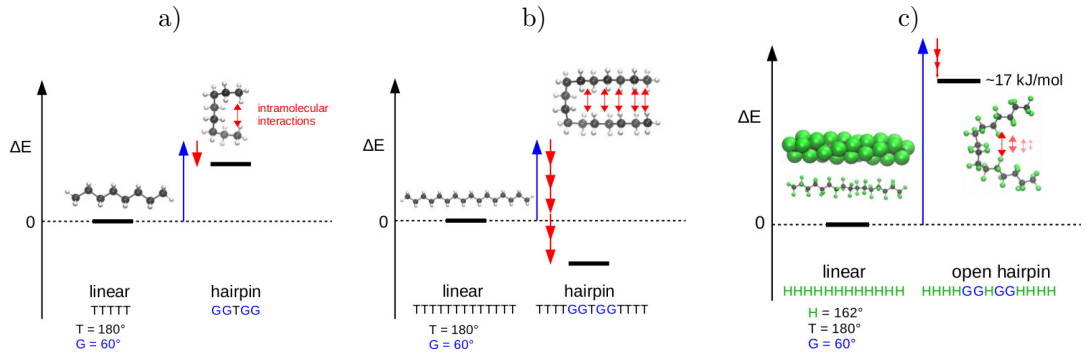


Figure 4.4: (a) and b)) The two twist turn (*ggtgg*) sequence of C-backbone dihedral angles leads to alkane hairpin conformers which are stabilized by intramolecular dispersion interactions between parallel running ends of the chain molecule. Perfluoroalkanes are stiffer than alkanes and do not form similar hairpins. Imposing the two twist turn, results in strong repulsive forces in the fold of the molecule. The relaxed structure is the V-shaped conformer in c). Clearly, the dispersion interactions per added pair of CF_2 groups decrease due to the unfavourable orientation of the chain ends.

lengths, we choose to focus on chains of even lengths from $n = 8$ (n-octane) and $n = 10$ (1-perfluormethylnonane and 1,8-diperfluormethyloctane), up to $n = 22$. As examples, the linear and hairpin structures of the systems are given for a total chain length of 16 carbon atoms in figure 4.3. These foldamers enable the study of self solvation, *i.e.* hairpin formation, driven by either alkyl-alkyl, perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions in a single molecule. The folding of the partially fluorinated chain molecules should in principle be observable by experiments similar to those performed by Lüttschwager *et al.* for the alkanes [32, 33] and by Drawe *et al.* for perfluoroalkanes [182]. The critical chain length for the hairpin formation is very sensitive to the magnitude of the interaction between the adjacent chains, as was shown for alkanes, and should provide experimental evidence on the magnitude of the perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions. Here, we want to predict the critical chain lengths for these systems using dispersion corrected densitiy functional theory and wave function based local correlation methods.

The important structural motive, which enables the hairpin formation of the partially fluorinated n-alkanes, is the central octyl (-C_8H_{16}-) sequence. This part provides the necessary flexibility and comparability between the different hairpins. Perfluorinated n-alkanes are known to be much more rigid than alkanes, and do not fold in the same way as alkanes [182]. For perfluoro-n-butane, the gauche minimum is 6.2 kJ/mol above the twist-anti global minimum, with a barrier of 10.1 kJ/mol, according to MP2/DZ+P cacluations by Dixon [183]. Thus, the introduction of gauche dihedral angles in perfluoroalkanes comes at an twofold increased energy cost in comarison to n-alkanes. To explore the hairpin motive with respect to perfluoro-n-alkanes, we performed DFT calculations using the TPSS functional and the D3(BJ) disperison correction with Becke-Johnson damping and the def2-TZVP basis set with Turbomole. Forcing the hairpin structure on the linear-helical perfluoro-n-alkanes resulted in an unfavourable repulsion between fluorine atoms within the bend. Upon optimization, the structure relaxed to an V-shaped open hairpin, which does not benefit from stabilizing dispersion interactions upon chain elongation. Irrespective of the chain length the open hairpin

was about 17 kJ/mol higher in energy than the linear perfluoroalkane, as shown in figure 4.4 c). Which is in accordance with the experiments of Drawe *et al.* who only observed linear perfluoro-*n*-alkanes in a supersonic jet expansion [182]. Hence, perfluorinated *n*-alkanes do not exhibit the observed folding behaviour as the alkanes. This does not rule out the possibility for folding of extremely long perfluoro-*n*-alkanes.

Perfluoroalkylalkanes are fluorophilic/lipophilic amphiphiles, which exhibit polar character due to the CH_2-CF_2 dipole. Addition to of perfluoroalkylalkanes to perfluoroalkane solvents increases the miscibility with hydrocarbons [16]. Their synthesis is straightforward. For example, the 1,8-diperfluoroalkyloctanes are synthesized by radical addition of perfluoroalkyl iodides to 1,7-octadiene followed by reductive deiodination [184]. Perfluoroalkylalkanes and diperfluoroalkylalkanes are used as additives for paraffines based ski-waxes. They reduce the water adhesion and thereby improve the sliding capabilities [184].

4.2 Methodology

In order to compare the energy of the hairpin to the linear conformer, we calculate the relative hairpin energy,

$$\Delta E(n) = E_{\text{hairpin}}(n) - E_{\text{linear}}(n), \tag{4.2}$$

as a function of the total chain length n. $\Delta E(n)$ is a measure of the electronic hairpin stabilization by dispersion interactions and the electronic part of the enthalpy for hairpin formation. The smallest $n = n_{\text{hp,el}}$ for which $\Delta E(n) < 0$ gives the critical chain length of the last electronically stable linear alkane or perfluoroalkylalkane as $n_{\text{c,el}} = n_{\text{hp,el}} - 1$. Here, we will mostly change the chain length in increments of 2. In this case, the last globally stable linear conformer with an even chain length is $n_{\text{c,el}} = n_{\text{hp,el}} - 2$. In our discussions, we will concentrate on n_{hp}, *i.e.* the chain lenght of the smallest stable hairpin.

To make predictions for the outcome of experiments, it is necessary to also include the effects of ZPE, and finite temperature on the vibrational, rotational and translational degrees of freedom, *i.e.* compute thermodynamic corrections. For each conformer, we compute vibrational frequencies for the normal modes and calculate the enthalpy at finite temperatures T and constant pressure of $P = 100$ kPa,

$$H(T) = E + ZPE + \delta H_{\text{vib}}(T) + \delta H_{\text{rot}}(T) + \delta H_{\text{trans}}(T). \tag{4.3}$$

From this, the enthalpy difference at a finite temperature can be compared,

$$\Delta H(T, n) = H_{\text{hairpin}}(T, n) - H_{\text{linear}}(T, n). \tag{4.4}$$

Or in terms of a correction to $\Delta E(n)$,

$$\Delta H(T, n) = \Delta E(n) + \Delta \delta H(T, n) \tag{4.5}$$

where $\Delta \delta H(T, n)$ holds information about the ZPE and temperature effect on the relative hairpin stability.

Entropic effects can also play a role, and will certainly be very important at elevated temperatures. We estimate the entropy $S(T)$ at a finite temperature from the normal mode frequencies.

Here, we also apply an approximation suggested by Grimme [37] and use the entropy of rigid rotors for vibrations below 100 cm^{-1}. Further details on this approximation are given below. With $S(T)$ we calculate the free energy of each conformer,

$$G(T) = H(T) - TS(T), \tag{4.6}$$

and the relative hairpin free energy by,

$$\Delta G(T, n) = \Delta E(n) + \Delta \delta H(T, n) - T\Delta S(T, n), \tag{4.7}$$

as a function of the chain length and the temperature. Of course, to resolve the whole experiment, one would have to also treat the statistics of individual conformers and the related problem of the conformational entropy, which is beyond the scope of this work. As for $\Delta E(n)$, also the critical chain lengths will be found for $\Delta H(T, n)$ and $\Delta G(T, n)$, which we will denote $n_{\mathrm{hp,H(T)}}$ and $n_{\mathrm{hp,G(T)}}$ for the first stable hairpin and $n_{\mathrm{c,H(T)}}$, $n_{\mathrm{c,G(T)}}$ for the last globally stable linear conformer with even chain length, respectively. $n_{\mathrm{hp,el}}$ is a measure for the performance of electronic structure methods, while $n_{\mathrm{hp,H(T)}}$ and $n_{\mathrm{hp,G(T)}}$ additionally depend on the approximations made for the thermodynamic corrections and can be directly compared to experimental results.

4.2.1 Procedure

Due to the size of the systems, we use a combination of DFT and wave function based local correlation methods to study the hairpin stability. An overview of the overall strategy is provided in figure 4.5.

Initial structure generation: Initial structures are generated by a script that builds Z-matrices from a provided list of carbon back-bone dihedral angles, tetrahedral angles, bond lengths and a string that determines the fluorination pattern. Initial back-bone dihedral angles are trans ($t = 180°$), gauche ($g = 60°$) and helical ($h_{\pm} = \pm 162°$), for either left or right-handed perfluoroalkyl helices. The initial bond lengths are for C-H: $d_{\mathrm{CH}} = 1.09$ Å, C-F: $d_{\mathrm{CF}} = 1.36$ Å and C-C: $d_{\mathrm{C_HC_H}} = 1.53$ Å, $d_{\mathrm{C_FC_F}} = 1.57$ Å, $d_{\mathrm{C_HC_F}} = 1.52$ Å. As an example we give the sequence of carbon back-bone dihedral angles for n-octane, which is 'ttttt' for the linear conformer and 'ggtgg' for the hairpin conformer. For the diperfluoroalkylalkane C_1F_3-C_8H_{16}-C_1F_3 the fluorination pattern is given by the string `'FHHHHHHHHF'`.

Structure optimization: For structure optimization, we employ DFT-D3 in form of the semi-local meta-GGA exchange and correlation functional TPSS [86] and the D3(BJ) Dispersion correction by Grimme [9], with Becke-Johnson damping [36] and the def2-TZVP basis set [185]. We choose to use the TPSS-D3/def2-TZVP level of theory, as it is well tested and provides accurate structures at low computational costs [37]. To speed up the calculations we further employ the multi-pole accelerated resolution of the identity (MARI) approximation for the evaluation of Coulomb integrals. For all MARI-DFT-D3 calculations we use the TURBOMOLE package of programs [186]. Initial structure optimizations were performed with the `M4` integration grid, and the standard SCF energy convergence criterion of 10^{-6} E_H. Then, in all cases, a second structure optimization with a tighter convergence criterion for the SCF energy

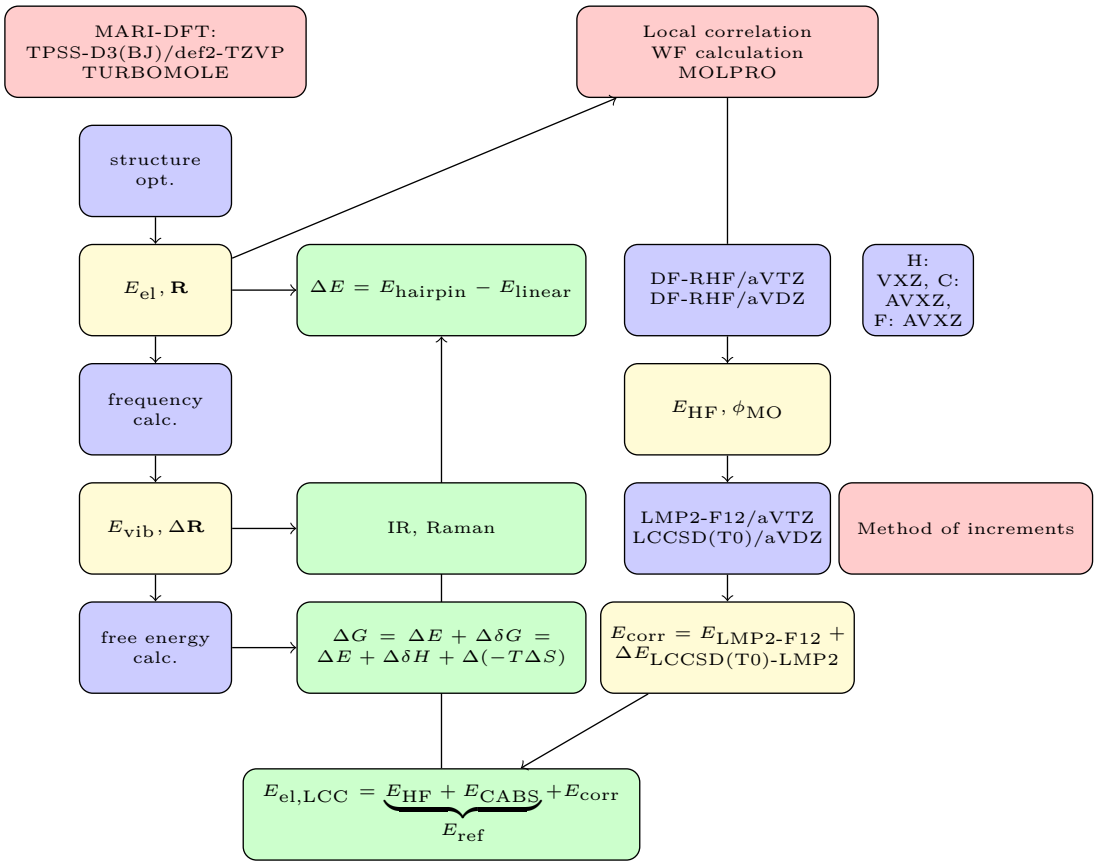


Figure 4.5: Overview of the computational methods applied for the calculation of the hairpin stability of n-alkanes and perfluoroalkylakanes. Left column: DFT-D3 is applied for structure optimization and normal mode analysis, yielding IR/Ramman spectra and thermodynamic corrections. Right column: Single point local correlation methods with the large basis MP2-F12, small basis CCSD(T) correction scheme are applied to calculate wave function based reference results for ΔE. We further use the Method of Increments to approximate the correlation energy.

(10^{-8} $\mathrm{E_H}$) was perfromed using the `M4` grid and the `weight derivatives` option, which ensures correct energy gradients by considering the derivatives of the quadrature weights, see the TURBOMOLE manual [186]. Structure optimizations of linear alkanes were performed under the constraint of C_S point group symmetry. All other structure optimizations were performed in point group C_1.

Vibrational frequencies: For the optimized structures we evaluated the vibrational frequencies and the force constants by numerical differentiation of the gradients using the `NUMFORCE` program from TURBOMOLE. For the `NUMFORCE` calculation the same tight convergence criteria and the option `weight derivatives` as in the second structure optimization step were used.

Free energy: The molecular partition functions, for the evaluation of the enthalpy and the free energy, were calculated from the unscaled normal mode frequencies and the rotational constants of the molecules using the `FREEH` script of TURBOMOLE.

Entropy of low frequency vibrational modes: The harmonic oscillator approximation often results in large amount of noise in the vibrational entropy, due to the poor representation of low frequency modes. A black box method, suggested by Grimme [37] for the determination of accurate and noise free thermodynamic functions for supra-molecular complexes, is to treat low frequency modes as rigid-rotors instead of harmonic oscillators. A switching function is applied to continuously interpolate between both approximations [37].

The entropy of a harmonic oscillator i with frequency ω_i is given by,

$$S_{\mathrm{HO},i} = R\left(\frac{h\omega}{kT\left(e^{\frac{h\omega_i}{kT}}-1\right)} - \ln\left(1-e^{-\frac{h\omega}{kT}}\right)\right), \tag{4.8}$$

where R is the universal gas constant, h is the Planck constant, k is the Bolzmann constant and T is the temperature. For $\omega_i \to 0$ the first term on the right hand side approaches infinity, which results in too large entropy contributions for low frequency modes. Treating these modes within the rigid-rotor instead can circumvent this behaviour. The rotational moment of inertia of a rotor with the same frequency as the oscillator is given by,

$$\mu_i = \frac{h}{8\pi^2\omega_i}. \tag{4.9}$$

For small ω, μ becomes large, thus effective moments of inertia μ' are introduced, which are limited to reasonable values. The effective moments

$$\mu_i' = \frac{\mu_i B_{\mathrm{av}}}{\mu_i + B_{\mathrm{av}}}, \tag{4.10}$$

are limited to the average moment of inertia $B_{\mathrm{av}} = \bar{\mu_i}$, *i.e.* the average over the moments of inertia of all respective normal modes of the molecule. The corresponding rigid-rotor entropy for one mode is then calculated by,

$$S_{\mathrm{RR},i} = R\left(\frac{1}{2} + \ln\left(\sqrt{\frac{8\pi^3 kT}{h^2}\mu_i'}\right)\right). \tag{4.11}$$

For each entropy contribution a switching function $w(\omega_i)$ is used to interpolate smoothly between both models,

$$S_{\mathrm{RRHO},i} = w(\omega_i)S_{\mathrm{HO},i} + [1-w(\omega_i)]\,S_{\mathrm{RR},i} \tag{4.12}$$

$$w(\omega_i) = \frac{1}{1+(\omega_0/\omega)^a}, \tag{4.13}$$

where $\omega_0 = 100\ \mathrm{cm}^{-1}$ denotes the threshold up to which the rigid-rotor model is applied, and $a = 4$ determines the behaviour of the interpolation. Using an individual script, we calculated the entropy contribution of low frequency ($\omega < 100\ \mathrm{cm}^{-1}$) vibrational modes using once the rigid-rotor model and once the harmonic oscillator (HO) model, according to the equations and parameters above, as suggested by Grimme [37]. From this data, we determined the correction to the total entropy from `FREEH`, *i.e.* without the quasi rotational modes, due to this model,

$$\Delta S_{\mathrm{RRHO}} = \sum_i S_{\mathrm{RRHO},i} - \sum_i S_{\mathrm{HO},i}, \tag{4.14}$$

where the sums run over all normal modes of the molecule. This method was also applied to alkane folding by Liakos and Neese [35].

Alternative DFT dispersion corrections: In order to study the performance of DFT dispersion corrections other than D3, we performed single point calculations on the optimized MARI-TPSS-D3(BJ)/def2-TZVP structures using two other approaches. Namely, the non-local correlation functional by Vydrov and Van Voorhis (VV10) [38] and the quantum harmonic oscillator model (QHO) by Silvestrelli [92, 93] in a recent parametrization by Partovi *et al.* [39]. All three dispersion correction methods are described in sub-section 2.2.3.5. For VV10 we performed single point calculations at the MARI-TPSS-VV10/def2-TZVP using TURBOMOLE. QHO results were obtained by Pouya Partovi-Azar at the PBE-QHO plane-wave basis level of theory using the program CP2K/Quickstep [187, 188] and provided to the author. Additionally, PBE-D3 plane wave basis results were obtained by Pouya Partovi-Azar using CP2K and provided to the author. The size of the unit cells in the periodic calculations were carefully evaluated and chosen large enough in order to avoid any interaction between molecular images. Both approaches are fundamentally different from the D3 correction and have, to our knowledge, not been tested for perfluoroalkyl interactions or alkane folding.

Wavefunction-based local correlation methods: An accurate description of ΔE, which to a large extent is governed by dispersion interactions, calls for the application of wavefunction-based electron correlation methods, ideally on the level of the 'gold standard' CCSD(T). The size of the systems only allows the use of low scaling coupled cluster methods. Hence, the relative alkane hairpin energies have been previously calculated by local correlation methods, *i.e.* a combination of PAO based LMP2-F12 and LCCSD(T) by Lüttschwager *et al.* [32] and DLPNO-CCSD(T) by Liakos *et al.* [35], or natural orbital methods with a truncated virtual space, *i.e.* CCSD FNO(T) by Byrd *et al.* [34]. In this work, we adopt the strategy of Lüttschwager *et al.* [32] and perform LMP2-F12 [57] calculations with a medium sized basis set (aVTZ, described below), and calculate a coupled cluster correction using LCCSD(T0) [40, 41, 65, 66] and a small basis set (aVDZ or VDZ). For the aVTZ basis we use Dunnings aug-cc-pVTZ basis functions [143] on C and F and cc-pVTZ basis functions on H centers. Likewise we use the abbreviation aVDZ, *i.e.* aug-cc-pVDZ basis functions on C and F and cc-pVDZ basis functions on H. If we use only the cc-pVXZ basis functions, we use the common abbreviation VDZ or VTZ. The LMP2-F12/aVTZ method should give correlation energies close to the MP2/CBS limit. Combined with the small basis CC correction and the CABS singles correction to HF, the approximation to the CCSD(T)/CBS limit used here is thus given by

$$E^{CBS}_{\mathrm{CCSD(T)}} \approx E_{\mathrm{LCC}} = \underbrace{E^{\mathrm{aVTZ}}_{\mathrm{HF}} + E^{\mathrm{aVTZ}}_{\mathrm{CABS}}}_{\approx E^{CBS}_{\mathrm{HF}}} + \underbrace{E^{\mathrm{aVTZ}}_{\mathrm{corr,\ LMP2\text{-}F12}} + \underbrace{\left(E^{\mathrm{aVDZ}}_{\mathrm{corr,\ LCCSD(T)}} - E^{\mathrm{aVDZ}}_{\mathrm{corr,\ LMP2}}\right)}_{\delta \mathrm{CC}/aVDZ}}_{\approx E^{CBS}_{\mathrm{corr,CCSD(T)}}}. \tag{4.15}$$

Results according to this approximation will be denoted by LCC. We will test this approximation on the interaction energy of the CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers. In all cases, we refer to results obtained within the frozen core approximation. All RHF and electron correlation (LMP2-F12, LCCSD(T0)) calculations have been performed using the MOLPRO package of programs [140]. For all steps, we used the density fitting (DF) approximations, *i.e.*

we performed DF-RHF, DF-LMP2, DF-LCCSD(T0) calculations. In all cases the default program choice for the DF basis sets, *i.e.* same cardinal number, was applied. We assume the errors introduced by the density fitting to be negligible. Local orbital [62] domains were generated using Boughton and Pulay domain completeness parameters [64] of 0.980 for aVDZ and VDZ basis sets and 0.985 for aVTZ basis sets. For the purpose of orbital localization, basis functions with the highest angular momentum were excluded (`PDEL`=1). For LMP2-F12, the standard domains were extended by local orbitals on neighbouring atoms by the connectivity criterion `IEXT`=1, which decreases the domain error. Explicit correlation (F12) is treated by the 3*A(LOC) or 3*A(LOC,FIX) ansatz. In LMP2-F12 all pair domains were treated at the LMP2-F12 level. For LCCSD(T) we choose distance criteria for close pairs according to the distance of the parallel chains in the hairpin structures, the individual ranges will be given in the results section. We have not attempted to correct the relative hairpin energies obtained by LCC for the basis set superposition error, because we assume it to be negligible, due to the local approach in combination with the explicit correlation treatment. To show this, we have also carried out counter-poise corrected LCC calculations for the CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers.

Method of increments: Accurate calculations on partially fluorinated n-alkanes with chain lengths n > 12 still pose a challenge, even for LCC like approaches. One main problem that we encountered, was a limitation in the maximal number of orbital domains. The method of increments (MoI) [44–46, 70–72, 189] allows to break the correlation calculation of a molecule into many smaller individual calculations, and thus allows to treat much larger systems than possible with the local correlation method alone, see also subsection 2.2.2.5. We thus applied the method of increments (MoI) to approximate the correlation part of the relative LMP2-F12 and LCCSD(T0) energies of the hairpins. The analysis of individual increments can provide insight into the spatial distribution of changes in intramolecular correlation energies, which is dominated by dispersion. The local molecular orbitals σ_{CH}, σ_{CC}, σ_{CF} and p_F resulting from Foster-Boys localization [63], have been grouped into bodies. Here, we have chosen bodies as depicted in figure 4.6, namely CH_3, CH_2, H_2CCH_2, CF_3, CF_2 and F_{LP}. The CH_3 body represents 8 valence electrons in the 3 σ_{CH} orbitals of the methyl group and the adjacent σ_{CC} orbital. Similarly, the CH_2 body holds 6 electrons in 2 σ_{CH} methylene orbitals and one σ_{CC} orbital. The same scheme is applied group σ_{CF} and σ_{CC} orbitals in perfluoroalkyl segments to CF_3 and CF_2 bodies. The 3 p_F lone pairs (LP) on each F atom, are grouped to one F_{LP} body representing 6 electrons. The central C_2H_4 body describes in total 10 electrons, which are distributed over the central σ_{CC} orbital and the in the four neighbouring σ_{CH} orbitals. It should be noted, that this scheme only applies to even numbered chain lengths, but can be easily adapted also to the uneven cases, by omission of the C_2H_2 body. For even chain lengths, the total number of one body increments is

$$N_1 = n_C - 1 + n_F,$$

with n_C the number of carbon atoms and n_F the number of fluorine atoms in the system. The number of higher order m body increments, where $m = 2, 3, 4, ...$ is given by the binomial coefficient

$$N_m = \binom{N_1}{m}.$$

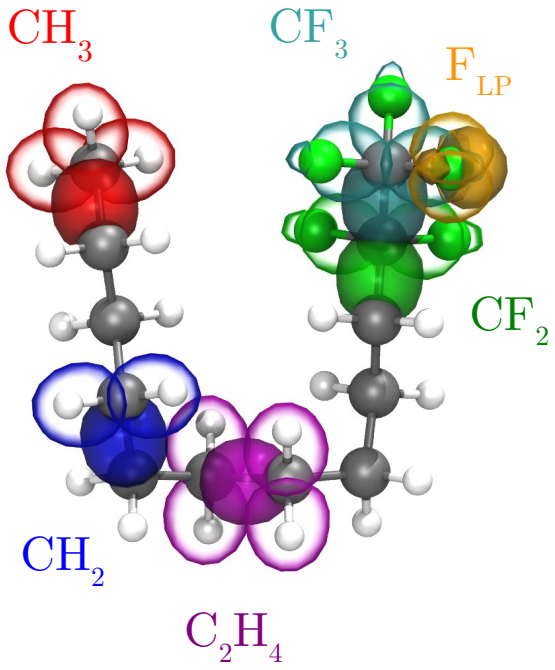


Figure 4.6: Types of one bodies for method of increments calculations, exemplified for the hairpin conformer of perfluoroalkylalkane C_2F_5-C_8H_{16}-C_2H_5. Local orbitals are represented by color coded isosurfaces (± 0.15 a.u.). Groups of local orbitals that form one bodies for the MoI calculations are denoted by color and labelling.

For example for octane, the smallest system considered in this work, we have 7, 21, 35 and 35, one, two, three and four body increments respectively. Considering one of the larger systems (1,8-diperfluorobutyloctane) the number of increments get quite large with 33, 528, 5456 and 40920. Thus, without additional approximations, incremental calculations up to the four-body level are limited to very fast methods and/or smaller systems. We will judge the accuracy of the MoI approximations by comparison to smaller systems, and on that basis omit higher-order terms for larger systems.

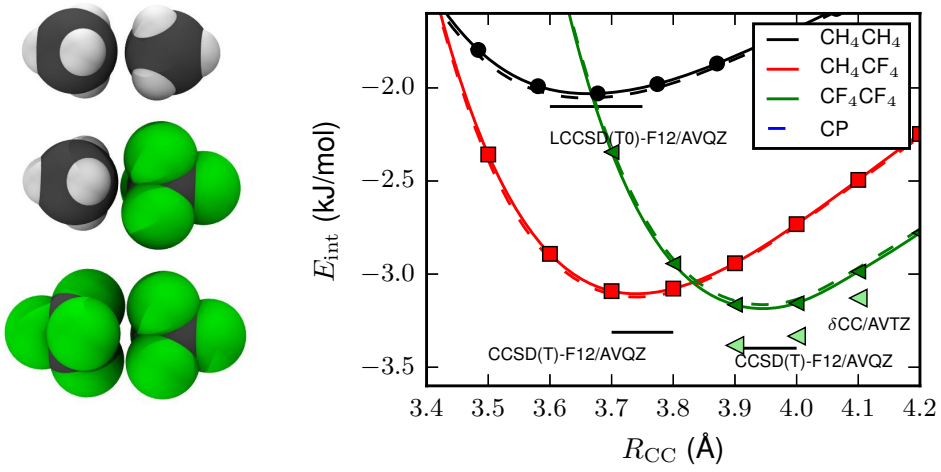
4.3 Results

4.3.1 Interaction energy of CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers.

The CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers are ideal systems to test the performance of the LCC approximation for the interaction energy against higher level of theories. They further enable us to investigate the amount of BSSE present in LCC interaction energies, because the counter poise scheme by Boys and Bernardi [103] can be applied with ease. A study of the dependence of mutual orientations on the interaction energy is beyond the scope of this work. Instead we choose to evaluate only the dimers in the staggered face to face orientation, *i.e.* point group D_{3d} for homo dimers and C_{3v} for CH_4/CF_4 respectively, see also figure 4.7a, which gives the global minimum of the PES according to previous studies [18, 19, 175, 176]. We further apply the fixed monomer approximation with a CH bond-length of $d_{CH} = 1.087$ Å and a CF bond-length of $d_{CF} = 1.321$ Å. Li *et al.* give $d_{CH} = 1.085$ Å from CCSD(T)/AVTZ structure optimization of methane. The CF bond-length was obtained by structure optimization of CF_4 at the CCSD(T)/AVTZ level of theory[1]. Chattoraj *et al.* use similar values of 1.088 Å and 1.326 Å for the CH and CF bond-lengths, respectively [18].

Initially, we calculated the interaction energies at several points close to the van der Waals minima using the LCC methodology. The respective curves are shown in figure 4.7b. For selected points near the minima, we then performed high level, *i.e.* LCCSD(T0)-F12/AVQZ [58] and CCSD(T)-F12/AVQZ [56], counter-poise corrected reference calculations. The results of these calculations are given in table 4.2.

[1] Work of Nils Niggemann during an internship in 2013.



(a) Dimer structures in the staggered face to face orientation.

(b) Interaction energy near the van der Waals minima.

Figure 4.7: Interaction energies (4.7b) near the van der Waals minima, of the CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers depicted in (4.7a). Full curves denote the LCC results, whereas counter poise corrected LCC results are denoted by dashed lines. Bars denote results from reference counter poise corrected single point calculations annotated by the respective levels of theory. Using AVTZ for the δCC correction (light green triangles) closes the gap to the reference, as exemplified for CF_4/CF_4.

High level reference calculations We first discuss the reference results. For the interaction energy of the methane dimer, the LCCSD(T0)-F12/AVQZ method, which gives −2.10 kJ/mol at $R_{CC} = 3.70$ Å, is in excellent agreement with the CCSD(T)/CBS limit determined by Li *et al.* [176], the respective error is −0.03 kJ/mol (1.5%). The canonical equivalent CCSD(T)-F12/AVQZ gives the same value of −2.18 kJ/mol at $R_{CC} = 3.70$ Å and $R_{CC} = 3.60$ Å, and is thus also in excellent agreement (+0.05 kJ/mol) with the CCSD(T)/CBS limit. Note, that we compare the energies for slightly different intermolecular distances, as the value of Li *et al.* refers to $R_{CC} = 3.63$ Å. We expect the resulting error in our comparison to be small, since the interaction energies obtained with LCC, see next paragraph and figure 4.7b, between 3.60 Å and 3.80 Å agree within 0.1 kJ/mol.

Our best estimate for the interaction energy of the CH_4/CF_4 dimer is −3.31 kJ/mol at the CCSD(T)-F12/AVQZ level of theory. The estimated CCSD(T)/CBS value −3.8 kJ/mol in ref. [18] was obtained for a slightly different orientation of the monomers, resulting in a dimer with D_3 symmetry. For the CF_4/CF_4 dimer CCSD(T)-F12/AVQZ gives an interaction energy of −3.40 Å. An indication, that the reference results for the CH_4/CF_4 and CF_4/CF_4 dimers are close to the CBS limit is given by the similarity to the results obtained with the AVTZ basis, which differ by just 0.05 kJ/mol and 0.02 kJ/mol, respectively. Again, the estimated CCSD(T)/CBS value in ref. [18] seems to be too low. We suspect, these differences to be due to the omission of the counter-poise correction in ref. [18]. Biller *et al.* reported large BSSE present for the dimers with CF_4 [175]. Our results for CF_4/CF_4 show, that even for the CCSD(T)-F12/AVQZ method a sizeable BSSE of 7% (15% AVTZ), with respect to the uncorrected energy, is present.

LCC interaction energy The LCC approximation gives results that are in good agreement with the high level references, see table 4.2. For the CH_4/CH_4 dimer LCC gives −2.03 kJ/mol at $R = 3.68$ Å. Thus, LCC underestimates the binding energy given by Li *et al.* by only 0.1 kJ/mol (5%). For CH_4/CF_4 the LCC interaction energy is −3.09 kJ/mol at $R = 3.70$ Å, hence strength of the Van der Waals bond is underestimated by 0.22 kJ/mol (7%) in comparison to CCSD(T)-F12/AVQZ. LCC gives a similar error (0.23 kJ/mol (7%)) for the CF_4/CF_4 interaction energy, which is −3.16 kJ/mol at $R = 3.90$ Å. These errors apply for the A*(LOC) F12 ansatz. If the fixed amplitude ansatz A*(LOC,FIX) is used for F12, the relative errors increase to 8%, 9% and 9% for CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 respectively. For the CH_4/CH_4 dimer, the coupled cluster correction δCC/aVDZ has almost no contribution. However, for CH_4/CF_4 and CF_4/CF_4 this correction becomes increasingly important, as noted also in ref. [18]. The small aVDZ basis set used in the correction δCC is the major source of error in the LCC approximation. Using δCC/aVTZ instead of δCC/aVDZ for CF_4/CF_4 results in an LCC(δCC/aVTZ) interaction energy of −3.40 kJ/mol, in perfect agreement with the CCSD(T)-F12/AVQZ reference. The influence of the domain extension beyond bounded neighbours (`IEXT=1`) is small, as shown by LCCSD(T0)-F12/AVTZ calculations with `IEXT=2` for CF_4/CF_4. We conclude, that the LCC approximation indeed gives interaction energies close to the CCSD(T)/CBS limit for CH_4/CH_4 with δCC/aVDZ, and for CF_4/CF_4, with δCC/aVTZ. The agreement of LCC for CH_4/CF_4 and CF_4/CF_4 with δCC/aVDZ to the reference values is still good, and most probably in the region of a CCSD(T)/AVTZ treatment.

CH_4/CH_4	R = 3.68 Å		
		E_{int} (kJ/mol)	Error (kJ/mol)
	DF-HF/aVTZ	1.70	−3.83
	DF-LMP2/aVTZ	−1.77	−0.36
	DF-LMP2-F12/aVTZ		
	A*(LOC)	−2.06	−0.07
	A*(LOC,FIX)	−1.99	−0.14
	δCC/aVDZ	0.03	
	LCC	−2.03	−0.10
	R = 3.70 Å		
	LCCD(T0)-F12/AVQZ (CP)	−2.10	−0.03
	CCSD(T)-F12/AVQZ (CP)	−2.18	0.05
	R = 3.60 Å		
	CCSD(T)-F12/AVQZ (CP)	−2.18	0.05
	R = 3.63 Å		
	CCSD(T)/CBS (CP)[176]	−2.13	0.00

CH_4/CF_4	R =3.70 Å		
		E_{int} (kJ/mol)	Error (kJ/mol)
	DF-HF/aVTZ	2.09	−5.41
	DF-LMP2/aVTZ	−2.92	−0.39
	DF-LMP2-F12/aVTZ		
	A*(LOC)	−2.98	−0.34
	A*(LOC,FIX)	−2.90	−0.42
	δCC/aVDZ	−0.12	
	LCC	−3.09	−0.22
	DF-LCCSD(T0)-F12/aVTZ (CP)	−3.19	−0.12
	CCSD(T)-F12/AVTZ (CP)	−3.36	0.05
	DF-LCCSD(T0)-F12/aVQZ (CP)	−3.19	−0.12
	CCSD(T)-F12/AVQZ (CP)	−3.31	0.00
	R = 3.77 Å		
	CCSD(T)/AVTZ (CP) [175]	−3.10	
	CCSD(T)/CBS est. (D_3) [18]	−3.80	

CF_4/CF_4	R = 3.90 Å		
		E_{int} (kJ/mol)	Error (kJ/mol)
	DF-HF/aVTZ	2.65	−6.04
	DF-LMP2/aVTZ	−3.41	0.02
	DF-LMP2-F12/aVTZ		
	A*(LOC)	−2.88	−0.52
	A*(LOC,FIX)	−2.81	−0.59
	δCC/aVDZ	−0.29	
	LCC	−3.16	−0.23
	LCC (δCC/AVTZ)	−3.40	0.00
	LCCSD(T0)-F12/AVTZ (CP)		
	`iext=1`	−3.38	−0.02
	`iext=2`	−3.41	0.01
	CCSD(T)-F12/AVTZ+BSSE	−4.02	0.62
	CCSD(T)-F12/AVQZ+BSSE	−3.65	0.25
	CCSD(T)-F12/AVTZ (CP)	−3.42	0.02
	CCSD(T)-F12/AVQZ (CP)	−3.40	0.00
	R = 4.0 Å		
	CCSD(T)/aug(df,pd)-6-311G(d,p) (CP) [175]	−3.26	
	CCSD(T)/CBS est. (D_3) [18]	−3.80	

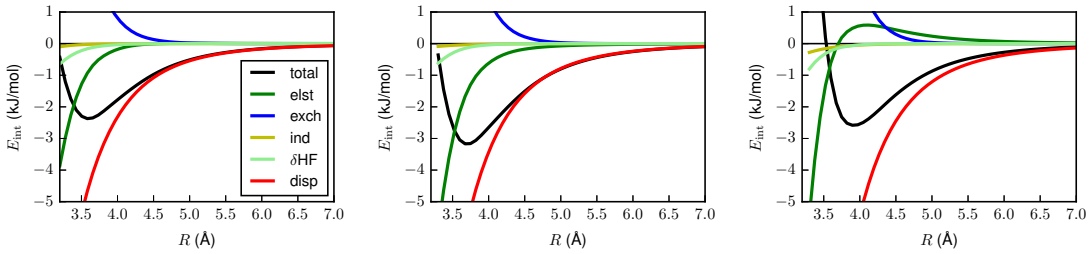
Table 4.2: Interaction energies for the methane and perfluoromethane dimers near the global minima. Energies subject to the counter-poise correction by Boys and Bernardi are marked by (CP). The local correlation plus explicit correlation (F12) methods are effectively BSSE free, for this see table 4.3.

		CH_4/CH_4 $R = 3.68$ Å			CH_4/CF_4 $R = 3.70$ Å			CF_4/CF_4 $R = 3.90$ Å		
E_{int} (kJ/mol)			CP	BSSE		CP	BSSE		CP	BSSE
E_{ref}	DF-HF/aVTZ	1.70	1.75	−0.05	2.09	2.51	−0.42	2.65	3.63	−0.98
	CABS	0.03	0.00	0.03	0.32	−0.01	0.33	0.85	0.03	0.82
E_{corr}	DF-LMP2/aVTZ	−3.48	−3.48	0.01	−5.01	−4.91	−0.10	−6.06	−5.84	−0.22
	F12 A*(LOC)	−0.31	−0.31	0.00	−0.38	−0.59	0.21	−0.31	−0.79	0.48
	DF-CCSD/aVDZ	0.32	0.29	0.03	0.37	0.39	−0.02	0.34	0.46	−0.12
	T0/aVDZ	−0.29	−0.30	0.01	−0.49	−0.50	0.01	−0.62	−0.63	0.00
	δCC	0.03	−0.01	0.04	−0.12	−0.11	−0.01	−0.29	−0.17	−0.12
E_{tot}	DF-LMP2/aVTZ	−1.77	−1.73	−0.04	−2.92	−2.40	−0.52	−3.41	−2.21	−1.20
	DF-LMP2-F12/aVTZ	−2.06	−2.04	−0.02	−2.98	−3.00	0.03	−2.88	−2.97	0.10
	LCC	−2.03	−2.05	0.02	−3.09	−3.11	0.02	−3.16	−3.14	−0.02

Table 4.3: Basis set superposition errors (BSSE) for the components of the LCC interaction energy for the methane and perfluoromethane dimers near the global minima. Relative BSSEs of the uncorrected LCC energies are below 1%, hence LCC is effectively BSSE free.

LCC and BSSE Local correlation methods in conjunction with explicit correlation are known to be effectively BSSE free or to have very small BSSEs, in comparison to canonical methods [105]; see also the sections 2.3.2, 2.2.2.3 and 2.2.2.4. To test this for the three dimers, we calculated the counter poise corrected LCC curves, shown by dashed lines in figure 4.7b. The counter poise corrected and uncorrected LCC interaction energy curves of all three dimers are virtually identical, due to relative BSSEs below 1%. Table 4.3 gives the BSSEs for individual components of the composite LCC energy. For CH_4/CH_4 all components show only small BSSEs. However, for CH_4/CF_4 and CF_4/CF_4 larger BSSEs are present in the components. Sizeable BSSEs arise for the HF energy, which are removed to a large extend by the BSSE of the CABS correction. The LMP2/aVTZ correlation energy contribution shows a sizeable BSSE of 3% for CF_4/CF_4. The explicit correlation energy components show the largest relative BSSEs, which partly cancels out with the BSSE of the LMP2/aVTZ correlation energy. Likewise, BSSEs of alternate signs for δCC and LMP2-F12 cancel each other out, which keeps the total BSSE small. Due to this cancellation of errors, the BSSE of LCC is lower than that of LMP2-F12, for the difficult CF_4/CF_4 system. The CP corrected DF-LMP2-F12/aVTZ interaction energies agree with the extrapolated MP2/CBS results of Biller *et al.* [175], see also table 4.1, for CH_4/CH_4, CH_4/CF_4, while a small deviation of 0.2 kJ/mol is present for CF_4/CF_4. The BSSE for the DF-LMP2-F12/aVTZ interaction energy of CF_4/CF_4 is positive (0.1 kJ/mol), due to the large positive BSSE contribution of the F12 energy, which effectively brings the uncorrected interaction energy closer to the MP2/CBS limit.

SAPT analysis To show the importance of individual interaction components, we performed HF-SAPT/AVTZ calculations for the three dimers using MOLPRO [140]. The resulting curves for the R dependence of the total interaction energy and its electrostatic, exchange, induction and dispersion components are given in figure 4.8. The interpolated values at the minimum of the curves are given in table 4.4. In terms of the well depths of CH_4/CH_4 and CH_4/CF_4, HF-SAPT/AVTZ gives results close to the ones obtained by CCSD(T)-F12/AVQZ theory. For CF_4/CF_4 this is different, as HF-SAPT/AVTZ underestimates the binding energy by about 0.8 kJ/mol. Hence, SAPT predicts the CH_4/CF_4 interaction to be stronger than the CF_4/CF_4



(a) CH_4/CH_4 **(b)** CH_4/CF_4 **(c)** CF_4/CF_4

Figure 4.8: HF-SAPT/AVTZ components of the interaction energy of the (a) CH_4/CH_4, (b) CH_4/CF_4 and (c) CF_4/CF_4 dimers depicted in (4.7a).

	CH_4/CH_4	CH_4/CF_4	CF_4/CF_4
R_e (Å)	3.60	3.70	3.90
E1pol (kJ/mol)	−0.86	−1.53	0.49
E1exch	3.31	4.14	3.26
E2ind+E2ind-exch	−0.02	−0.04	−0.06
E2disp+E2disp-exch	−4.52	−5.70	−6.45
δ HF	−0.15	−0.13	−0.05
HF-SAPT/AVTZ	−2.37	−3.18	−2.58

Table 4.4: HF-SAPT/AVTZ components and interaction energy of CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers.

interaction. The most notable differences between the interactions of the latter dimers, is the electrostatic interaction which is binding for CH_4/CF_4 but anti-binding for CF_4/CF_4.

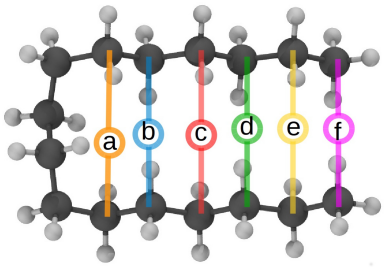
The relative magnitudes of the SAPT interaction components show the pattern expected for apolar molecules. In terms of magnitudes at the minima, the dispersion interaction is the most important, followed by the exchange interaction, electrostatic interactions are small and induction interactions are absent. The δHF component plays no important role either. Taking into account only the dispersion interactions at the minima, we see that they increase by about 1 kJ/mol from CH_4/CH_4 over CH_4/CF_4 to CF_4/CF_4. The actual reason for the discrepancy of the CCSD(T) and HF-SAPT results for CF_4/CF_4 are most probably correlation contributions to the electrostatic interaction energy and dispersion contributions of higher order, which are not covered by HF-SAPT. In the regions with $R > 4.5$ Å (CF_4/CF_4: $R > 5$ Å) the exchange and electrostatic interactions cease and dispersion dominates the interaction energy. Since similar large C-C distances are present between parallel chains of the hairpin conformers, we can assume that dispersion is the main diving force for hairpin stabilization.

Conclusion The LCC approximation, which can be also applied to larger systems, gives accurate and effectively BSSE-free interaction energies for CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers near their respective global minima, underestimating the van der Waals interacion energy by less than 10%. The ordering of the magnitude of interactions, $CH_4/CH_4 < CH_4/CF_4 < CF_4/CF_4$, is also reproduced. However, the small difference between CH_4/CF_4 and CF_4/CF_4 is slightly underestimated. Although there is no trivial way to transfer the errors for the interaction energy of the three dimers to the hairpin systems, due to the obvious structural differences, we are confident that the LCC approximation gives reasonable results within about 1 kJ/mol of the actual relative CCSD(T)/CBS hairpin energies.

4.3.2 n-Alkane folding

In this section, we present the results on the relative hairpin energy of the n-alkanes. We focus on hairpins with even chain lengths n from $n = 8$ to $n = 22$. The section is organized in the following way. First, we discuss the TPSS-D3/def2-TZVP hairpin structures and compare them to those obtained by others using different levels of theory. Second, we present the relative hairpin energy ΔE obtained with the LCC approximation and discuss the effect of the individual components within LCC. Third, we compare the LCC results to wave-function-based reference results obtained by others. Fourth, we compare DFT results for ΔE using the D3, VV10 and QHO dispersion corrections, to the LCC results. Fifth, we add thermodynamic corrections obtained at the TPSS-D3/def2-TZVP level of theory to the LCC results and predict the first stable hairpin at experimental conditions ($T = 100$ K). Last, we asses the convergence of the method of increments in the framework of the LCC for ΔE.

n-Alkane hairpin structures The n-alkane hairpin structures with even n up to $n = 18$ obtained at the TPSS-D3/def2-TZVP level of theory are shown in figure 4.9. The structures of short hairpins (n = 8-12) are more open compared to those with $n \geq 14$. Additionally, the 'arms' of the hairpins become gradually more oriented towards each other upon chain elongation. The distance between facing carbon atoms on both 'arms' alternates between long (≈ 4.5 Å) and short (≈ 4.1 Å), due to the 'zig-zag' structure of the alkyl chains. This is exemplified for the hairpin conformer of $C_{16}H_{34}$ in table 4.5. As a measure of the mean distance between the parallel alkyl chains we introduce the mean distance between facing carbon atoms on the parallel chains. For $C_{16}H_{34}$ for example, this distance is calculated by adding up the values a to f, given in table 4.5, and dividing by six. The mean C-C distances calculated in this manner are given figure 4.10a. This measure is in accordance with the visual inspection of the individual structures in figure 4.9. The $C_{10}H_{22}$ hairpin is the most open structure, with a mean C-C distance of 5.2 Å. The mean C-C distance between the 'arms' decreases significantly for $n = 12$, and are almost constant for hairpins with $n = 14$ to $n = 22$ with ≈ 4.3 Å.



Ref.	Method	a (Å)	b (Å)	c (Å)	d (Å)	e (Å)	f (Å)
this work	TPSS-D3/def2-TZVP	4.64	4.09	4.54	4.13	4.34	4.11
Lüttschwager *et al.* [32]	LMP2/VTZ	4.67	4.09	4.62	4.13	4.41	4.06
Byrd *et al.* [34]	MP2/VTZ	4.55	3.90	4.34	3.85	4.09	3.83
Liakos *et al.* [35]	B2PLYP-D3/def2-pVQZ	4.59	3.98	4.42	3.94	4.16	3.88

Table 4.5: C-C distances between carbons on parallel chains of the C16 alkane hairpin, as indicated by the above figure.

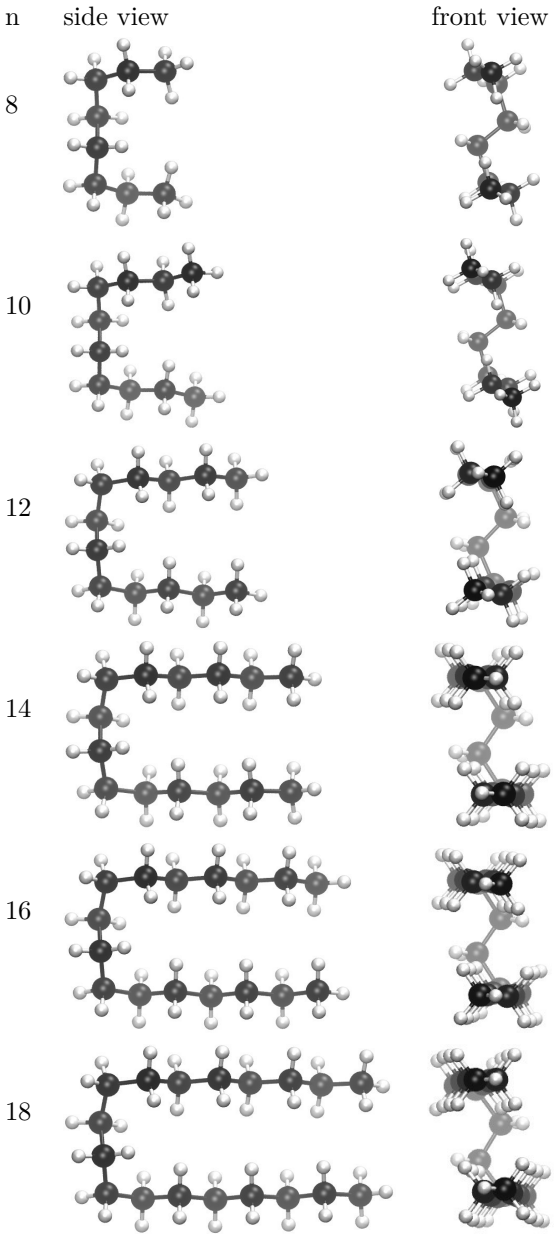
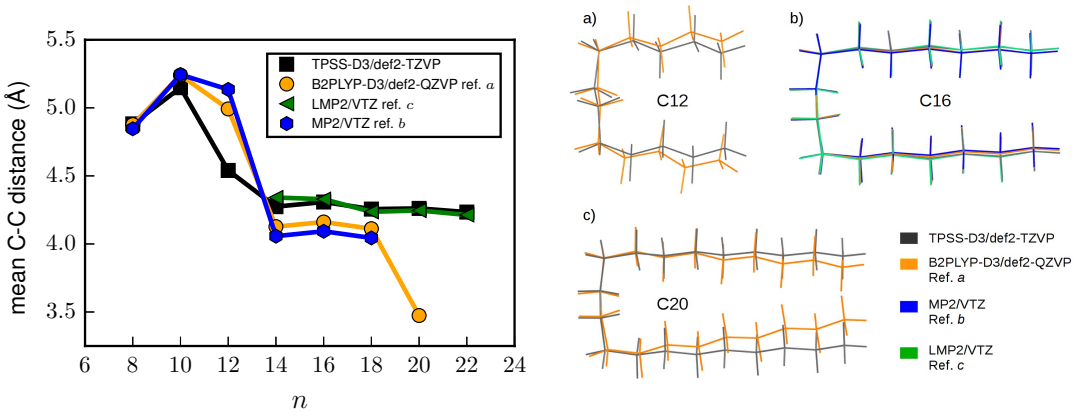


Figure 4.9: Alkane hairpin TPSS-D3/def2-TZVP structures for even n from $n = 8$ to $n = 18$.

In previous works other levels of theory have been applied for the structure optimization of the n-alkane hairpins. Before we compare the relative hairpin energies from previous works with respect to those obtained for the TPSS-D3/def2-TZVP structures, we have to first compare the structures. For this purpose, we have calculated the mean C-C distances between the arms of hairpins with even n for the LMP2/VTZ structures of Lüttschwager *et al.* [32], the MP2/VTZ structures of Byrd *et al.* [34] and the DFT double hybrid B2PLYP-D3/def2-QVZP structures of Liakos *et al.* [35]. The n dependence of the mean C-C distances are given in figure 4.10a. For hairpins with $n = 8$ and $n = 10$ the mean C-C distances are very similar, indicating a good match between the TPSS-D3, MP2 and B2PLYP-D3, structures. The TPSS-D3 hairpin

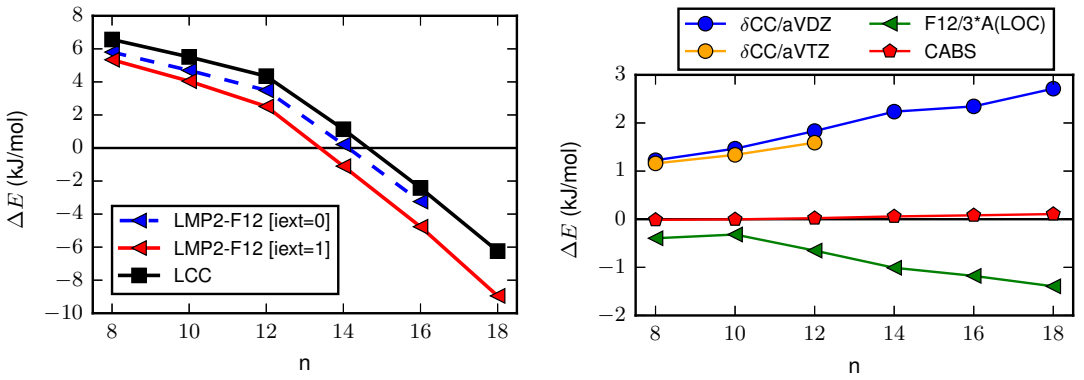


(a) Mean C-C distances of facing carbons on adjacent chains in the hairpin conformers. For an example with individual C-C distances see table 4.5.

(b) Selected aligned hairpin structures from this work and from others.

Figure 4.10: Comparison of alkane hairpin structures for even n. Ref. a [35], ref. b [34], ref. c [32]

with $n = 12$ is less open in comparison to the MP2 and B2PLYP-D3 structures, which can be also seen by comparison of the aligned TPSS-D3 and B2PLYP-D3 $n = 12$ hairpins in inset a) of figure 4.10b. The mean C-C distances of the LMP2 structures with $n = 14$ to $n = 22$ are in very good agreement to those of the TPSS-D3 structures. The B2PLYP and the MP2 hairpin structures for $n = 14$ to $n = 18$ show slightly lower mean C-C distances than the respective TPSS-D3 and LMP2 structures. Inset b) of figure 4.10b shows the aligned hairpin structures for $n = 16$. Despite the slight difference in the mean distances, the structures obtained by the different methods align well. For the example of $n = 16$, we also give the individual C-C distances between facing carbons on the two 'arms' with respect to the TPSS-D3 values in table 4.5. The C-C distances (a to f) of the LMP2 structure deviate from the TPSS-D3 results by 0.08 Å to −0.05 Å. The least agreement is found for the MP2 structure with deviations ranging from −0.09 Å to 0.29 Å. The respective B2LYP hairpin shows on average better agreement (−0.05 Å to −0.23Å) to the TPSS-D3 structure, in comparison to the MP2 structure, but also results in a hairpin with 'arms', that are slightly closer to each other. Thus when using the same method for relative hairpin energies, we can expect slightly lower energies, *i.e.* more stable hairpins, in the range $n = 14$ to $n = 18$, for the B2PLYP-D3/def2-QVZP structures of Liakos *et al.* [35] and the MP2/VTZ structures of Byrd *et al.* [34] in comparison to our TPSS-D3/def2-TZVP structures. In the same range, we expect the effect of structural differences to be negligible when comparing results based on the TPSS-D3/def2-TZVP structures to those based on the LMP2/VTZ structures of Lüttschwager *et al.* [32]. The largest deviation between hairpin structures is found for $n = 20$, where the 'arms' of the B2LYP-D3 structure are in significantly closer contact than in our TPSS-D3 structure, as depicted in inset c) in figure 4.10b. Considering the discussed differences, we have to be particularly careful when comparing results for $n = 12$ and $n = 20$. The relative hairpin energies obtained by TPSS-D3 will be discussed and compared to the LCC results at a later point in the text. They are given in figure 4.14 and table 4.7.



(a) Change of the relative hairpin energy of n-alkanes with the chain length n. Linear and hairpin structures were optimized at the TPSS-D3/def2-TZVP level of theory. LMP2-F12/AVTZ results with standard domains are given by blue triangles. LMP2-F12/AVTZ results with extended orbital domains are given by red triangles. LCC results are given by black squares.

(b) Change of components of the LCC energy relative to LMP2/AVTZ results with n. Explicit correlation (F12) decreases ΔE, whereas the coupled cluster correction δCC increases ΔE.

Figure 4.11: Energy of the n-alkane hairpin conformers relative to the linear conformers. Results for LCC and the TPSS-D3/def2-TZVP structrues.

Relative n-alkane hairpin energies (ΔE) obtained by LCC: Here we present the results for the relative hairpin energy (ΔE) obtained for the TPSS-D3/def2-TVZP optimized structures with the LCC approximation for n-alkane chains with even n. As shown in equation 4.15, LCC consist of an LMP2-F12/aVTZ part and a δCC/aVDZ correction. Concerning the LCCSD(T0)/aVDZ energy for the δCC/aVDZ correction, we choose to treat pairs of orbital domains within a distance of 4.23 Å as strong, thus at the local coupled cluster level of theory ($R_{\text{close}} = 8$ a$_0$). Pairs between 4.23 Å and 7.94 Å were treated as close pairs ($R_{\text{weak}} = 15$ a$_0$), *i.e.* the MP2 amplitudes influence the coupled cluster triples amplitudes of the strong pairs. All other pair-domains were treated at the LMP2 level of theory. The same cut-off for strong pairs was also employed by Lüettschwager *et al.* [32] and ensures that the most of the facing C-H and C-C domains in the hairpin are treated at LCCSD(T0) level of theory. We also determined the pair-list for the hairpin conformer and used domain merging to ensure that the same respective pairs of the linear conformer are treated at the same level of theory. We found however, that omitting the above step does not introduce a notable error in the δCC/aVDZ part of ΔE. For the LMP2-F12/aVTZ part, we use extended domains (`IEXT=1`) and treat all pair-domains at the LMP2-F12 level of theory.

Figure 4.11a shows the dependence of the relative hairpin energy (ΔE) on the chain length n, obtained by LMP2-F12/aVTZ for standard domains (`IEXT=0`), extended domains (`IEXT=1`) and by LCC. The use of extended domains in the LMP2-F12/aVTZ lowers ΔE considerably, that is by on average -0.27 kJ/mol per added pair ($\Delta n = 2$) of methylene (-CH_2-) groups. For $n = 16$ the domain extension results in an additional hairpin stabilization of -1.52 kJ/mol, with respect to the standard domains. Thus, it is of similar importance as the explicit correlation treatment, which for $n = 16$ leads to a stabilization by -1.18 kJ/mol with respect to the

LMP2/aVTZ (`IEXT=1`) energy, see also 4.11b. The CABS-singles correction to the HF part of ΔE gives a negligible contribution to ΔE, figure 4.11b. Thus in the case of n-alkanes, HF/aVTZ already gives results for ΔE close to the HF/CBS limit. From here on, we will omit the (`IEXT=1`) label and refer to LMP2-F12/aVTZ (`IEXT=1`) as just LMP2-F12/aVTZ. In comparison to LCC, LMP2-F12/aVTZ alone underestimates ΔE, *i.e.* overestimates dispersion interactions. The coupled cluster correction to ΔE in the form of δCC/aVDZ thus raises ΔE by 1.2 kJ/mol for $n = 8$ to 2.7 kJ/mol for $n = 18$, which is about 0.3 kJ/mol per chain elongation by $\Delta n = 2$, see figure 4.11b. To investigate the error introduced by δCC/aVDZ, we have calculated the δCC/aVTZ contribution to ΔE for $n = 8$ to $n = 12$, figure 4.11b. The δCC/aVTZ and δCC/aVDZ corrections give similar values. For $n = 12$ the difference between δCC/aVTZ and δCC/aVDZ is 0.24 kJ/mol. Assuming a linear increase in this error with even n, the extrapolated error for $n = 18$ is 0.60 kJ/mol. On this basis, a conservative global error estimate of ± 1 kJ/mol for ΔE at the LCC level of theory seems reasonable up to $n = 20$.

Due to the overestimation of the hairpin stability, LMP2-F12/aVTZ predicts $n_{\mathrm{hp,el}} = 14$ as the shortest n-alkane of even chain length where the hairpin has a lower electronic energy than the linear all-trans conformer. Respectively, the last globally stable linear n-alkane with even n at the LMP2-F12/aVTZ level of theory is $n_{\mathrm{c,el}} = 12$. Addition of the δCC/aVDZ correction to the LMP2-F12/aVTZ energies gives the LCC result, with $n_{\mathrm{hp,el}} = 16$ and $n_{\mathrm{c,el}} = 14$. The dependence of LCC results for ΔE on even n shows two approximately linear gradients. Starting at $n = 8$ with $\Delta E(8) = 6.57$ kJ/mol, ΔE changes by about -1 kJ/mol per $\Delta n = 2$ until $n = 12$. From $n = 12$ to $n = 18$ the average change in ΔE is -3.5 kJ/mol per $\Delta n = 2$. The occurrence of the two slopes can be explained by the 'closing' of the hairpin structures and the increased orientation of their arms for increasing n, as shown in figures 4.9 and 4.10a.

Comparison of ΔE_{LCC} to reference results obtained by others: We are now in the position to compare the LCC results for ΔE, to wave-function-based reference results obtained by Lüttschwager *et al.* [32], Byrd *et al.* [34] and Liakos *et al.* [35]. Note, that the herein compared results differ in the underlying structures and methods. Lüttschwager *et al.* [32] used an approach similar to LCC. They determined the LMP2-F12/VTZ-F12 energy to which they added a short range coupled cluster correction $\delta\mathrm{CC}_{\mathrm{sr}}$/AVTZ and a long range coupled cluster correction $\delta\mathrm{CC}_{\mathrm{lr}}$/VTZ = LCCSD(T0)$_{\mathrm{lr}}$/VTZ − LCCSD(T0)$_{\mathrm{sr}}$/VTZ. Explicit correlation at the LMP2-F12 level was restricted to orbital pairs with atom distances of up to 4.23 Å. For the short range LCCSD(T) energy the standard pair approximation was applied ($R_{\mathrm{close}} = 1$ a$_0$, $R_{\mathrm{weak}} = 3$ a$_0$), while for the long range correction orbital pairs with atom distances up to 4.23 Å ($R_{\mathrm{close}} = 8$ a$_0$) were treated at the coupled cluster level of theory. The main difference of this method to LCC is the use of standard domains instead of extended domains in the LMP2-F12 calculation and the separation of short range and long range local coupled cluster contributions. Upon comparison of the present LCC results to the results by Lüttschwager *et al.* [32] for ΔE for $n \geq 14$, as shown in figure 4.12, we notice that method used by Lüttschwager *et al.* [32] gives relative hairpin energies that are about 2 kJ/mol higher than our LCC results. Since the comparison of the LMP2/VTZ hairpin structures used by Lüttschwager *et al.* [32] and our TPSS-D3/def2-TZVP structures showed that they are very similar, the difference in the results, are most probably based on the difference in the composite local correlation methods. Unfortunately, we know only the hairpin structures from reference [32] and not the linear structures, thus we cannot repeat the LCC calculation for these structures. Omission of

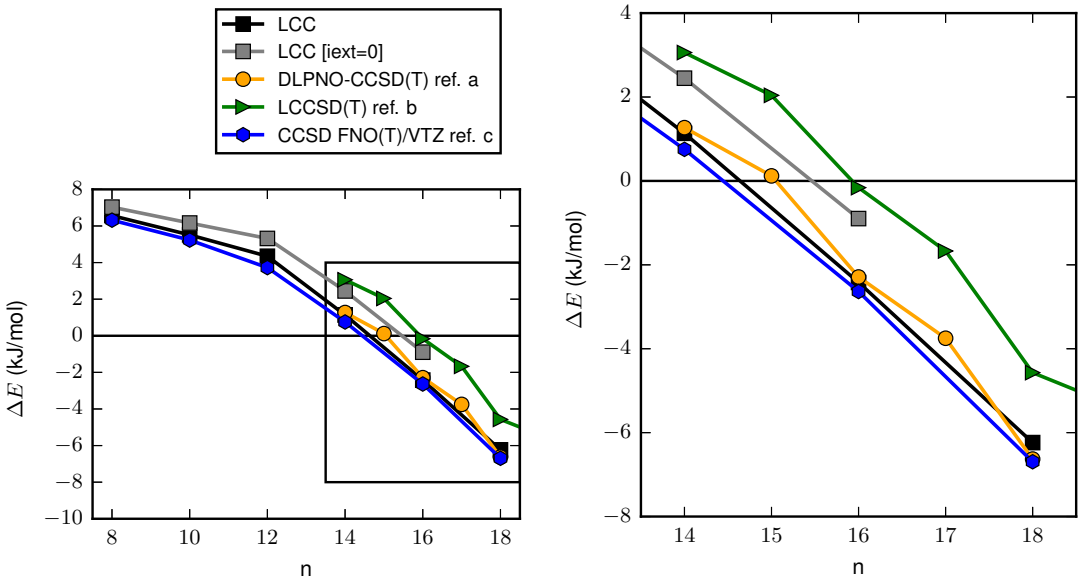
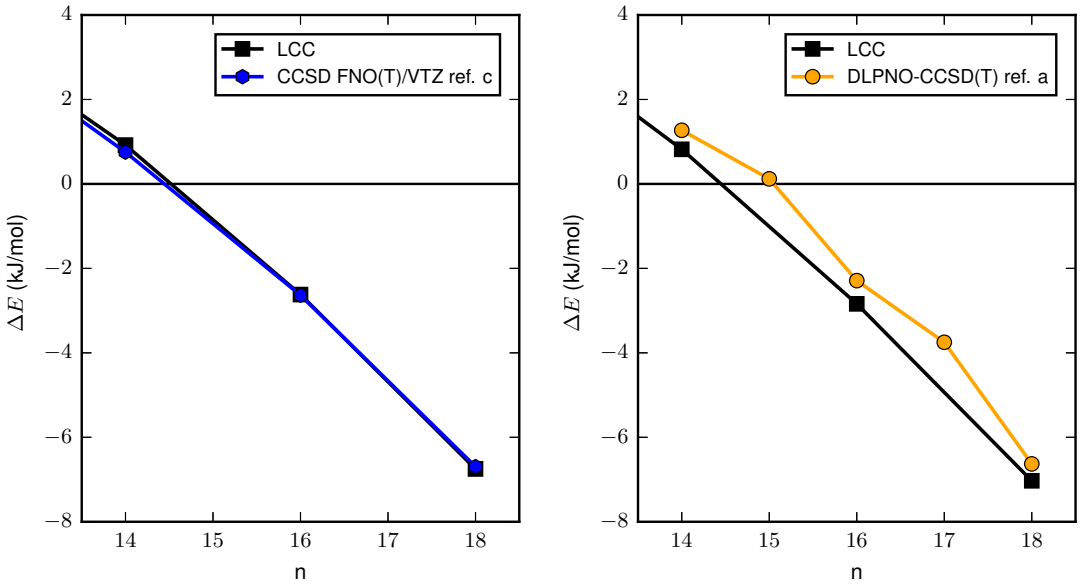


Figure 4.12: Comparison of relative n-alkane hairpin energies obtained by LCC with and without domain extension to the wave-function-based results from references a) [35], b) [32] and c) [34]. Note that in each case different methods for structure optimization have been applied, see also figure 4.10.

the domain extension in the LMP-F12/aVTZ part of the LCC calculation increases ΔE, the respective LCC (`IEXT=0`) relative hairpin energies agree with the values of Lüttschwager *et al.* within 1 kJ/mol, see figure 4.12. We therefore argue, that the main discrepancy between the results of Lüttschwager *et al.* [32] and the present LCC results originate from a domain error, which is reduced in LCC by the use of extended LMP2-F12 domains. The remaining difference of about 0.6 kJ/mol might be attributed to the different approaches used to obtain the coupled cluster corrections, which we will not analyse further.

Byrd *et al.* calculated ΔE for n-alkanes with even n ranging from $n = 8$ to $n = 18$ using the CCSD FNO(T)/VTZ level of theory [34]. They compared estimated CCSD(T)/CBS results for $n = 8$ to $n = 14$, with CCSD FNO(T)/VTZ and assign a global error estimate of ± 1 kJ/mol. The present LCC results for ΔE are in excellent agreement with the CCSD FNO(T)/VTZ results of Byrd *et al.* [34], see figure 4.12. The mean absolute deviation of LCC with respect to CCSD FNO(T)/VTZ is 0.26 kJ/mol, where this value also contains deviation due to the structural differences. We have therefore calculated the LCC ΔE using the MP2/VTZ structures of Byrd *et al.* [34]. Comparing results for the same structures reduces the mean absolute deviation of LCC and CCSD FNO(T)/VTZ to 0.12 kJ/mol, see figure 4.13 and table 4.6. The mean absolute deviation between the LCC energies obtained for the MP2/VTZ and TPSS-D3/def2-TZVP structures is 0.28 kJ/mol. Here the maximal deviations of about 0.5 kJ/mol arise for $n = 12$ and $n = 18$, whereas the absolute differences for the other chain lengths are 0.2 kJ/mol. Thus, the structural differences in the linear and hairpin conformers of n-alkanes resulting from optimization with either MP2 or TPSS-D3, do not lead to errors outside of the estimated global error bounds of LCC. Consequently, both methods, *i.e.*



(a) ΔE for the MP2/VTZ structures of reference c) [34]

(b) ΔE for the B2PLYP-D3/def2-QVTZ structures of reference a) [35]

Figure 4.13: Comparison of relative n-alkane hairpin energies obtained for the same structures.

CCSD FNO(T)/VTZ and LCC, give the same result for the critical chain lengths: $n_{\mathrm{hp,el}} = 16$ and $n_{\mathrm{c,el}} = 14$.

Liakos *et al.* gave the most recent estimate for ΔE using the advanced DLPNO-CCSD(T) theory extrapolated to the CBS limit using VTZ and VQZ basis sets and adding a CCSD(T) correction to DLPNO-CCSD(T)/CBS using the AVDZ basis set [35]. From an extensive error analysis for small n-alkanes, they estimate the error with respect of their method to CCSD(T)/CBS to be 0.5 kJ/mol for $n = 17$, on the side of overestimating the hairpin stability [35]. Their results for all n between $n = 14$ to $n = 18$ are compared to our LCC results in figure 4.12. We find that LCC is in good agreement with the DLPNO-CCSD(T) results of Liakos *et al.* [35], as the mean deviation is 0.21 kJ/mol, which includes deviations due to structural differences. Comparing both methods for the B2PLYP-D3/def2-QZVP structures [35] gives a slightly higher mean absolute deviation of 0.47 kJ/mol, where LCC predicts lower relative hairpin energies than DLPNO-CCSD(T), see figure 4.13 and table 4.6. The mean absolute deviation between LCC results calculated for the B2PLYP-D3/def2-QZVP optimized structures with even n for $n = 8$ to $n = 18$ of Liakos *et al.* [35] and our TPSS-D3/def2-TZVP structures is 0.33 kJ/mol. Irrespective of the two choices for structure optimization, the relative hairpin energies for n-alkanes from LCC and DLPNO-CCSD(T) agree within their accuracy estimates. Thus, both methods predict the same even critical chain lengths.

The good agreement of LCC with the accurate non-local CCSD FNO(T)/VTZ and the local DLPNO-CCSD(T) theory for ΔE confirms the accuracy of the method and our choices for the local approximations in the LCC method for n-alkanes. We could also confirm that the TPSS-D3/def2-TZVP structures are of good quality. On this basis, we estimate the error of our approach to be ± 1.5 kJ/mol. Nevertheless, we have confirmed the reference character of the LCC results which we will compare to more approximate but more economic methods, in the form of dispersion corrected DFT theory.

structure ΔE / n	TPSS-D3	MP2 [34]			B2PLYP-D3 [35]		
	LCC	CCSD FNO(T) [34]	LCC	Δ	DLPNO-CCSD(T) [35]	LCC	Δ
8	6.57	6.32	6.50	0.18		6.54	
10	5.51	5.23	5.48	0.25		5.53	
12	4.35	3.72	3.80	0.08		3.91	
14	1.14	0.75	0.92	0.17	1.27	0.82	−0.45
16	−2.42	−2.64	−2.62	0.01	−2.29	−2.84	−0.55
18	−6.24	−6.69	−6.75	−0.06	−6.63	−7.03	−0.40

Table 4.6: Relative n-alkane hairpin energies (in kJ/mol) obtained at the LCC level for the present TPSS-D3/def2-TZVP structures, the MP2/VTZ structures of Byrd *et al.* [34] and the B2PLYP-D3/def2-QZVP structures of Liakos *et al.* [35]. Additionally the CCSD FNO(T) results obtained by Byrd *et al.* [34] and those obtained by Liakos *et al.* [35] with DLPNO-CCSD(T) are given for n =14, 16 and 18. Differences of LCC with respect to the results from the literature are denoted by Δ.

Performance of three dispersion corrections with TPSS and PBE DFT functionals for the relative hairpin energy of n-alkanes: Density functional theories with (m)GGA functionals fail at describing long range dispersion interactions, due to the non-local correlation nature of these interactions. A remedy for this shortcoming are dispersion corrections (D3, QHO) and non-local correlation functionals (VV10). Because dispersion corrected (m)GGA DFT scales much lower than typical wave-function-based correlation methods, it can be applied to larger systems and used for computationally demanding tasks such as structure optimizations and frequency analysis of large molecules. Here, we compare the performance of the TPSS-D3(BJ), TPSS-VV10 and PBE-QHO methods for the relative hairpin energy of n-alkanes, with respect to our accurate LCC results. This comparison is carried out for the TPSS-D3/def2-TZVP optimized structures. Figure 4.14 and table 4.7 show the dependence of ΔE on n as obtained by the three DFT methods, LCC and LMP2-F12/aVTZ, for even n in the range $n = 8$ to $n = 18$.

The relative hairpin energies obtained by TPSS-D3(BJ) and TPSS-VV10 in this range differ by only 1 kJ/mol, where the use of VV10 results in slightly more stable hairpins compared to D3. In comparison to LCC, both methods overestimate the hairpin stability. In the range $n = 8$ to $n = 18$ TPSS-D3 underestimates ΔE on average by 1.5 kJ/mol with respect to LCC, whereas the average error for TPSS-VV10 is 2.4 kJ/mol. Since both methods perform slightly better for the shorter alkanes with $n = 8$ and $n = 10$, we also calculated the mean deviation for the range $n = 12$ to $n = 18$. In this range TPSS-D3 and TPSS-VV10 consistently underestimate ΔE by 2 kJ/mol and 3 kJ/mol, respectively. Thus, both methods give results similar to LMP2-F12/aVTZ.

The PBE-QHO results for short n-alkane hairpins with $n = 8$ and $n = 10$ are similar to the LCC results. For $n = 12$ and $n = 14$ PBE-QHO gives results similar to TPSS-D3 and TPSS-VV10. However for $n \geq 16$ ΔE is significantly underestimated by PBE-QHO. The average error of PBE-QHO with respect to LCC in the range $n = 12$ to $n = 18$ is −4.4 kJ/mol. At $n = 18$ PBE-QHO underestimates the relative hairpin energy by 6.3 kJ/mol.

All three dispersion corrected DFT methods overestimate the relative hairpin stability, PBE-QHO predicts a higher stabilization per chain elongation by two carbons, compared to TPSS-D3

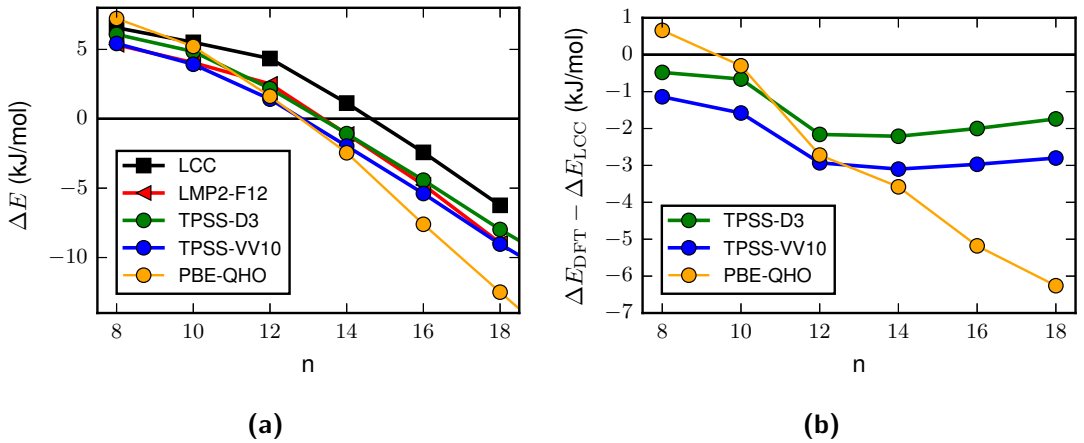


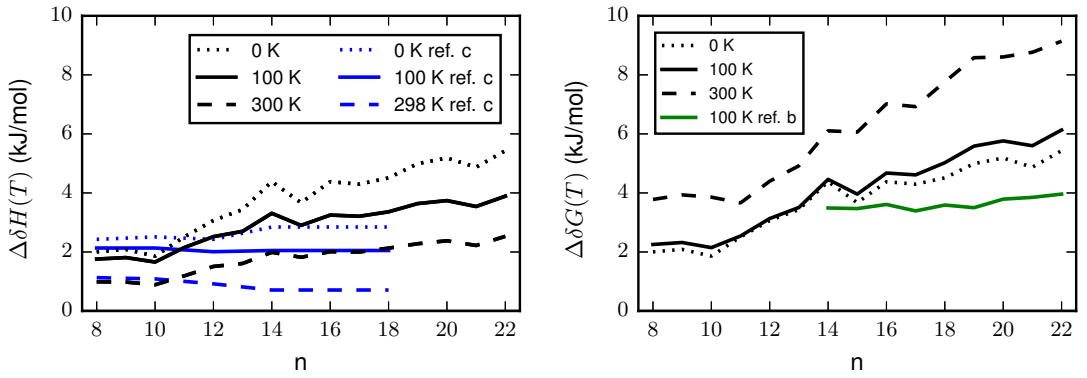
Figure 4.14: Comparison of relative n-alkane hairpin energies (a) for the TPSS-D3 structures obtained by the wave-function-based LCC and LMP2-F12/aVTZ methods and the DFT results obtained with TPSS-D3/def2-TZVP, TPSS-VV10/def2-TZVP and PBE-QHO with a plane wave basis set. (b) Differences with respect to LCC.

n \ ΔE	LCC	TPSS-D3	TPSS-VV10	PBE-QHO
8	6.57	6.09	5.43	7.23
10	5.51	4.85	3.93	5.21
12	4.35	2.18	1.42	1.62
14	1.14	−1.07	−1.97	−2.45
16	−2.42	−4.42	−5.39	−7.60
18	−6.24	−7.98	−9.03	−12.50
20	(−10)	−11.19	−12.28	−17.32
22	(−14)	−14.68	−15.84	−22.22

Table 4.7: Relative n-alkane hairpin energies (kJ/mol) obtained by LCC, TPSS-D3(BJ)/def2-TZVP, TPSS-VV10/def2-TZVP and PBE-QHO with plane wave basis functions. Estimates for the LCC values for $n = 20$ and $n = 22$, given in brackets, have been obtained by linear extrapolation from the $n = 16$ and $n = 18$ data-points.

and TPSS-VV10, where the latter two methods give gradients for ΔE very similar to the LCC results. Concerning the critical chain length for hairpin formation, the three DFT methods give the result $n_{\mathrm{hp,el}} = 14$ and $n_{\mathrm{c,el}} = 12$. We can recommend TPSS-D3 and TPSS-VV10 for the calculation of interaction energies of molecules involving short and long alkyl chains. Both theories are good candidates for studies on the folded conformers of very long n-alkanes, for example double hairpins and paper-clips. For PBE-QHO dispersion interactions between short alkyl chains ≤ 5 are predicted reliably, while interactions between longer chains are overestimated. It should be noted, that double hybrid DFT with the D3 dispersion correction gives ΔE in good agreement to the CCSD(T) FNO(T) and DLPNO-CCSD(T) results [34, 35].

Relative n-alkane hairpin enthalpy and Gibbs free energy: In order to predict the critical chain length for n-alkane hairpin formation at the experimental conditions described by Lüttschwager *et al.* [32, 33], we have calculated thermodynamic corrections to ΔE for the



(a) Temperature and chain length dependence of the enthalpy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies. Ref. c: Byrd *et al.* [34] data based on scaled MP2/VTZ frequencies.

(b) Temperature and chain length dependence of the Gibbs free energy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies and the RRHO entropy correction. Ref. a: Lüttschwager *et al.* [32] based on LMP2/VTZ data.

Figure 4.15: Thermodynamic corrections to ΔE for n-alkanes.

entropy and the Gibbs free energy at $T = 100$ K based on structures and unscaled normal mode frequencies obtained at the TPSS-D3/def2-TZVP level of theory. Additionally, we also calculated the corrections at 0 K and 300 K. The resulting thermodynamic corrections are shown in figure 4.15. Qualitatively, we obtain a similar trend for the temperature dependence of the enthalpy correction $\Delta\delta H$ as Byrd *et al.* [34]. In the considered window of temperatures $\Delta\delta H$ is positive, hence hairpins are destabilized with respect to ΔE. With raising temperature $\Delta\delta H$ decreases. For shorter n-alkanes ($n \leq 10$) our corrections are similar to those of Byrd *et al.*, but they increase for longer n-alkanes, while those of Byrd *et al.* stay constant in comparison. At $n = 18$ our value for $\Delta\delta H(100\text{ K})$ is 3.3 kJ/mol while the value of Byrd *et al.* is 2.1 kJ/mol. Including also an estimation for the entropic contributions to the hairpin stability, we obtain a Gibbs free energy correction to ΔE termed $\Delta\delta G(T)$, see figure 4.15. As for the enthalpy, the entropy difference between the linear and the hairpin conformer results in a destabilization of the hairpin conformer. For low temperatures ($T = 100$ K) the entropic destabilization ranges from 0.5 kJ/mol to 2 kJ/mol, depending on the chain length, thus it is smaller than the enthalpic destabilization. At 300 K, the entropic destabilization clearly dominates $\Delta\delta G(T)$. We therefore expect that the critical chain length for n-alkane hairpin formation will shift to higher n for increased temperatures. Because the entropic destabilization also increases with n, any hairpin formation will be disfavoured for temperatures above a certain critical temperature. Lüttschwager *et al.* predicted $\Delta\delta G(100\text{ K})$ in the range of $n = 14$ to $n = 22$ to be about 3.5 kJ/mol, based on LMP2/VTZ data [32], see figure 4.15. Liakos *et al.* give values for $\Delta\delta G(100\text{K})$ in the range of $n = 8$ to $n = 19$, ranging from about 2.3 kJ/mol to about 3.4 kJ/mol, using PW6B95-D3/def2-QZVP results [35]. Our results agrees with the correction obtained by Liakos *et al.* in the range $n = 8$ to $n = 12$. For longer n-alkanes our results overestimate $\Delta\delta G(100\text{ K})$ in comparison to the results by Lüttschwager *et al.* and Liakos *et al.*, for example by about 2 kJ/mol at $n = 18$. These differences arise due to many factors, and is difficult to decide, which methodology is suited best. PW6B95-D is an excellent

approach for the thermochemistry of alkanes and outperforms TPSS-D in terms of atomization energies, linear to branched isomerization reaction enthalpies [190]. For isodesmic reactions (formal bond breaking reaction of an alkane with methane to give ethane) TPSS-D performs better than PW6B95-D [190].

In order to give an error estimate for the thermodynamic correction $\Delta\delta G(T)$ obtained with TPSS-D3 we turn to the free energy of association of a crown ether - alkyl ammonium ion supramolecular complex. For this complex, Achazi *et al.* [191] determined $\Delta G(635\text{ K})$ using TPSS-D3/def2-TZVP with the same methodology as applied here, and reported LCCSD(T)/CBS reference values for ΔE. The calculated free energies were also compared to experimental gas-phase values obtained in a mass-spectroscopy experiment [192]. Form $\Delta E_{\text{LCCSD(T)}} - \Delta G^{T}_{\text{exp}}$ an accurate value for $\Delta\delta G_{\text{ref}}$ can therefore be compared to the $\Delta\delta G_{\text{DFT-D3}}$ value. For the complex in ref. [191, 192] $\Delta\delta G_{\text{DFT-D3}}$ based on TPSS-D3 has an estimated relative error of 7%. On this basis, we estimate the relative error of $\Delta\delta G(T)$ in the case of the hairpins as 10%. The estimated errorbars of $\Delta G(T)$ are therefore $\pm(1.5 + 0.1\Delta\delta G(T))$. The first term includes the error of LCC and the variability due to structural differences, while the last term denotes the uncertainty in the thermodynamic correction.

Our composite estimate for the enthalpy and Gibbs free energy for the n-alkane hairpin stability relative to the linear conformers at $T = 100$ K is shown in figure 4.16. For comparison the figure also shows the results for ΔH of Byrd *et al.* [34] and ΔG by Lüttschwager *et al.* [32]. The DLPNO-CCSD(T) results obtained by Liakos*et al.* [35] are not shown, they however lie in-between the curves of Lüttschwager *et al.* and Byrd *et al.* [34] and are similar to our results for ΔH. The ΔG curves obtained by Liakos *et al.* [35] and Lüttschwager *el al.* [32] lie within the error bounds of our results. The results for $\Delta G(100\text{ K})$ obtained by TPSS-D3/def2-TZVP closely follow the results obtained by the wave-function-based methods. The reason for this is that the underestimation of ΔE is cancelled by the overestimation of $\Delta G(100\text{ K})$, hence the uncertainty of the TPSS-D3/def2-TZVP results is larger ± 3.5 kJ/mol in comparison to LCC results. Within the estimated errors, our results predict $n_{\text{hp},G(100\text{ K})} = 18$ as the chain length of the first stable n-alkane hairpin and $n_{\text{c},G(100\text{ K})} = 16$ as the last globally stable linear n-alkane with even n. This prediction is in accordance with the theoretical results by Liakos *et al.* [35] and the theoretical and experimental results of Lüttschwager *el al.* [32]. The results by Byrd *et al.* [34] for ΔH suggest, that the hairpin could be favoured already at $n = 16$. However, within their error bounds (± 1.3 kJ/mol) also $n_{\text{hp},H(100K)} = 18$ is more likely [34]. Liakos *et al.* [35] who calculated ΔG also for odd n, found that hairpin formation might start at $n_{\text{hp},G(100\ K)} = 17$ based on their DLPNO-CCSD(T) results.

Finally, the theoretical predictions for the n-alkane hairpin stability $\Delta G(100\text{ K})$ lead to the following conclusions. Linear conformers with $n \leq 15$ are with certainty more stable than the respective hairpin conformers. The turning point, where the hairpin becomes more stable than the linear conformer lies between $n = 16$-18. With $n = 18$ being the most likely, based on the most accurate results obtained by Liakos *et al.* [35]. To give a definite answer, $\Delta G(100\text{ K})$ in this range has to be predicted with a certainty of ± 0.5 kJ/mol, which would foremost require an more exact treatment of thermodynamic corrections. All four studies agree, that for $n \geq 18$ the hairpin conformers are more stable than the linear conformers at $T = 100$ K.

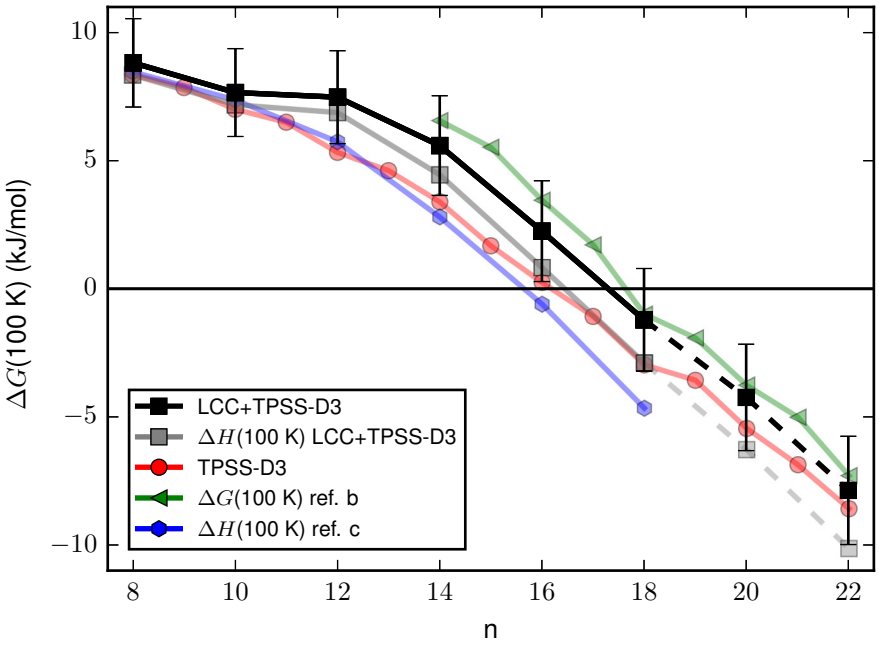


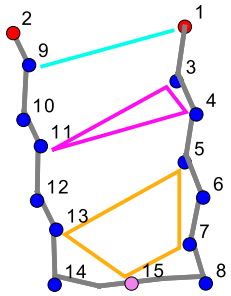
Figure 4.16: Chain length dependence of the the relative Gibbs free energy for n-alkane hairpins at 100 K. Additionally we show also the respective enthalpy. The results for $n = 20$ and $n = 22$ are based on linearly extrapolated values for ΔE, see table 4.7. Ref. b: Lüttschwager *et al.* [32], ref. c: Byrd *et al.* [34]. The results by Liakos *et al.* [35] lie between those of ref. b and c.

n \ ΔE	LMP2-F12	LCC	$\Delta H(100\ \mathrm{K})$	$\Delta G(100\ \mathrm{K})$
8	5.34	6.57	8.33	8.82
10	4.05	5.51	7.17	7.66
12	2.51	4.35	6.87	7.48
14	−1.10	1.14	4.45	5.59
16	−4.76	−2.42	0.83	2.25
18	−8.95	−6.24	−2.88	−1.21
20	-	(−10)	(−6.26)	(−4.24)
22	-	(−14)	(−10.11)	(−7.87)

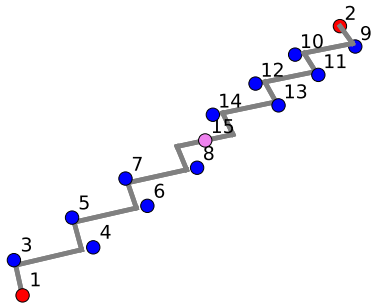
Table 4.8: Relative n-alkane hairpin energies (kJ/mol) obtained by LMP2-F12/aVTZ, LCC, and by adding thermodynamic TPSS-D3/def2-TZVP corrections to LCC ΔE giving $\Delta H(100\ \mathrm{K})$ and $\Delta G(100\ \mathrm{K})$ Estimates for the LCC values for $n = 20$ and $n = 22$, given in brackets, have been obtained by linear extrapolation from the $n = 16$ and $n = 18$ data-points.

4.3.2.1 Method of increments approach to LCC ΔE

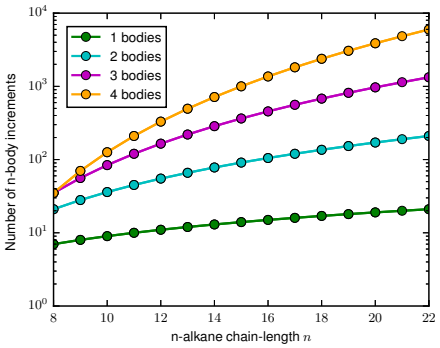
We have demonstrated that the relative hairpin energy of n-alkanes can be accurately described with the composite local correlation LCC method. Next, we investigate the correlation energy part of LCC ΔE as obtained by the method of increments. For this, we will determine which n-body level is necessary to accurately describe the two parts, *i.e.* LMP2-F12/aVTZ and δCC/aVDZ, of the LCC correlation energy difference between the two conformers. Exemplary for $C_{16}H_{34}$ figure 4.17 shows the backbone superimposed by the charge centres of the local orbital groups that define the 1-body increments in our partitioning scheme. For both the hairpin and the linear conformer the internal numbering of the 1-body increments along the backbone is the same. As the 2-body, 3-body and n-body increments are generated by pairing, tripling and so forth of 1-bodies, we obtain the same lists of n-body increments for both con-



(a) $C_{16}H_{34}$ hairpin conformer backbone (gray) superimposed by coloured circles representing the charge centres of the one-body increments. Arbitrary 2-body, 3-body and 4-body increments are visualized by coloured lines connecting 1-body increments.



(b) Like (a) but for the linear conformer of $C_{16}H_{34}$.



(c) Semilogarithmic plot of the dependence of the number of n-body increments on the n-alkane chain length.

Figure 4.17: (a,b): Schematic depiction of increments for $C_{16}H_{34}$. (c) Numbers of increments for n-alkanes.

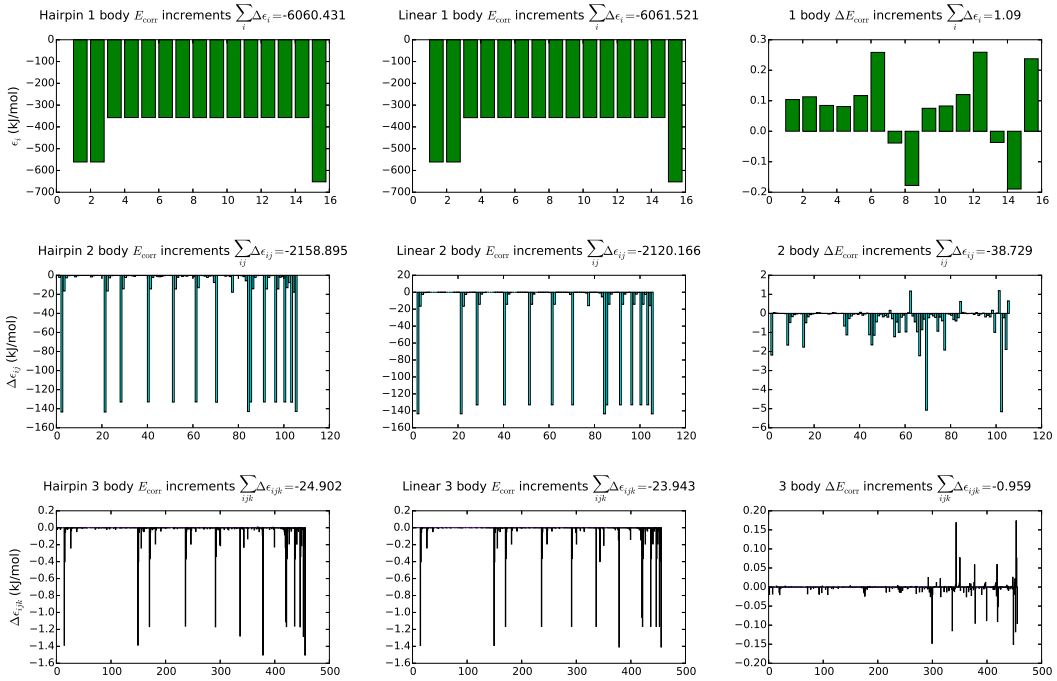


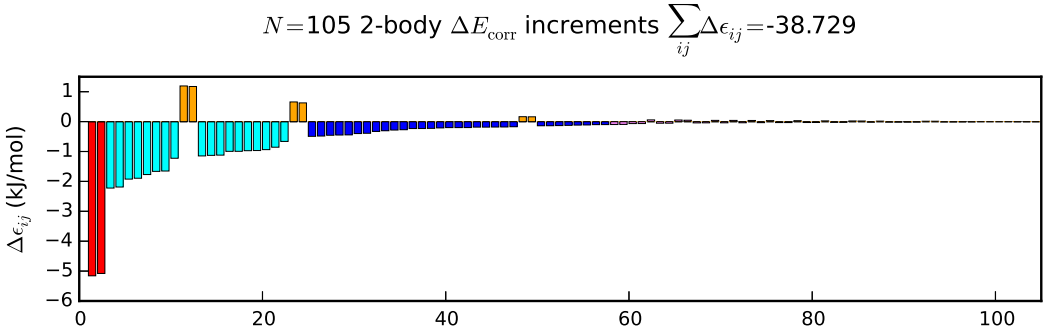
Figure 4.18: Bar plots of individual LMP2-F12/aVTZ [IEXT=1] correlation energy increments for $C_{16}H_{34}$. Columns: hairpin (left), linear (middle) and ΔE_{corr} (right). Rows: 1-body (top), 2-body (middle) and 3-body (bottom). Also given are the sums over the respective n-body increments in (kJ/mol).

formers. Thus, ΔE_{corr} is separated into incremental contributions and thier importance can be determined. One challenge in incremental calculations is that the number of 3-body and higher order body increments for even medium sized molecules is rather large, see figure 4.17. For $C_{16}H_{34}$ the number of 1-body, 2-body, 3-body and 4-body increments is 15, 105, 455 and 1365, respectively. Thus, to obtain the correlation energy at the 4-body level 1940 single point calculations have to be performed for each conformer. Additionally, higher-order increments correlate more electrons than those of lower orders, which makes them computationally more expensive. Moreover, with raising chain length the numbers of increments increase approximately exponentially. The actual number of increments to be evaluated, can be reduced by symmetry considerations, distance criteria, embedding schemes and extrapolation methods. Which of these methods applies, depends on the systems. The structures of the linear conformers of the *n*-alkanes can be set-up as to provide symmetry equivalent increments. However, as we intend to apply the method to mainly non-symmetric hairpin molecules, namely perfluoroalkylalkanes and diperfluoroalkyloctanes as shown in figure 4.3, we omit symmetry considerations in this work. First, we focus on the MoI-LMP2-F12/aVTZ [`IEXT=1`] description of the hairpin and linear conformer of $C_{16}H_{34}$. The respective increments up to the 3-body level are given in figure 4.18. The major part of the correlation energy of the individual conformers is described at the 1-body level, where each 1-body increment gives a contribution approximately proportional to the number of correlated electrons, ranging from about -350 kJ/mol to about -650 kJ/mol. The part of the correlation energy represented by the sum over 2-body increments amounts to about one third of the sum over the 1-body increments.

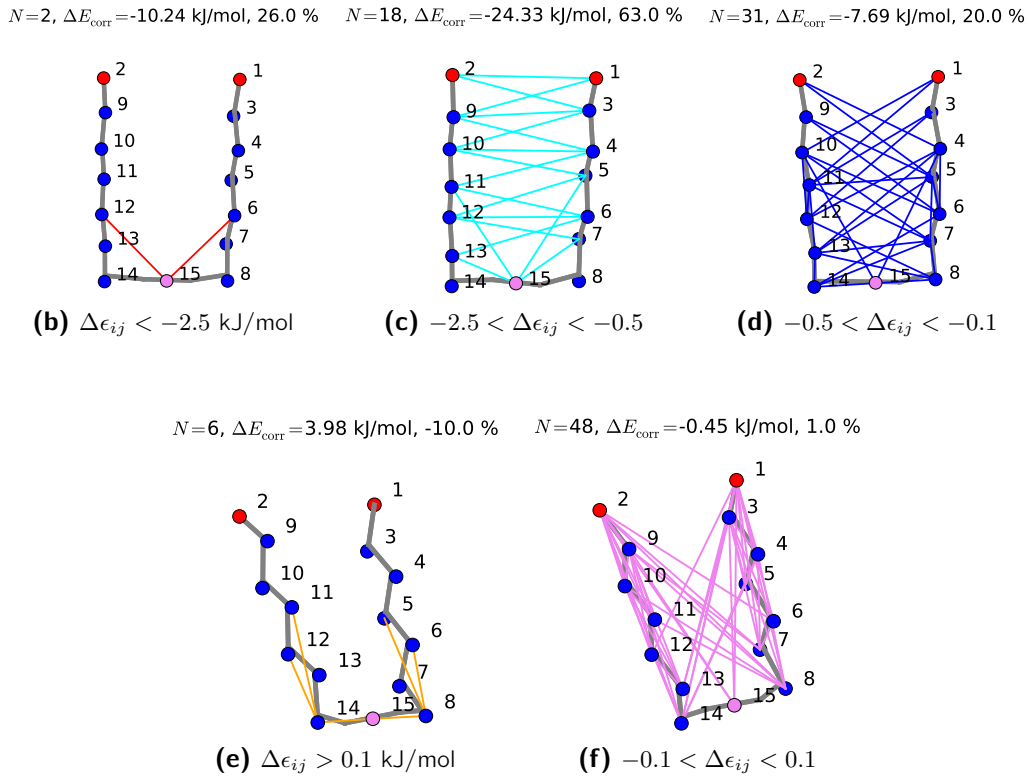
The 2-body increments differ considerably in size, ranging from about −140 kJ/mol to about 0 kJ/mol. Many of the 2-body increments give a negligible contribution at the single conformer level, due to weak correlation between distant 1-bodies. The sum of 3-body increments is two orders of magnitude smaller than the sum over 2-body increments, which indicates convergence of the MoI expansion for the LMP2-F12 correlation energy at the 3-body level. This is the expected behaviour for LMP2, a method which describes mainly pair correlation. Which leads to the question: What is the origin of the 3-body incremental contributions in MoI-LMP2? The answer lies in the different environments present in which pair excitations are treated in LMP2 and MoI-LMP2 [76, 193]. In LMP2, *i.e.* for the whole molecule, the amplitude for a given pair excitation is evaluated in the environment of all other pair excitations in the molecule. In MoI-LMP2, the same pair excitation within a 2-body depends on the other pair-excitations within that 2-body, while the rest of the electrons within the molecule are treated at the HF level. Expanding the 2-body to a 3-body, more correlated electrons are added to the environment of the same pair-excitation compared to the situation in the 2-body, the respective 3-body increment then captures this change in the environment. Thus, 3-body increments will give sizeable contributions, if the respective 2-body increments to be corrected contribute to the correlation energy. Hence, one strategy to reduce the amount of 3-body calculations in MoI-LMP2, is to screen 2-body increments by size and omit 3-bodies that correct negligible 2-body increments.

Since, we are interested in the difference between the correlation energies of the hairpin and linear conformers, we now turn to the analysis of the respective increments for the relative hairpin correlation energy, given in the right column of figure 4.18. Despite being very large for the individual conformers, the 1-body increments give almost no contribution to the relative hairpin correlation energy. Size-wise they range between −0.2 kJ/mol and 0.25 kJ/mol, where the 1-body increments with the largest contributions are 6, 8, 12, 14 and 15, which are located where the n-alkane bends. The 3-body increments give a similar small contribution to ΔE_{corr}, with a majority of the increments being negligibly small. Interestingly, the sum of 1-body and 3-body increments effectively cancel each other resulting in a net contribution of only 0.1 kJ/mol to ΔE_{corr}. The LMP2-F12 ΔE_{corr} is thus effectively given by 2-body increments, which sum up to −38.7 kJ/mol for $C_{16}H_{34}$. The 2-body increments for ΔE_{corr} show an interesting structure, with two large increments (−5 kJ/mol), many increments with energies around −1 kJ/mol, four increments with sizable positive energies and many increments which are close to 0 kJ/mol.

To analyse the 2-body increments for ΔE_{corr} in more detail, they are presented in figure 4.19 sorted by their magnitude (a) and superimposed onto the hairpin structure (b-f) for color coded energy ranges. The 2-body increments shown in (b-d) describe the main binding intramolecular dispersion interactions between the two 'arms' of the hairpin. The two largest increments, shown in (b), are 12_15 and 6_15 and can be attributed to dispersion interactions between the central H_2CCH_2 body and the CH_2 bodies which are three carbons further along the chain. These two increments alone contribute 26% of the total 2-body MoI-LMP2-F12 correlation energy difference. In total 18 increments, given in (c), fall into the energy range between −2.5 kJ/mol and −0.5 kJ/mol. Most of these increments can be attributed to intramolecular dispersion interactions between neighbouring and second nearest neighbouring pairs of 1-bodies located on each of the two 'arms' of the hairpin. Additionally, four pairs involving the one body increment 15 also fall into this range. These interactions amount to 63% and thus the largest part of the total 2-body sum. A number of 31 increments has energies ranging



(a) The 105 2-body increments for LMP2-F12/aVTZ [IEXT=1] ΔE_{corr} of $C_{16}H_{34}$ sorted by their absolute value. Increments in certain energy ranges are color coded and depicted as lines of the same color superimposed on the $C_{16}H_{34}$ hairpin structure in (b-f).



(b) $\Delta\epsilon_{ij} < -2.5$ kJ/mol **(c)** $-2.5 < \Delta\epsilon_{ij} < -0.5$ **(d)** $-0.5 < \Delta\epsilon_{ij} < -0.1$

(e) $\Delta\epsilon_{ij} > 0.1$ kJ/mol **(f)** $-0.1 < \Delta\epsilon_{ij} < 0.1$

Figure 4.19: Importance of individual 2-body increments for the relative LMP2-F12 correlation energy of the $C_{16}H_{34}$ hairpin. (a) Overview of the increments ordered by magnitude. (b-f) The depiction of the location of 2-body increments within given energy ranges (kJ/mol) provides a map of the mainly dispersive interactions within the hairpin conformer. Also given are the number of increments (N), their sum (kJ/mol) and their fraction (%) relative to the sum over all 2-body increments.

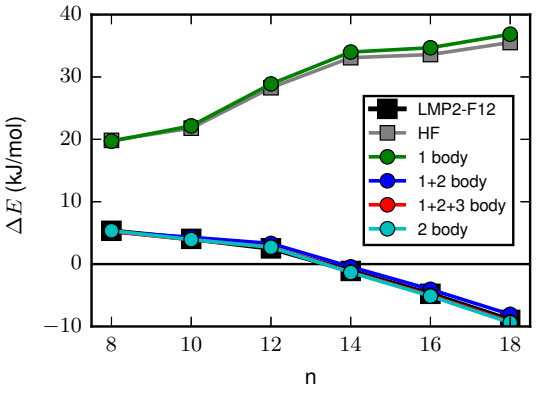
from −0.5 kJ/mol to −0.1 kJ/mol, of which most can be described as third and fourth nearest neighbour interactions of one-bodies located on both 'arms'. These 'weak' interactions amount to 20% of the total 2-body relative hairpin correlation energy. The last sizable contribution of −10% is given by 6 increments with positive energies > 0.1 kJ/mol, which are located along the bend of the hairpin, all involving 1-bodies 8 and 14. The positive contribution can be explained by considering that in the hairpin conformer the C-H bonds at the 'corners' are facing away from the C-H bonds of their next neighbours. This reduces intra-molecular C-H dispersion interactions compared to their more or less parallel arrangement in the linear conformer.

Hence, intra-molecular dispersion interactions stabilize the hairpin conformer, when present in the 'inside' of the bend and destabilize the hairpin, when present on the 'outside' near the 'corners'. However, the destabilization is by far out-weight by stabilizing interactions. Almost half of the 2-body increments, in total 48, fall into the range of very small energies, given in (f), and in sum contribute only 1% to the total 2-body sum. Increments with almost zero effect on the hairpin stability are those correlating the 1-bodies located on the same 'arm', fifth nearest neighbours and more distant pairs between different 'arms'.

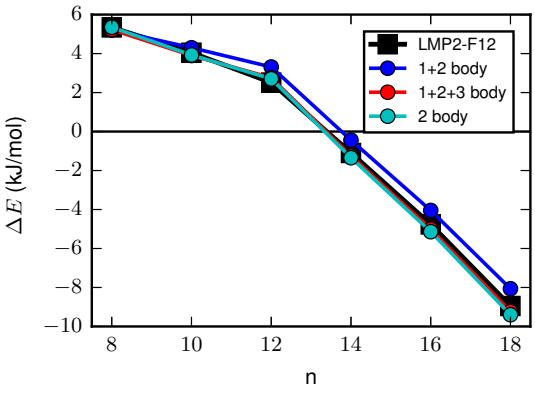
To conclude, for the example of the MoI-LMP2-F12 calculation of $C_{16}H_{34}$ we obtain similar results for ΔE_{corr}, *i.e.* differing by less than 1 kJ/mol from the conventional LMP2-F12 relative hairpin correlation energy −38.36 kJ/mol, by: a) Summing up all 1-body, 2-body and 3-body increments, which is based on $2 \times 575 = 1150$ incremental single point calculations. At the 3-body level, $\Delta E_{\mathrm{corr}} = -38.60$ kJ/mol is obtained, which is the most accurate result within this comparison. b) Only considering the sum over 2-body increments, which reduces the number of incremental calculations to $2 \times 120 = 210$. At this level which we denote 2-body-only $\Delta E_{\mathrm{corr}} = -38.73$ kJ/mol is obtained. c) Evaluation of only those 2-body increments which contribute significantly to conformational correlation energy changes, which further reduces the number of calculations to $2 \times 72 = 144$, *i.e.* 2×15 1-body and 2×57 2-body correlation energy calculations. This approach gives $\Delta E_{\mathrm{corr}} = -38.28$ kJ/mol. d) At the 2-body level, *i.e.* summing over all 1-body, 2-body increments, $\Delta E_{\mathrm{corr}} = -37.64$ kJ/mol is obtained, which also gives sub-kJ/mol accuracy. The deviation at the 2-body level with respect to the reference value is larger, compared to the 2-body-only approximation, which benefits from cancellation of 1-body and 3-body correlation contributions.

The interaction maps shown in figure 4.19 would provide more detail if each 1-body would represent a single local orbital. In this case, the analysis would show also σ_{CC} and σ_{CH} interactions. On the basis of detailed interaction maps, one could develop a local correlation method for conformational correlation energy changes. Such a method would introduce pair approximations based on the distance, orientation and spread changes of local orbitals with respect to both conformers. Orbital pairs that do not contribute to the correlation energy change could be discarded, if their distance, orientation and spread is not changing within some thresholds, from one conformer to the other. For conformational energy differences in large molecules, where large parts of the structure do not change, such a scheme may provide significant computational savings. The development of the method as described above is beyond the scope of this work, which merely focusses on the brute force incremental approach, and was thus not undertaken.

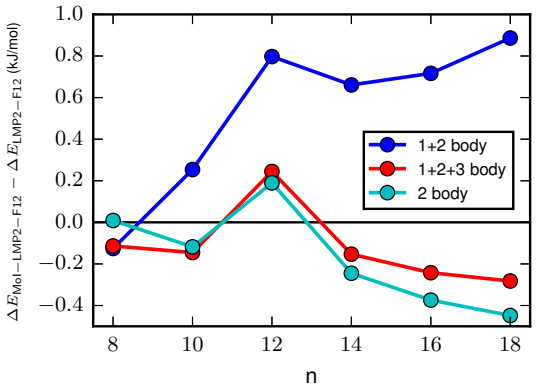
Next, we discuss the MoI-LMP2-F12 results for ΔE for n-alkanes with even n in the range $n = 8$ to $n = 18$. Figure 4.20 provides respective plots for ΔE in (a) and (b). At the 1-body level, the MoI-LMP2-F12 and HF curves coincide, see figure 4.20a, as the sum over the 1-body



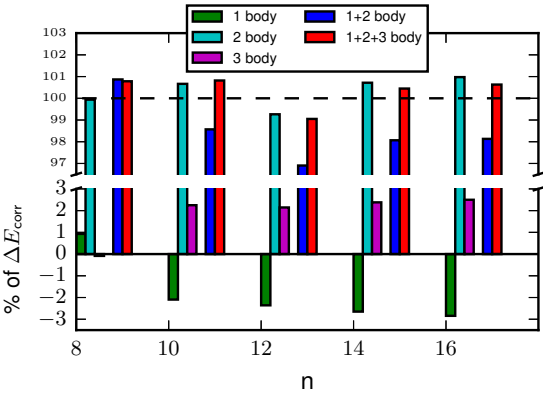
(a) Relative n-alkane hairpin energy at the HF+[CABS singles] and LMP2-F12 level of theory and using the MoI-LMP2-F12 up to the 3-body level. The notations 1+2 body etc. refer to sums over all the respective n-body increments which approximate the LMP2-F12 correlation energy contribution to ΔE.



(b) Zoom-in of the plot in (a).



(c) Absolute error in the relative hairpin LMP2-F12 energy obtained for different MoI approximations.



(d) Fraction of the relative hairpin LMP2-F12 correlation energy obtained at different levels of the MoI.

Figure 4.20: Convergence of the relative hairpin energy of n-alkanes for the MoI-LMP2-F12/aVTZ [iext=1] level of theory with increasing incremental orders.

increments gives a negligible correlation contribution to ΔE. At the 2-body level, summing over all 1-body and 2-body increments, ΔE is overestimated by about 0.8 kJ/mol in the range $n = 12$ to $n = 18$, see also figures 4.20a-c. At the 3-body level ΔE obtained by MoI-LMP2-F12 agrees with the LMP2-F12 results within ± 0.3 kJ/mol. Summation of only the 2-body increments, provides a result similar to the 3-body level MoI-LMP2-F12. Figure 4.20d shows the percentage of the LMP2-F12 correlation energy part of ΔE as obtained by summing over 1-body, 2-body and 3-body increments. For $n = 10$ to $n = 16$ the 1-body sum and the 3-body sum effectively cancel each other, which explains the good performance of the 2-body-only MoI-LMP2-F12. Thus, the analysis performed for $n = 16$ also holds for other chain lengths.

The second component for LCC is the local coupled cluster correction δCC/aVDZ, *i.e.* the post-LMP2 correction. Here we investigate at which level the MoI-δCC converges to the exact result obtained by the standard local correlation treatment. Figure 4.21 shows the

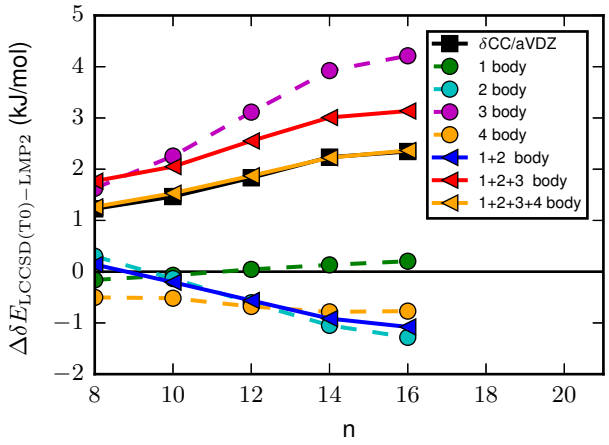


Figure 4.21: Convergence of the MoI-δCC/aVDZ correction for different incremental sums up to the 4-body level.

dependence of MoI-δCC on the chain length and on increasing n-body levels up to the 4-body level. Additionally individual n-body sums for MoI-δCC are given for each chain length. The sum of 1-body increments has an almost zero contribution to MoI-δCC. At the 2-body level, MoI-δCC has the wrong sign with respect to the reference. The absolute values of the 2-body contribution increase with the chain length. At the 3-body level, MoI-δCC shows the correct chain length dependence, but the magnitude of the δCC correction is overestimated by about 1 kJ/mol. It is obvious, that the 3-body sum provides the main part of the MoI-δCC correction. One reason for this is that LCCSD(T0) describes Axilrod-Teller-Muto dispersion, a 3-body body effect, which is not covered by LMP2 [194, 195]. At the 4-body level, the MoI-δCC is converged to the non-MoI reference result. In contrast to sums over the 2-body and 3-body increments the 4-body sum is much less dependent on the chain length. It gives a negative contribution ranging from -0.5 kJ/mol at $n = 8$ to -1 kJ/mol at $n = 16$.

The correct description of the δCC correction within the framework of the MoI thus affords a huge number of incremental calculations. The necessity of up to 4-body increments makes brute force MoI calculations for δCC for large chain molecules therefore impractical. In the present case one could calculate the MoI-δCC at the 3-body level and add the 4-body sum obtained for a smaller molecule, without introducing a large error. Of course, a future analysis of the importance of individual increments for δCC will lead to further insights and a more efficient evaluation of the term. Although this data is available from the present work, at this point we do not undertake this elaborate analysis. Instead of using brute force MoI, we will employ the standard local methods and smaller basis sets to calculate δCC for the systems with perfluoroalkylgroups, additionally we will extrapolate δCC when necessary.

4.3.3 Perfluoroalkylalkane folding

In the preceding section we have demonstrated that the LCC approach gives accurate relative hairpin energies for the alkanes. In this section we apply the method to perfluoroalkylalkane hairpins of the type C_mF_{2m+1}-C_8H_{16}-C_mH_{2m+1}, up to even chain lengths of $n = 22$. First, we begin with a discussion of the structural features obtained at the TPSS-D3/def2-TZVP level of theory. Second, we give the results on the electronic part of the relative hairpin energy, for wave function based methods, *i.e.* the LMP2-F12 and LCC methods. Third, we compare results obtained by dispersion corrected DFT with the LCC results. Fourth, we present results on thermodynamic corrections and on the stability of perfluoroalkylalkane hairpins at finite temperature. Fifth, we analyse the convergence of the MoI-LMP2-F12 results and present a map of the 2-body increments for the intramolecular correlation interaction energy on the hairpin structure of C_4F_9-C_8H_{16}-C_4H_9 ($n = 16$).

Perfluoroalkylalkane hairpin structures Structure-optimizations of the linear and hairpin conformers of the perfluoroalkylalkanes C_mF_{2m+1}-C_8H_{16}-C_mH_{2m+1} have been performed at the TPSS-D3/def2-TZVP level of theory. We choose to limit the investigation to even chain lengths n in the range $n = 10$ to $n = 22$. In the case of $n = 10$, one -CF_3 and one -CH_3 group is added to each end of the central -C_8H_{16}- unit. For $n = 22$, one -C_7F_{15} group and one -C_7H_{15} is added to each end of the central -C_8H_{16}- unit. The chain length m of the perfluoroalkyl group is related to the overall chain length n by,

$$m = \frac{n-8}{2}. \tag{4.16}$$

Although other folded hairpin structures are possible, we limit our investigation to hairpins, which are folded through the central -C_8H_{16}- group. The structures of the investigated perfluoroalkylalkane hairpins are shown up to $n = 20$ in figure 4.22. It should be noted, that an extension of the investigation to odd n would present two new possible hairpins cases for each odd n. Namely, the case where the perfluoroalkyl group is one CF_2 unit longer than the alkyl group and vice versa. Thus, two distinct even-odd effects can be expected for the proposed perfluoroalkylalkane hairpins. Again the main question of this work is: For which even n are hairpin conformers lower in energy compared to linear conformers?

First, we shall discuss the structural changes in the perfluoroalkylgroup with m. From the crystal-structure of polytetrafluoroelthylene (-CF_2-)$_n$ it is known, that the trans-like structure of long perfluoroalkyl groups occurs not for carbon backbone dihedral angles of 180° as in alkanes, but is shifted by about 17° to 163°. This implies a helical structure, where the helix repeats after 13 carbon atoms [196]. The origin of the helicity of perfluoroalkanes is attributed to electrostatic interactions[2], which stabilize the helical conformation over the all-trans structure [197]. Therefore, we have biased our structure optimization towards helical structures, by using initial dihedral angles of ±162° for all perfluoroalkyl groups.

In the case of the perfluoroalkylalkanes we only discuss the results for −162° and note, that the DFT results for the other handedness (with +162°) obtained are similar. For short perfluoroalkyl groups, $m = 1$–$m = 3$, figure 4.22 shows that the structure optimization resulted in a relaxation to non-helical structures. The onset of helicity begins with the perfluorobutyl group ($m = 4$) and is kept for all longer groups. This behaviour was also found in other

[2]Note, that this explanation is based on a force field model.

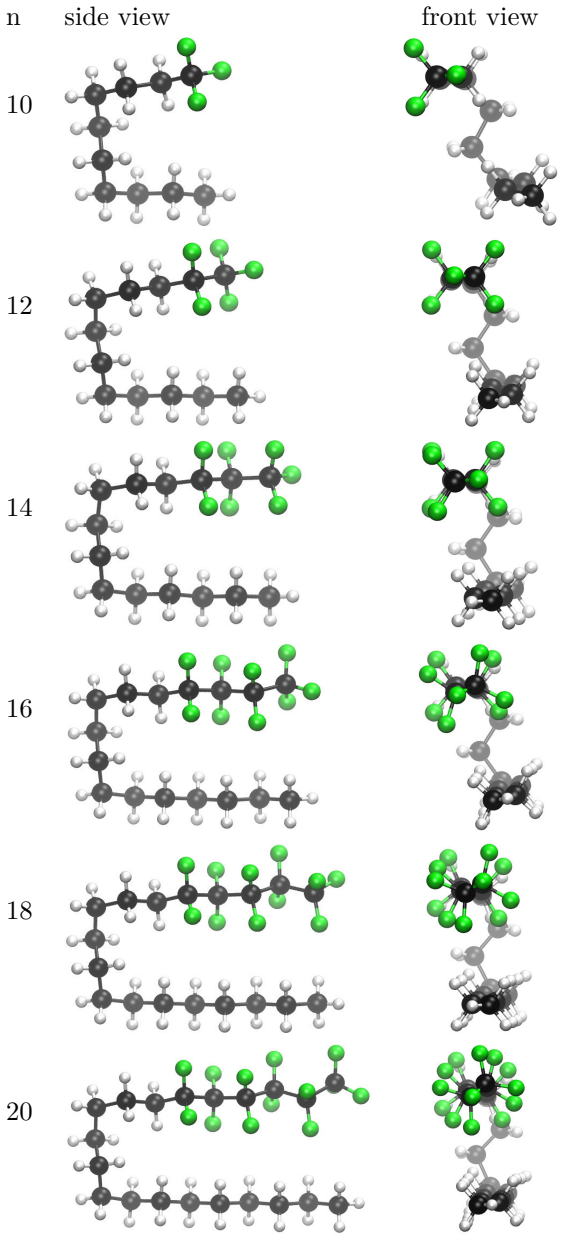
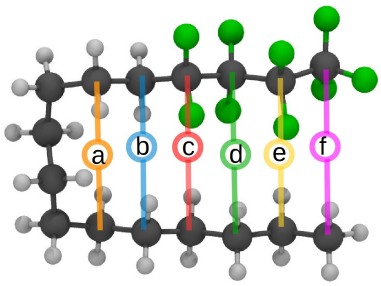


Figure 4.22: Perfluoroalkylalkane (C_mF_{2m+1}-C_8H_{16}-C_mH_{2m+1}) hairpin TPSS-D3/def2-TZVP structures for even n from $n = 10$ to $n = 20$. The length of the perfluoroalkyl group is $m = (n - 8)/2$.

ab initio studies on the structures of perfluoroalkanes [183, 197–199]. The optimized helical dihedral angles range from $-161°$ to $-166°$, with most angles being close to $-163°$. Backbone dihedral angles become closer to $-180°$ for every involved CH_2 carbon, when measured where the perfluoroalkyl group connects to the alkyl chain. Dihedral angles for $n = 16$ are given table 4.9.

The next question is: How close are the arms of the perfluoroalkylalkane hairpins? To answer this question, we measure interarm C-C distances for each hairpin conformer as shown in figure 4.9. Form this data, we calculate mean interarm C-C distances for each chain length n.



Method	a (Å)	b (Å)	c (Å)	d (Å)	e (Å)	f (Å)	∅ (Å)
TPSS-D3/def2-TZVP	4.86	4.34	5.03	4.62	4.88	4.85	4.76

dihedral angles	d_{abcd}	d_{bcde}	d_{cdef}
perfluoroalkyl	-178.2	-168.7	-165.8
alkyl	-179.6	-174.9	-180.0

Table 4.9: The upper table gives C-C distances between carbons on parallel chains of the $n = 16$ perfluoroalkylalkane hairpin, as indicated by the above figure. The mean of the C-C distances is denoted by ∅. The lower table gives carbon dihedral angles for both arms. Carbon atoms are labelled as their distance lines (a,$\cdots$,f) in the figure above.

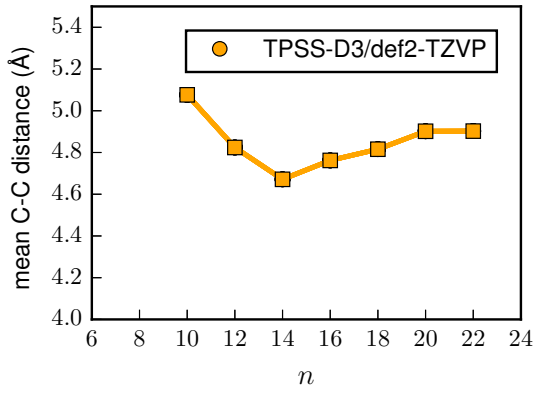
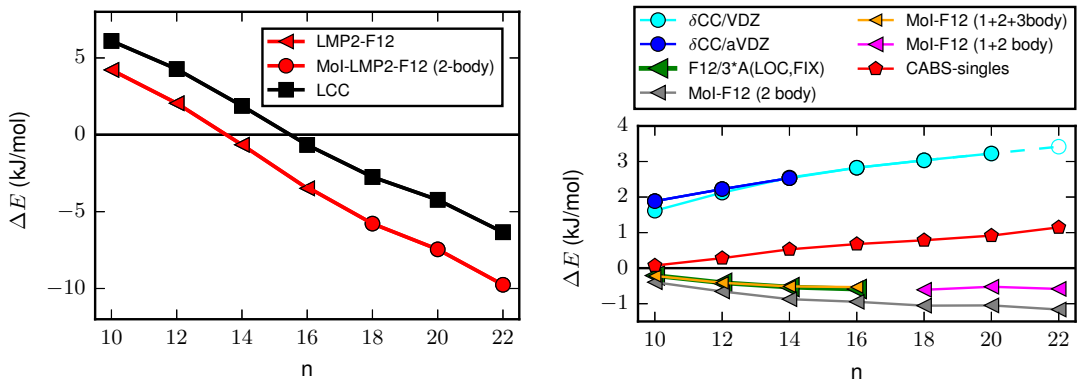


Figure 4.23: Mean interarm C-C distances of facing carbons on adjacent chains in the perfluoroalkylalkane hairpin conformers. For an example with mean and individual C-C distances see table 4.9.

From table 4.9 and figure 4.23 we see, that the mean interarm C-C distances range from 5.1 Å ($n = 10$) to 4.6 Å ($n = 14$). Mean interarm C-C distances, decrease from 5.1 Å at $n = 10$ to 4.6 Å at $n = 14$ and increase from there to 4.9 Å at $n = 22$. Hence, for $n \geq 12$ mean interarm C-C distances of perfluoroalkylalkane hairpins are about 0.5 Å larger than those of the respective alkane hairpins. For $n = 16$ the inter arm C-C distances lie between 4.8 and 5.0 Å, while for the respective alkane the distances range from 4.6 to 4.1 Å. This has two reasons. Firstly, the C-F bond is longer (1.34-1.37 Å) than the C-H bond (1.1 Å), while the CF_2-CF_2 distance of 1.57 Å is comparable the CH_2-CH_2 distance of 1.53 Å. Secondly, the helicity of the perfluroalkylgroups also increases mean interarm C-C distances, compare figure 4.23 and structures in figure 4.22.



(a) Change of the relative hairpin energy of perfluoroalkylalkanes with the chain length n. Linear and hairpin structures were optimized at the TPSS-D3/def2-TZVP level of theory. LMP2-F12/AVTZ results with extended orbital domains are given by red triangles and by red circles as obtained by 2-body only MoI. LCC results are given by black squares.

(b) Change of components of the LCC energy relative to LMP2/AVTZ results with n. Explicit correlation (F12) decreases ΔE, whereas the coupled cluster correction δCC increases ΔE. The δCC value for $n = 22$ has been linearly extrapolated from the values for $n = 18$ and $n = 20$.

Figure 4.24: Energy of the perfluoroalkyalkane hairpin conformers relative to the linear comformers.

Relative perfluoroalkylalkane hairpin energy: LCC results In order to obtain the relative hairpin energy ΔE for the perfluoroalkylalkanes up to $n = 22$, as shown in figure 4.24a, we use local correlation methods and introduce further approximations, relative to the n-alkane case. LMP2-F12/aVTZ results where all pair domains are included have been obtained up to $n = 16$. For perfluoroalkylalkanes with $n > 16$ we use the 2-body only MoI-LMP2-F12/aVTZ approximation for the correlation energy, *i.e.* we neglect the 1-body sum. This approximation gives results in excellent agreement to the true LMP2-F12/aVTZ values. A detailed comparison for $n \leq 16$ is given at the end of this chapter. For MoI and normal LMP2-F12/aVTZ results extended domains (`IEXT=1`) where used in combination with the F12/3*A(LOC,FIX) explicit correlation treatment. The F12 contribution to the relative hairpin correlation energy leads to a small (-0.5 kJ/mol) stabilization of the hairpin conformer and is relatively independent from the chain length, see figure 4.24b. Concerning the MoI results, the F12 correction is nicely recovered at the 2-body level, while its magnitude is overestimated at the 2-body only level. The latter finding partly explains the good agreement of the 2-body only MoI-LMP2-F12 results to the converged 3-body level MoI-LMP2-F12/aVTZ. This is discussed at the end of this section, together with the convergence of the MoI results. The CABS-singles correction to the HF/aVTZ relative hairpin energies is destabilizing. It begins with 0.07 kJ/mol ($n = 10$) and reaches 1.15 kJ/mol at $n = 22$. Due to the similar magnitude of the CABS-singles correction and the F12 correction both cancel out. Therefore, LMP2/aVTZ alone should already be within 1 kJ/mol of the LMP2/CBS limit for ΔE. At the LMP2-F12/aVTZ level, perfluoroalkylalkane hairpins are lower in energy than their respective linear conformer for chainlenghts $n \geq 14$, see figure 4.24a.

The post-MP2 correction δCC was obtained with the aVDZ basis set up to $n = 14$, where pair domains with distances of 5.3 Å were treated at the LCCSD(T0) level ($R_{\text{close}} = 10$ a$_0$) and all other pairs were treated at the LMP2 level ($R_{\text{weak}} = 0$ a$_0$). For chain lengths up to $n = 20$,

δCC was calculated using the VDZ basis, where the distance criterion for the LCCSD(T0) pairs was raised to 6.4 Å ($R_{\text{close}} = 12$ a$_0$). The δCC value for $n = 22$ was linearly extrapolated from the δCC/VDZ values at $n = 18$ and $n = 20$. Both methodologies for δCC give very similar results, as evident by the agreement of their values up to $n = 14$, see figure 4.24b. δCC is the overall most important correction with a destabilizing contribution of 1.9 kJ/mol ($n = 10$) to 3.4 kJ/mol ($n = 22$). The respective linear gradient for δCC is 0.25 kJ/mol per $\Delta n = 2$. LMP2-F12 therefore overstabilizes perfluoroalkylalkane hairpins by the respective amounts given by δCC.

The relative hairpin energy of the perfluoroalkylalkanes obtained via the LCC method, see figure 4.24a, begins at 6.10 kJ/mol for $n = 10$ and decreases approximately linearly by 2 kJ/mol per added pair of a methlyene (CH_2) and a perfluoromethylene (CF_2) group ($\Delta n = 2$). For $n \geq 16$ the perfluoroalkylalkane hairpin conformers are lower in energy relative to the linear conformers. However, the stabilization at $n = 16$ is less pronounced compared to the respective alkane, which is true also at the LMP2-F12 level. Concerning the error estimation for ΔE as obtained by the LCC method for the perfluoroalkanes, we stick to the same error bounds as estimated for the alkanes, *i.e.* $\pm$1.5 kJ/mol.

Relative perfluoroalkylalkane hairpin energy: DFT results With the accurate wave-function based LCC results at hand, we now turn to the evaluation of the DFT results obtained with TPSS-D3/def2-TZVP, TPSS-VV10/def2-TZVP and PBE-QHO with plane waves. Figure 4.25a shows the relative hairpin energy ΔE as obtained by DFT and LMP2 and LCC. Differences of the DFT results with respect to LCC are given in figure 4.25b. As found for the alkanes, TPSS/def2-TZVP with the D3(BJ) correction gives results similar to LMP2, in that the relative hairpin stability is overestimated by 1.8 to 4.8 kJ/mol with respect to LCC. The slope of ΔE with TPSS-D3/def2-TZVP is with -2.5 kJ/mol per $\Delta n = 2$ also similar to the wave function based results. In consequence, TPSS-D3/def2-TZVP predicts $n_{\text{hp,el}} = 14$.

Switching to the non-local dispersion correction functional VV10 with TPSS-VV10/def2-TZVP results in a strong over-stabilization of the perfluoroalkylalkane hairpin conformers with increasing n. The gain in stabilization per $\Delta n = 2$ with TPSS-VV10/def2-TZVP is 3.4 kJ/mol. The first stable hairpin in terms of the electronic energy is predicted for $n_{\text{hp,el}} = 12$. At a chain length of $n = 22$ the deviation of TPSS-VV10 with respect to LCC amounts to -11 kJ/mol.

An even more extreme over-stabilization of perfluoroalkylalkane hairpins is found for PBE-QHO. In this case, the linear slope is -5.2 kJ/mol per $\Delta n = 2$, which leads to $n_{\text{hp,el}} = 12$, as all three DFT methods give comparable results for $n = 10$. At $n = 22$ a deviation with respect to LCC of -21 kJ/mol is reached.

All three dispersion corrections overestimate the intermolecular interactions which stabilize the perfluoroalkylalkane hairpin conformers. Out of the tested methods, only TPSS-D3/def2-TZVP gives results for ΔE within chemical accuracy ($\pm$4.18 kJ/mol) over the whole range of chain lengths. The non-local correlation functional VV10 and the orbital depended QHO correction in conjunction with the TPSS and PBE functionals result in a very pronounced overestimation of perfluoroalkyl-alkyl interactions, where PBE-QHO shows the largest error with respect to LCC. The reason for this behaviour remains an open question. One may speculate, that their parametrization is less ideal with respect to perfluoroalkyl interactions. Another reason could be, that the densities and Wannier orbitals of the perfluoroalkylgroups

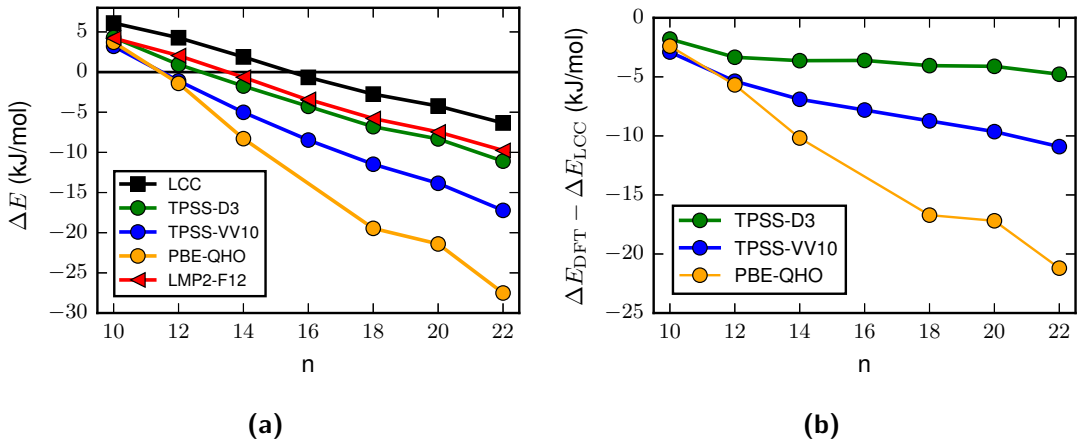
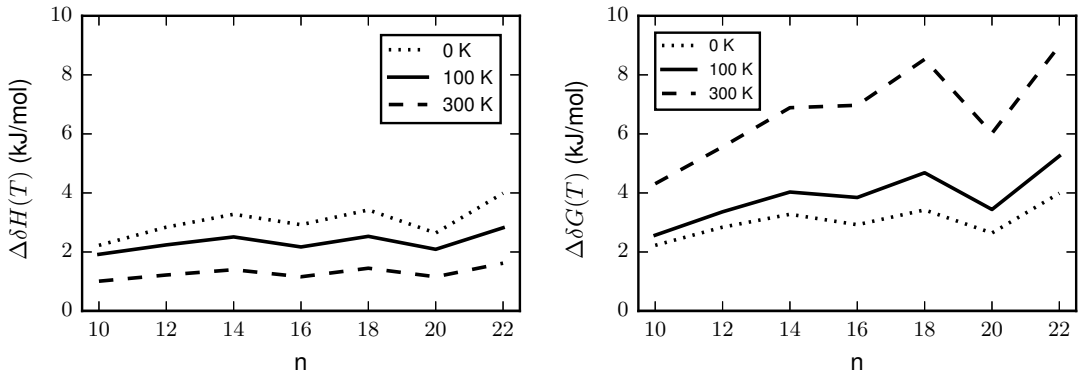


Figure 4.25: Comparison of relative perfluoroalkylalkane hairpin energies (a) for the TPSS-D3 structures obtained by the wave-function-based LCC and LMP2-F12/aVTZ methods and the DFT results obtained with TPSS-D3/def2-TZVP, TPSS-VV10/def2-TZVP and PBE-QHO with a plane wave basis set. (b) Differences with respect to LCC.

obtained from TPSS or PBE are not well suited for the non-local correlation functional or the QHO model, *i.e.* a different functional may give better results.

Relative perfluoroalkylalkane hairpin energy: thermodynamic corrections As for the alkanes, thermodynamic corrections $\Delta\delta H(T)$ and $\Delta\delta G(T)$ have been obtained at temperatures of 0 K, 100 K and 300 K, using the TPSS-D3/def2-TZVP level of theory, without scaling of normal mode frequencies and with the rigid-rotor approximation for low frequency modes. The chain length and temperature dependence of both corrections are given in figure 4.26. The enthalpic correction $\Delta\delta H(T)$ is relatively insensitive to the chain length. With increasing temperature the destabilization due to $\Delta\delta H(T)$ drops from 2-4 kJ/mol ($T = 0$ K), over 2-3.5 kJ/mol ($T = 100$ K) to 1-2 kJ/mol ($T = 300$ K). At $T = 100$ K, entropic effects are comparable to the enthalpic effects, leading to a further destabilization of the perfluoroalkylalkane hairpin conformers by 0.5-2.5 kJ/mol. At $T = 300$ K, entropic effects dominate the $\Delta\delta G(T)$ correction and give rise to a total destabilization by 4-9 kJ/mol, depending on the chain-lengh. Interestingly, the entropy and, to a lesser extend, also the enthalpy correction show an even-odd behaviour. Corrections for perfluoroalkylalkanes with odd m are larger in comparison to those than of neighbouring even m. Compare for example $\Delta\delta G(T)$ for $n = 16$, 18 and 20. The reason for this effect is unclear. Finding it would certainly require a deeper analysis of the corrections for the individual conformers in terms of contribution of different modes. Again, we estimate the relative error of the thermodynamic corrections to be 10%.

Relative perfluoroalkylalkane hairpin free energy Our estimate for the relative perfluoroalkylalkane hairpin free energy $\Delta G(T)$ at $T = 0$ K, $T = 100$ K and $T = 300$ K is shown in figure 4.27. The temperature $T = 100\ K$ is close to the estimated temperature in the experiments on alkanes by Lüttschwager [32]. Table 4.10 provides relative hairpin energy values for ΔE, ΔG and ΔH at $T = 100$ K. Adding ΔE from LCC and $\Delta\delta G(100\text{ K})$ from TPSS-D3, gives the chain length of the last stable investigated linear perfluoroalkylalkane hairpin with even n as $n_{\mathrm{c},\Delta G(100\text{ K})} = 18$. The shortest chain length for a hairpin conformer that is more stable



(a) Temperature and chain length dependence of the enthalpy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies.

(b) Temperature and chain length dependence of the Gibbs free energy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies and the RRHO entropy correction.

Figure 4.26: Thermodynamic corrections to the relative hairpin energy (ΔE) for perfluoroalkylalkanes.

than the linear conformer at $T = 100$ K is $n_{\mathrm{hp},G(100\ \mathrm{K})} = 20$. The same prediction is made for $T = 0$ K, as the hairpins are only about 1 kJ/mol more stable in comparsion to $T = 100$ K. It should be mentioned, that the calculated relative hairpin free energy at $T = 100$ for $n = 20$ is very small in magnitude, with $\Delta G(100\ \mathrm{K}) = -0.8$ kJ/mol. This is also the case for $n = 22$, with $\Delta G(100\ \mathrm{K}) = -1.1$ kJ/mol. In view of our estimated errorbars for $\Delta G(100\ \mathrm{K})$ at $n = 20$ and $n = 22$ of about ± 2 kJ/mol, our results allow for the possibility, that perfluoroalkylalkane hairpin formation may start at even higher n. On the other hand the prediction, that for $n = 18$ the linear conformer is preferred at $T = 100$ K is more certain. Hence, our results clearly show that perfluoroalkylalkane hairpin formation starts at longer chainlenghts in comparison to n-alkanes. At $T > 100$ K, the critical chain length shifts to larger values, as at $T = 300$ K none of the investigates perfluoroalkane hairpins are stable.

For comparison, we also show the results for $\Delta G(100\ \mathrm{K})$ obtained by using TPSS-D3, given by the red triangles in figure 4.27. These results lead to the prediction of shorter critical chain lengths $n_{\mathrm{c},G(100\ \mathrm{K})} = 14$ and shorter hairpin critical chain lengths $n_{\mathrm{hp},G(100\ \mathrm{K})} = 16$. These changes arise due to the discussed overestimation of the perfluoroalkylalkane hairpin stabilization by about 4 kJ/mol.

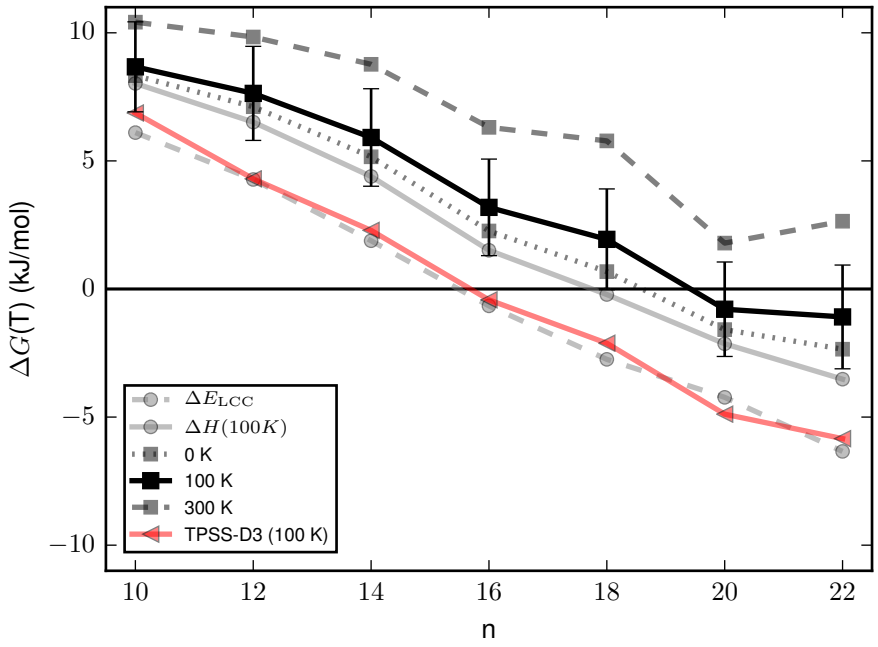
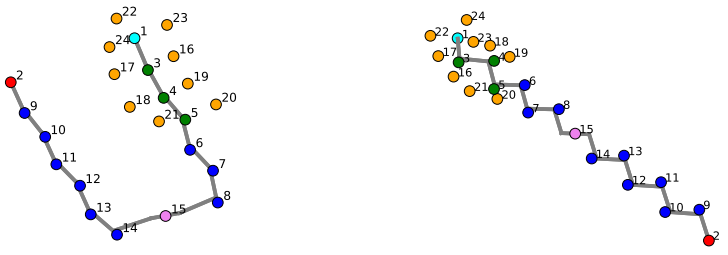


Figure 4.27: Chain length dependence of the the relative Gibbs free energy for perfluoroalkylalkane hairpins at 100 K. Additionally we show also the respective enthalpy and results for 0 K and 300 K.

n \ ΔE	(MoI)-LMP2-F12	LCC	ΔH(100 K)	ΔG(100 K)
10	4.22	6.10	8.02	8.67
12	2.05	4.27	6.51	7.64
14	−0.65	1.88	4.39	5.91
16	−3.48	−0.66	1.51	3.19
18	−5.78	−2.75	−0.22	1.94
20	−7.45	−4.23	−2.14	−0.79
22	−9.76	−6.34	−3.52	−1.09

Table 4.10: Relative perfluoroalkylalkane hairpin energies (kJ/mol) obtained by (MoI)-LMP2-F12/aVTZ, LCC, and by adding TPSS-D3/def2-TZVP thermodynamic corrections to LCC ΔE giving ΔH(100 K) and ΔG(100 K).

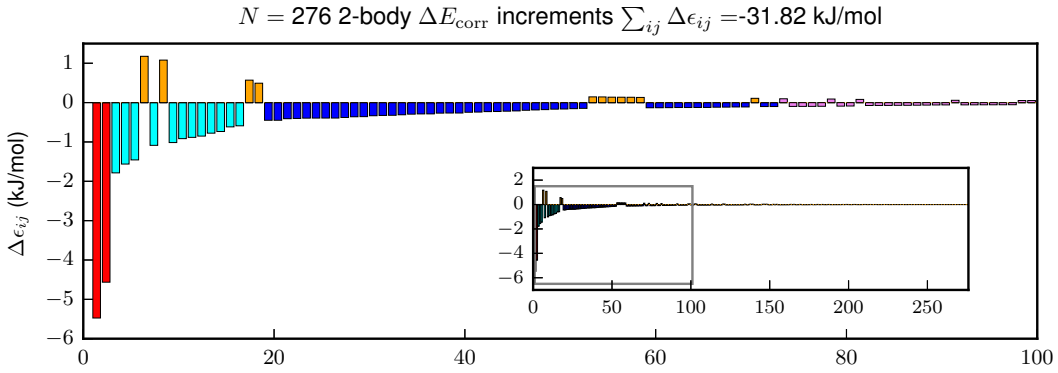
Relative perfluoroalkylalkane hairpin electronic energy: method of increments Here, we analyse the MoI-LMP2-F12/aVTZ results for the perfluoroalkylalkanes, of which the 2-body only results for $n \geq 18$ were used to approximate the LMP2 part of the LCC energy. Figure 4.28 shows the one-bodies of $C_4F_9C_{12}H_{25}$ ($n = 16$) represented by circles, which are superimposed on the carbon backbone. The internal numbering of the 1-bodies is the same for both conformers. With respect to the $n = 16$ alkane, the total number of 1-bodies for the $n = 16$ perfluoroalkylalkane is increased by 9, *i.e.* 15 vs. 24 1-bodies, due to the additional fluorine lone-pair 1-bodies (F_{LP}). The n-alkane with the same number of 1-bodies has a chain length of $n = 25$. The largest perfluoroalkylalkane system investigated in this work has $n = 22$. In terms of numbers of increments this is equivalent to the n-alkane with $n = 36$. As for the alkanes, we find that the 2-body increments give a good representation of the intra-molecular correlation energy at the LMP2 level of theory. We turn again to the example of the perfluoroalkylalkane with $n = 16$ and the magnitude of individual 2-body increments for the relative hairpin energy in figure 4.29. Let us begin with increments that can be attributed to hairpin stabilizing intramolecular dispersion interactions. Out of 276 2-body increments, two increments located in the fold of the hairpin contribute 32% of the total 2-body only relative hairpin correlation energy of -31.82 kJ/mol, see figure 4.29b. Twelve increments, shown figure 4.29c, fall into the range of -2 to -0.5 kJ/mol and contribute 39% to the overall 2-body sum. In the case of the respective alkane, similar sized increments have a contribution of 63%. The 12 increments describe 7 interactions located within the bend of the hairpin and 5 interactions involving the two F_{LP} bodies 17 and 18 to close CH_2 increments on the other arm. A similar sized part (36%) of the 2-body sum can be attributed to 47 increments with energies between -0.5 to -0.1 kJ/mol. As shown in figure 4.29d, these increments include interactions between F_{LP}-CH_2 as well as CF_2-CH_2 interactions. Destabilizing interactions, shown in figure 4.29e, amount to -13% of the 2-body sum, where the largest destabilizing increments (1 to 0.5 kJ/mol) are located in the fold. Small increments in the range -0.1 to -0.05 kJ/mol give a contribution of 5%, see figure 4.29f. The number of very small increments, shown in figure 4.29f, is 181. In sum, these increments give 2% of the total 2-body sum. They include increments between bodies located on the same arm and increments between bodies that are far away from each other. The analysis of the 2-bodies shows that two thirds of the 2-body increments could be neglected, without a significant loss of accuracy.



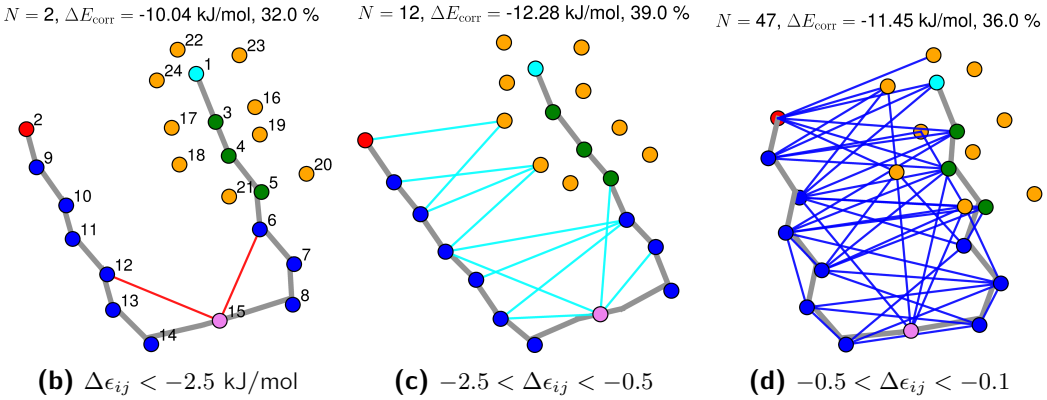
(a) $C_4F_9C_{12}H_{25}$ hairpin conformer backbone (gray) superimposed by coloured circles representing the charge centres of the one-body increments.

(b) Like (a) but for the linear conformer of $C_4F_9C_{12}H_{25}$.

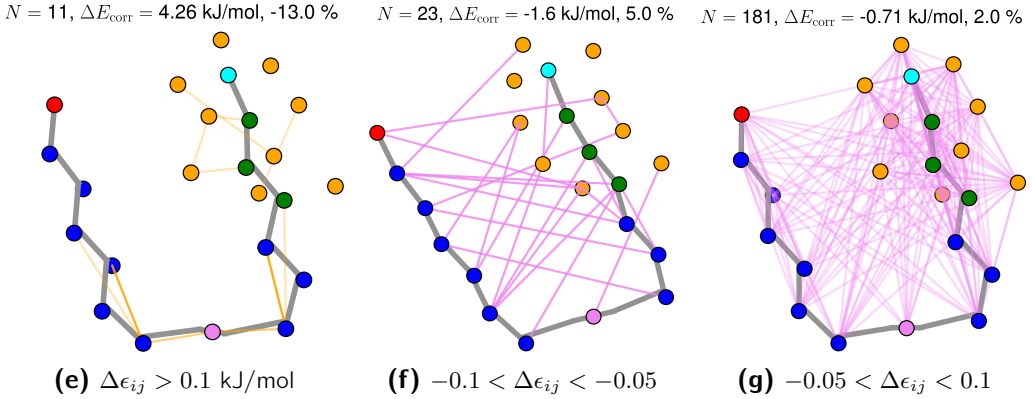
Figure 4.28: (a,b): Schematic depiction of one-body increments for $C_4F_9C_{12}H_{25}$.



(a) The 276 2-body increments for LMP2-F12/aVTZ [iext=1] $\Delta E_{\rm corr}$ of $C_4F_9C_{12}H_{25}$ sorted by their absolute value. Increments in certain energy ranges are color coded and depicted as lines of the same color superimposed on the $C_4F_9C_{12}H_{25}$ hairpin structure in (b-f).



(b) $\Delta\epsilon_{ij} < -2.5$ kJ/mol **(c)** $-2.5 < \Delta\epsilon_{ij} < -0.5$ **(d)** $-0.5 < \Delta\epsilon_{ij} < -0.1$

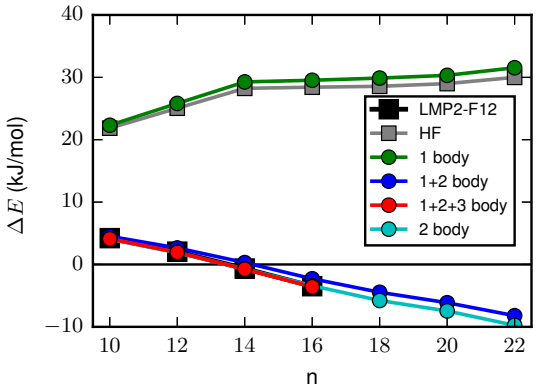


(e) $\Delta\epsilon_{ij} > 0.1$ kJ/mol **(f)** $-0.1 < \Delta\epsilon_{ij} < -0.05$ **(g)** $-0.05 < \Delta\epsilon_{ij} < 0.1$

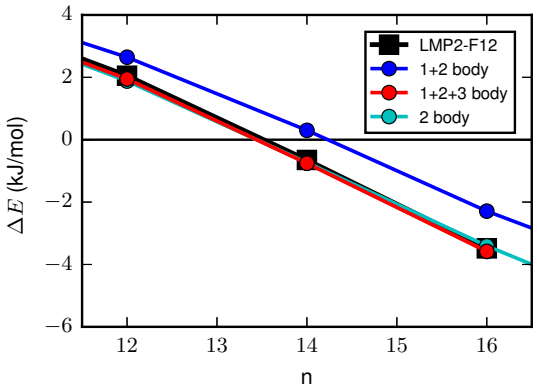
Figure 4.29: Importance of individual 2-body increments for the relative LMP2-F12 correlation energy of the $C_4F_9C_{12}H_{25}$ hairpin. (a) Overview of the increments ordered by magnitude. (b-f) The depiction of the location of 2-body increments within given energy ranges (kJ/mol) provides a map of the mainly dispersive interactions within the hairpin conformer. Also given are the number of increments (N), their sum (kJ/mol) and their fraction (%) relative to the sum over all 2-body increments.

Next, we briefly discuss the convergence of the chain length dependence of the relative hairpin energy of the perfluoroalkylalkanes with respect to the order of the incremental expansion of the LMP2-F12/AVTZ correlation energy. Up to $n = 16$ we calculated ΔE_{corr} with MoI-LMP2-F12/AVTZ at the 3-body level and up to $n = 22$ at the 2-body level. For $n \leq 16$ we obtained the errors due to the truncated incremental expansions with respect to the full LMP2-F12/AVTZ calculations.

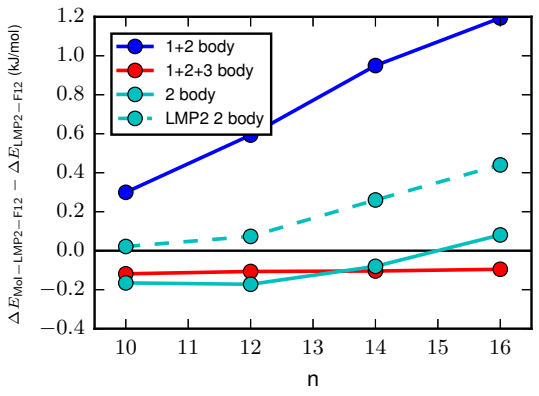
Truncation at the 1-body level results in a small hairpin destabilization with respect to the HF+[CABS singels] energy, see figure 4.30a. The sum over the 1-body correlation energies gives -2% ($n = 10$) to -4% ($n = 16$) of the normal LMP2-F12 correlation energy contribution to ΔE, see figure 4.30d. If the MoI is truncated at the 2-body level, 96% to 98% of the stabilizing LMP2-F12 correlation is described, see figure 4.30d. Hence, at the 2-body level, the LMP2-F12 part of ΔE is overestimated by 1.2 kJ/mol at $n = 16$, see figures 4.30b and c. A converged LMP2-F12 ΔE curve is obtained at the 3-body level, see figures 4.30a and b, where the absolute error in correlation energy at the 3-body level is -0.1 kJ/mol for $n \leq 16$, see figure 4.30c. As for the alkane folding, the 1-body sum and the 3-body sum are of similar magnitude but of opposite sign, see figure 4.30d. Due to this cancellation effect, the sum over the 2-body increments gives a good approximation to the correlation energy. The correlation energy contribution from the 2-body-only approach deviates by -0.2 kJ/mol ($n = 10$) to 0.1 kJ/mol ($n = 16$) from the actual LMP2-F12 value, see figure 4.30c. The good performance of the sum over the 2-body increments is partly related to the F12 part of the LMP2-F12 energy. The F12 contribution to ΔE_{corr} converges at the 2-body level, but its magnitude is overestimated by about 0.4 kJ/mol for $n = 16$, by only accounting for the sum over 2-body increments, *i.e.* neglect of the 1-body sum. The error in the 2-body-only LMP2 energy, given in figure 4.30c, is effectively counter balanced by the error in the 2-body-only F12 part. Using linear extrapolation of the errors, we estimate that the 2-body-only approach underestimates the hairpin stability by 0.6 kJ/mol at $n = 22$, while the full 2-body level results in an underestimation of 2.1 kJ/mol at $n = 22$. For $n = 22$ the number of increments on both conformers is 1332 at the 2-body level and 15612 at the 3-body level. We have therefore employed the sum over the 2-body increments approach for the MoI-LMP2-F12 part of LCC used for perfluoroalkylalkanes with $n > 16$.



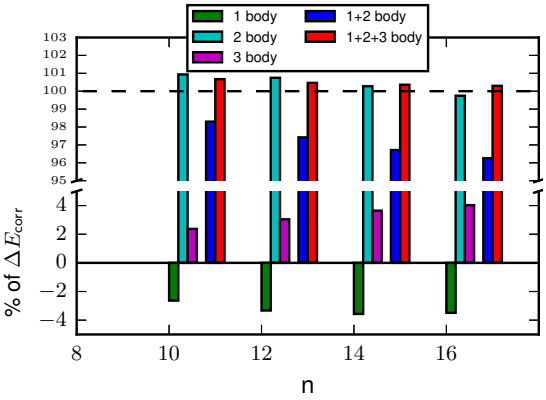
(a) Relative perfluoroalkylalkane hairpin energy at the HF+[CABS singles] and LMP2-F12 level of theory and using the Mol-LMP2-F12 up to the 3-body level. The notations 1+2 body etc. refer to sums over all the respective n-body increments which approximate the LMP2-F12 correlation energy contribution to ΔE.



(b) Zoom-in of the plot in (a).



(c) Absolute error in the relative hairpin LMP2-F12 energy obtained for different Mol approximations. Additionally, the error of the 2-body incremental sum for the expansion of the LMP2 relative hairpin energy is given.



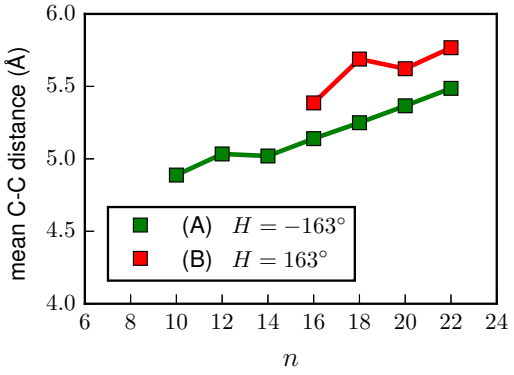
(d) Fraction of the relative hairpin LMP2-F12 correlation energy obtained at different levels of the Mol.

Figure 4.30: Convergence of the relative hairpin energy (ΔE) of perfluoroalkylalkanes for the Mol-LMP2-F12/aVTZ [iext=1] level of theory with increasing incremental orders.

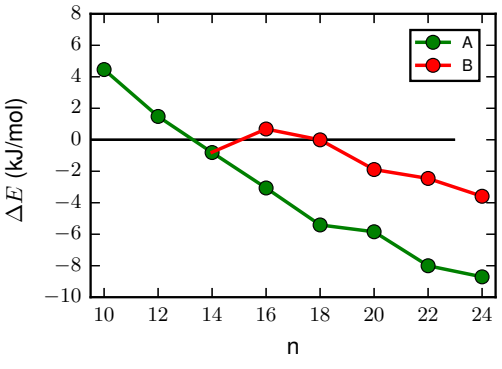
4.3.4 1,8-Diperfluoroalkyloctane folding

Here, we present the results on the folding of 1,8-diperfluoroalkyloctanes with the formula C_mF_{2m+1}-C_8H_{16}-C_mF_{2m+1}. First, we discuss the TPSS-D3 structures and relative hairpin energies for two combinations of differently handed perfluoralkylhelices. Second, LCC relative hairpin energies for the more stabilized diperfluoroalkyloctanes hairpins are discussed. Third, dispersion corrected DFT results and their deviations with respect to LCC are compared. Fourth, we present TPSS-D3/def2-TZVP thermodynamic corrections. Fifth, we combine LCC and the thermodynamic corrections to predict the critical chain length for hairpin formation. Sixth, we analyse the MoI-LMP2-F12 results in terms of a 2-body map and the convergence of ΔE at the n-body level.

1,8-Diperfluoroalkyloctanes hairpin structures The TPSS-D3/def2-TZVP structures of the 1, 8-diperfluoroalkyloctane hairpin conformers (C_mF_{2m+1}-C_8H_{16}-C_mF_{2m+1}) are shown in figure 4.32. The overall structures of the perfluoroalkylgroups are as already discussed for the perfluoroalkylalkane hairpins, *i.e.* helix formation begins at $m = 4$. We optimized structures with initial helical dihedral angles of $-162°$ (case A) and $+162°$ (case B). Hence, in both cases the handedness of the perfluoroalkyl helices is interchanged. It should be noted, that a third case, where one perfluoroalkylgroup has helical dihedral angles of $-162°$ and one has helical dihedral angles of $+162°$ is not considered in this investigation. Both structure optimizations lead to considerably different structures for $n \geq 16$. In this range, case A gave rise to less open hairpins, where the mean C-C distances are about 0.25 to 0.5 Å shorter compared to case B, see figure 4.31a. The hairpin conformers of case A are stabilized by 4-5 kJ/mol over the ones of case B, as can be seen in the relative hairpin energies at the TPSS-D3/def2-TZVP in figure 4.31b. The different interaction behaviour of perfluoroalkylhelices within hairpin conformers is very interesting. However, we are unsure if this is a genuine effect or merely an artefact of our structure generation process. Because we are mainly interested in the comparison of CH-CH, CH-CF and CF-CF interactions, we focus on the more stable hairpins of case A, *i.e.* with helix dihedrals of $-163°$. The mean interarm C-C distances of the hairpins (A) range from 4.9 Å ($n = 10$) to 5.5 Å ($n = 22$).



(a) Mean C-C distances of facing carbons on adjacent chains of the diperfluoroalkyloctane hairpin conformers. For an example with individual C-C distances see table 4.11.



(b) TPSS-D3 relative hairpin energy for diperfluoroalkylalkanes A and B.

Figure 4.31

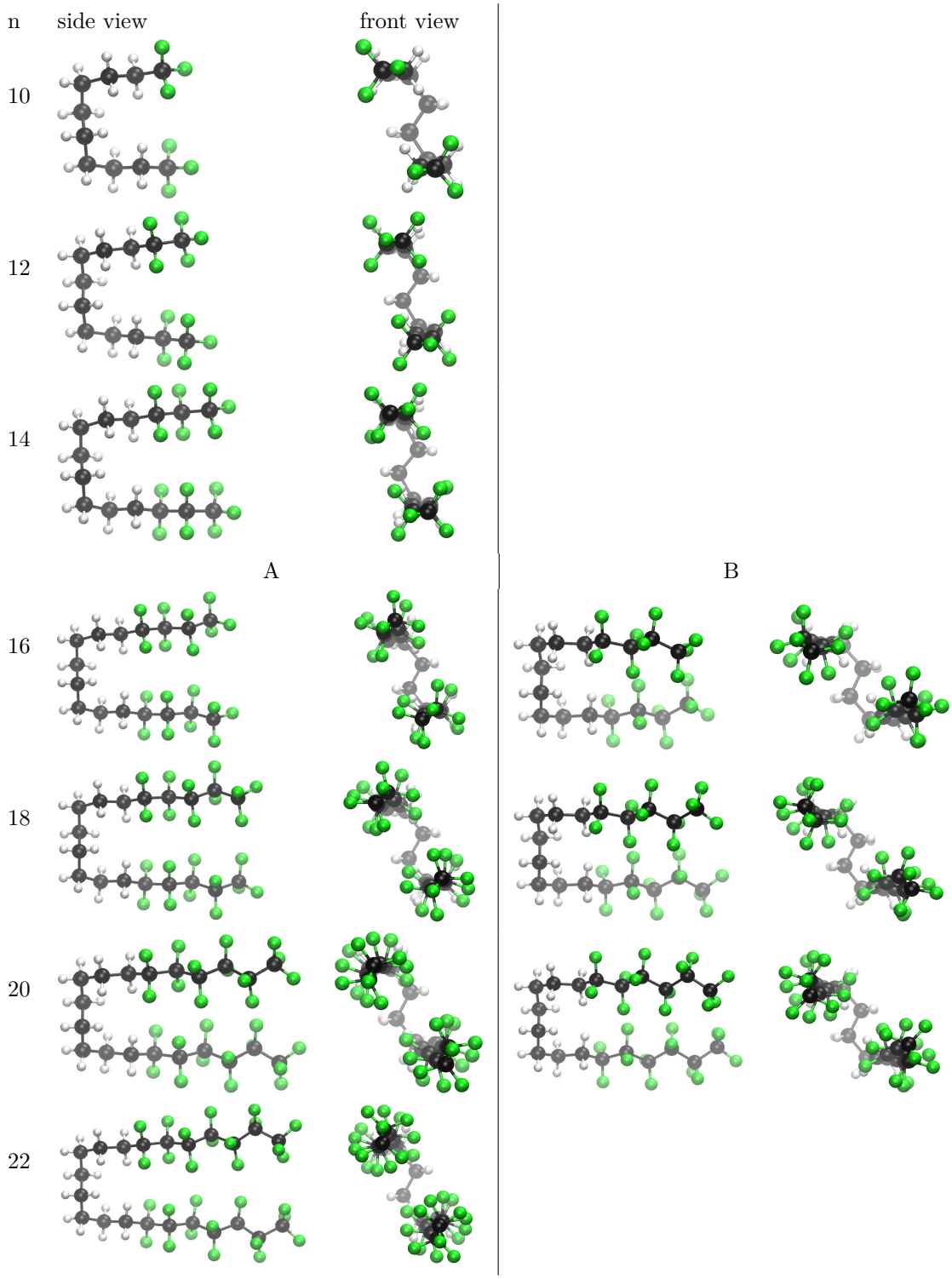
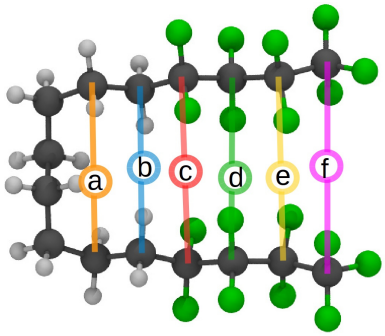


Figure 4.32: Diperfluoroalkyloctane (C_mF_{2m+1}-C_8H_{16}-C_mF_{2m+1}) hairpin TPSS-D3/def2-TZVP structures for even n from $n = 10$ to $n = 20$. The length of the perfluoroalkyl groups is $m = (n - 8)/2$.

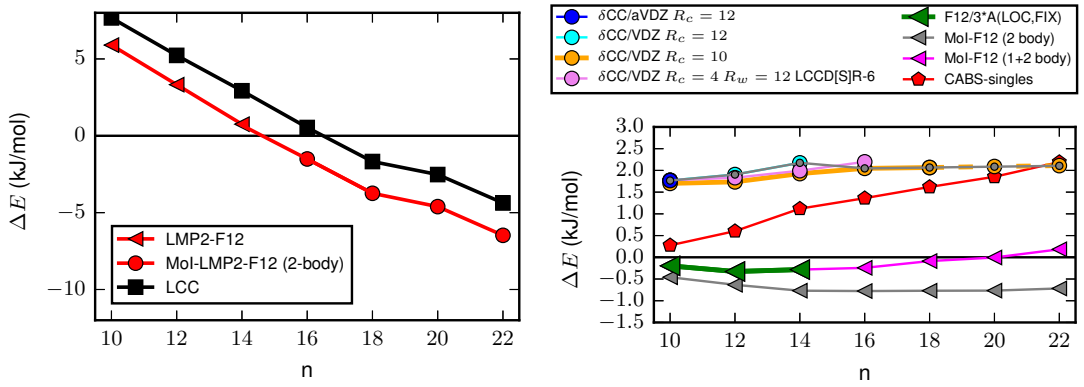


Method	a (Å)	b (Å)	c (Å)	d (Å)	e (Å)	f (Å)	∅ (Å)
TPSS-D3/def2-TZVP	4.95	4.55	5.23	5.08	5.23	5.79	5.14

Table 4.11: C-C distances between carbons on parallel chains of the $n = 16$ 1,8-diperfluoroalkyloctane hairpin, as indicated by the above figure. The mean interarm C-C distance is denoted by ∅.

The C-C distances between the arms are not only larger than for comparable n-alkane hairpins, but also increase with n. Individual C-C distances are given for $n = 16$ in table 4.11. The TPSS-D3 relative hairpin energy of case A shows a similar dependence on n as found for the perfluoroalkylalkanes and alkanes. TPSS-D3/def2-TZVP predicts $n_{\mathrm{hp,el}} = 14$ as the critical chain length for hairpin stability in terms of the electronic energy.

Relative 1,8-diperfluoroalkyloctane hairpin energy: LCC results The calculation of the LCC ΔE energy for the 1,8-diperfluoroalkyloctanes is more demanding compared to the alkanes and perfluoroalkylalkanes. It was therefore necessary, to introduce further approximations in the determination of ΔE up to chain lengths of $n = 22$. The LMP2-F12/aVTZ energy was obtained for extended domains and all pair domains for even chain lengths $n \leq 14$ using the standard local correlation method. For chain lengths $n \geq 16$, the LMP2-F12 energy was obtained using the MoI 2-body-only approximation. For a description of the convergence behaviour of the MoI results we refer to the end of this section. The resulting composite LMP2-F12 ΔE curve is shown in figure 4.33a. The CABS singles correction to the HF/aVTZ part of ΔE is considerably larger than the F12 part, see figure 4.33b. The CABS singles correction increases with n from 0.25 kJ/mol ($n = 10$) to 2.25 kJ/mol ($n = 22$). This indicates that the HF/aVTZ description of interacting perfluoroalkylgroups is not converged with respect to the complete basis set limit, which is also the case for the perfluoromethane dimer. The contribution of explicit correlation (F12/3*A(LOC,FIX)) to ΔE ranges between -0.3 kJ/mol ($n = 12$) to 0.2 kJ/mol ($n = 22$). It has to be noted, that the 2-body-only approach overestimates the F12 contribution considerably, while the F12 contribution at the 2-body level agrees with the values obtained without the MoI. The δCC/aVDZ correction, where pair domains with distances smaller than 6.4 Å were treated at the LCCSD(T0) level ($R_{\mathrm{close}} = 12$ a$_0$), was obtained only for $n = 10$. Up to $n = 14$ we obtained values from δCC/VDZ using the same distance criteria as for δCC/aVDZ. Reducing the distance criterion for strong pairs to 5.4 Å ($R_{\mathrm{close}} = 10$ a$_0$) and neglect of pairs with distances larger than 7.9 Å ($R_{\mathrm{vdist}} = 15$ a$_0$) allowed us to obtain δCC/VDZ up to $n = 18$. Up to $n = 16$ we also employed another pair approximation where LCCSD(T0) was used for pair domains with distances below 2.1 Å ($R_{\mathrm{close}} = 4$ a$_0$) and close



(a) Change of the relative hairpin energy of 1,8-diperfluoroalkyloctanes with the chain length n. Linear and hairpin structures were optimized at the TPSS-D3/def2-TZVP level of theory. LMP2-F12/AVTZ results with extended orbital domains are given by red triangles and by red circles as obtained by 2-body only Mol. LCC results are given by black squares.

(b) Change of components of the LCC energy relative to LMP2/AVTZ results with n. Explicit correlation (F12) decreases ΔE, whereas the coupled cluster correction δCC increases ΔE. The δCC values for $n = 20$ and $n = 22$ have been linearly extrapolated from the values for $n = 16$ and $n = 18$ of δCC/VDZ $R_c = 10$.

Figure 4.33: Energy of the 1,8-diperfluoroalkyloctane hairpin conformers relative to the linear conformers.

pairs, with distances up to 6.4 Å ($R_{\text{weak}} = 12$ a$_0$), were treated with LCCSD[S]R^{-6}, which is a full local CCD with perturbative singles and neglect of all diagrams that decay quicker than R^{-6} [200]. All methods yield similar results for δCC, which destabilizes hairpins by about 2 kJ/mol, without a strong chain length dependence. For the final LCC relative hairpin energy, we employ the corrections δCC/aVDZ at $n = 10$, δCC/VDZ ($R_{\text{close}} = 12$ a$_0$) for $12 \geq n \leq 14$, δCC/VDZ ($R_{\text{close}} = 10$ a$_0$) for $16 \geq n \leq 18$ and extrapolate the δCC/VDZ correction at $n = 20$ and $n = 22$. This is indicated by small grey circles in figure 4.33b. The corresponding LCC ΔE curve, see figure 4.33a, begins at 7.7 kJ/mol ($n = 10$) and decreases approximately linearly to -4.4 kJ/mol ($n = 22$). The mean stabilization that is gained by adding an extra pair of CF_2 groups ($\Delta n = 2$) is 2.3 kJ/mol between $n = 10$ and $n = 18$ and decreases to 1.3 kJ/mol between $n = 18$ and $n = 22$. Averaging over the range $n = 10$ to $n = 22$ gives a mean stabilization of 2 kJ/mol per $\Delta n = 2$. The resulting hairpin critical chain length at the LCC level of theory based on the electronic energy is $n_{\text{hp,el}} = 18$. Again, we estimate the error of the LCC relative hairpin energy for the 1,8-diperfluoroalkyloctanes to be ± 1.5 kJ/mol. The main part of this error is attributed to the approximations in the δCC correction, and to the uncertainty due to the TPSS-D3 structures.

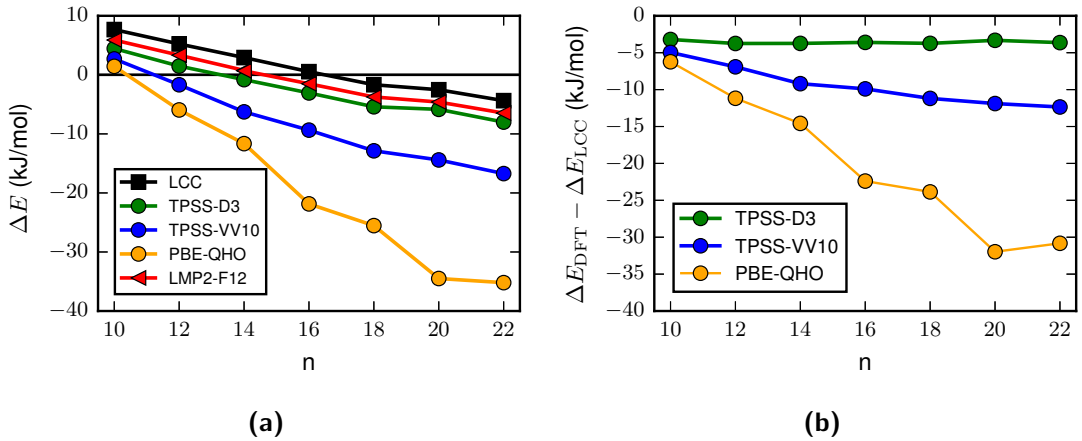


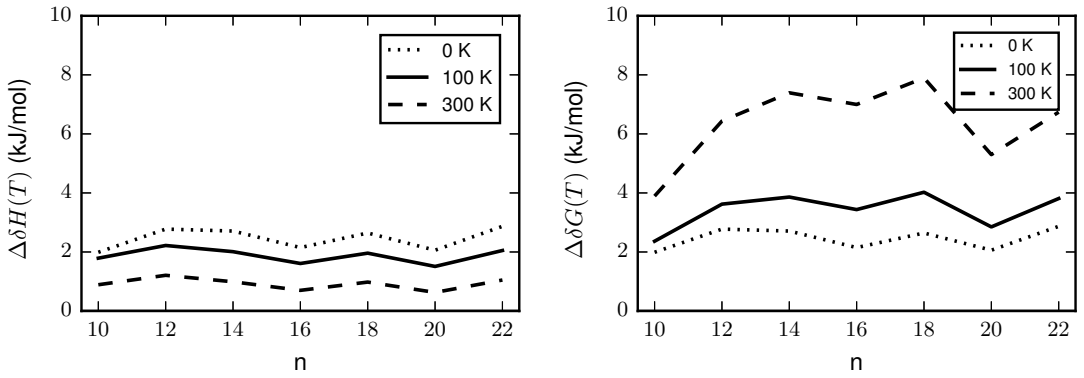
Figure 4.34: Comparison of relative diperfluoroalkyloctane hairpin energies (a) for the TPSS-D3 structures obtained by the wave-function-based LCC and LMP2-F12/aVTZ methods and the DFT results obtained with TPSS-D3/def2-TZVP, TPSS-VV10/def2-TZVP and PBE-QHO with a plane wave basis set. (b) Differences with respect to LCC.

Relative 1,8-diperfluoroalkyloctane hairpin energy: DFT results Comparing the dispersion corrected DFT relative hairpin energies of the 1,8-diperfluoroalkyloctanes with respect to the LCC results, gives a similar picture as in the case of the perfluoroalkylalkanes, see figure 4.34.

TPSS-D3/def2-TZVP overestimates the relative hairpin stability by 2 kJ/mol compared to LMP2-F12 and by 3.7 kJ/mol in comparison to LCC. The TPSS-D3 stabilization gain per added pair of CF_2 units agrees with the wave function based results. The critical chain length for hairpin formation obtained by TPSS-D3 is $n_{\mathrm{hp,el}} = 14$. This shows again, that the measure $n_{\mathrm{hp,el}}$ for the 1,8-diperfluoroalkyloctane hairpins is extremely sensitive on small deviations in ΔE.

The TPSS-VV10/def2-TZVP method overestimates the hairpin stability by 5 kJ/mol to 12 kJ/mol, relative to LCC. Other than D3, VV10 not only shifts ΔE, but also overestimates the stability gain per pair of CF_2 units, yielding 3.2 kJ/mol. The critical chain length for hairpin formation obtained by TPSS-V10 is $n_{\mathrm{hp,el}} = 12$.

PBE-QHO shows an even stronger hairpin stabilizing tendency than TPSS-VV10. The method over-binds 1,8-diperfluoroalkyloctane hairpins by 6 ($n = 10$) to 31 kJ/mol ($n = 22$), which results in an average stabilization gain of 6.1 kJ/mol per pair of CF_2 units. However, the ΔE curve drops more drastically between $n = 14$ to $n = 16$, while almost no change occurs from $n = 20$ to $n = 22$. The critical chain length for hairpin formation obtained by PBE-QHO is $n_{\mathrm{hp,el}} = 12$. But it should also be noted, that ΔE at $n = 10$ is just 1.4 kJ/mol.



(a) Temperature and chain length dependence of the enthalpy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies.

(b) Temperature and chain length dependence of the Gibbs free energy correction to ΔE from unscaled TPSS-D3/def2-TZVP vibrational frequencies and the RRHO entropy correction.

Figure 4.35: Thermodynamic corrections to the relative hairpin energy (ΔE) for 1,8-diperfluoroalkyloctanes.

Relative 1,8-diperfluoroalkyloctane hairpin energy: Thermodynamic corrections The thermodynamic corrections $\Delta\delta H(T)$ and $\Delta\delta G(T)$ obtained at the TPSS-D3 level of theory are similar to those obtained for the perfluoroalkylalkanes. The enthalpic correction $\Delta\delta H(T)$ is small and shows an almost constant behaviour with respect to n and T, see figure 4.36. At $T = 100$ K, the correction $\Delta\delta H(T)$ for 1,8-diperfluoroalkyloctane hairpins amounts to 2 kJ/mol. The Gibbs free energy correction naturally shows a stronger dependence on T due to the added entropic correction. At $T = 100$ K, $\Delta\delta G(T)$ destabilizes hairpins by 2.3 kJ/mol ($n = 10$) to 4.0 kJ/mol ($n = 18$). At $T = 300$ K, entropic effects dominate $\Delta\delta G(300\text{ K})$. Here, the destabilization raises to 4 kJ/mol ($n = 10$) to 8 kJ/mol ($n = 18$). We estimate the relative error of the $\Delta\delta G(T)$ correction to be 10%.

Relative 1,8-diperfluoroalkyloctane hairpin free energy Addition of the thermodynamic corrections to ΔE increases the critical chain length for hairpin formation, see also table 4.12. Taking into account the enthalpy correction at $T = 0$ K raises $n_{\mathrm{hp,el}} = 18$ to $n_{\mathrm{hp},\Delta G(0\ \mathrm{K})} = 20$, see figure 4.36. The value of $\Delta G(0\ \mathrm{K}) = -0.5$ kJ/mol at $n = 20$ implies nearly equal energies for hairpin and linear conformers, as does the value at $n = 22$ ($\Delta G(0\ \mathrm{K}) = -1.5$ kJ/mol). At $T = 100$ K, *i.e.* at a temperature as estimated for experiments on alkanes [32, 33], the hairpin critical chain length is formally reached at $n_{\mathrm{hp},\Delta G(100\ \mathrm{K})} = 22$, with $\Delta G(100\,\mathrm{K}) = -0.6\,\mathrm{kJ/mol}$. The change in $\Delta G(100\ \mathrm{K})$ between $n = 16$ and $n = 22$ is 1.5 kJ/mol. Using this gradient, the extrapolated value at $n = 24$ is $\Delta G(100\ \mathrm{K}) = -2.1$ kJ/mol and thus within our estimated error bound of ± 2 kJ/mol. At $n = 20$ the relative hairpin free energy $\Delta G(100\ \mathrm{K})$ is 0.3 kJ/mol. Thus, the 1,8-diperfluoroalkyloctane with $n = 20$ can not be excluded with total certainty as a candidate for a stable hairpin at $T = 100$ K. Nevertheless, the prediction for the most likely range for $n_{\mathrm{hp},\Delta G(100\ \mathrm{K})}$ of 1,8-diperfluoroalkyloctane at $T = 100$ K is $n = 22$ to $n = 24$. According to our results, the 1,8-diperfluoroalkyloctanes with $n \leq 18$ are predicted to be linear. At a temperature of 300 K, no hairpin conformer up to $n = 22$ is lower in free energy than the respective linear conformers. The introduction of perfluoroalkyl-perfluoroalkyl interactions results in less stabilized hairpins in comparison to the n-alkane case. Therefore 1,8-diperfluoroalkyloctane hairpin formation begins at longer chainlenghts than for n-alkanes. The predicted ranges for hairpin critical chainlenghts of 1,8-diperfluoroalkyloctanes and perfluoroalkylalkanes are similar.

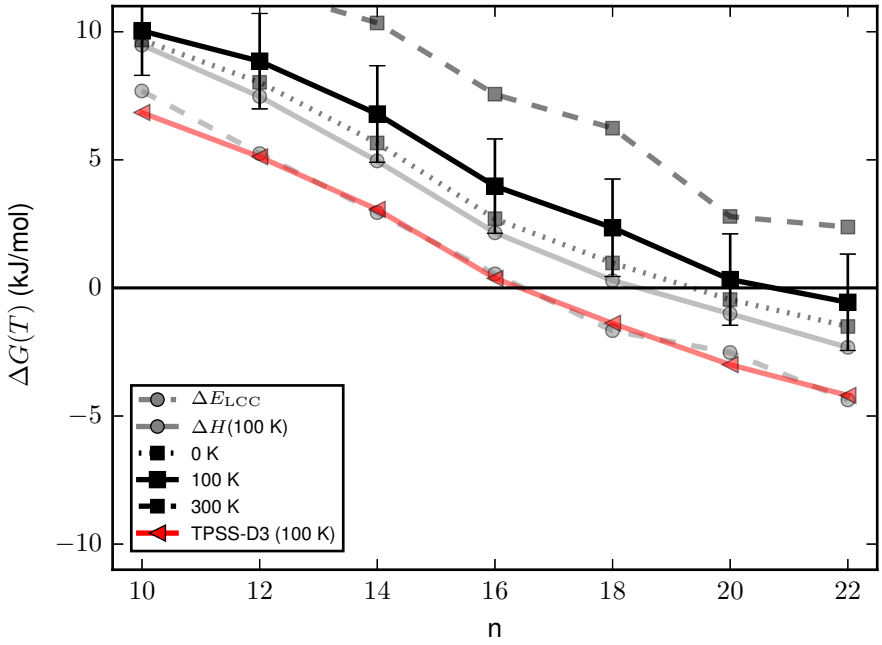
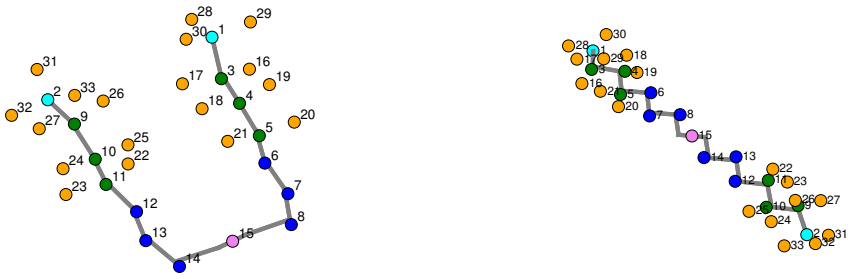


Figure 4.36: Chain length dependence of the the relative Gibbs free energy for 1,8-diperfluoroalkyloctane hairpins at 100 K. Additionally we also show the respective enthalpy and results for 0 K and 300 K.

n \ ΔE	(MoI)-LMP2-F12	LCC	$\Delta H(100\ \mathrm{K})$	$\Delta G(100\ \mathrm{K})$
10	5.90	7.67	9.46	10.03
12	3.23	5.23	7.45	8.85
14	0.75	2.93	4.94	6.79
16	−1.51	0.54	2.15	3.97
18	−3.74	−1.68	0.28	2.35
20	−4.61	−2.52	−1.01	0.33
22	−6.48	−4.37	−2.32	−0.56

Table 4.12: Relative 1,8-diperfluoroalkyloctane hairpin energies (kJ/mol) obtained by (MoI)-LMP2-F12/aVTZ, LCC, and by adding TPSS-D3/def2-TZVP thermodynamic corrections to LCC ΔE giving $\Delta H(100\ \mathrm{K})$ and $\Delta G(100\ \mathrm{K})$.

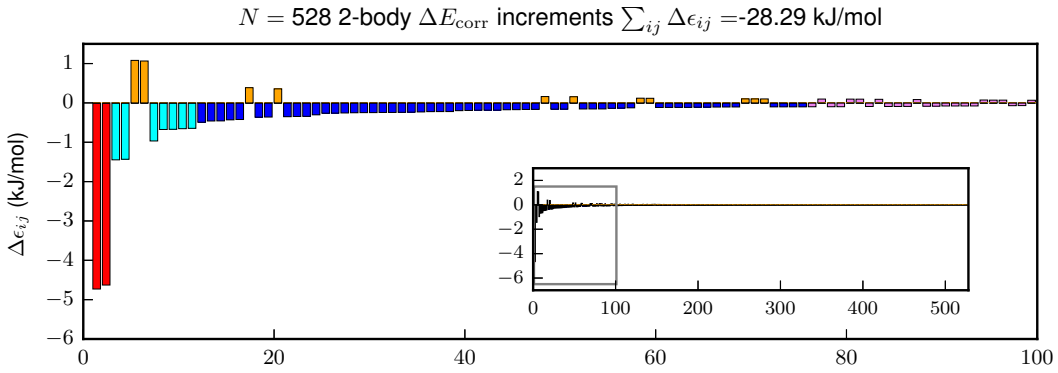
Relative 1,8-diperfluoroalkyloctane hairpin electronic energy: method of increments As in the case of the perfluoroalkyalkanes, we rely on the MoI to calculate the LMP2-F12 energy for 1,8-perfluoroalkyloctanes with $n > 14$ without the neglect of any pairs. The 33 one-bodies of 1,8-diperfluorobutyloctane ($n = 16$) are given in figure 4.37 for both conformers. The higher number of one-body increments, compared to the respective alkane (15) and perfluoroalkylalkane (24), results in a even larger number of two-body (528) and three-body (5456) increments, that in principle have to be evaluated for each conformer. In terms of the number of increments, the 1,8-diperfluoroalkyloctane with $n = 22$ is comparable to the n-alkane with $n = 52$. Due to limited resources, the full MoI-LMP2-F12/aVTZ [`IEXT=1`] 3-body level is only explored for 1,8-diperfluoromethyloctane ($n = 10$). Two-body level MoI calculations are performed up to $n = 22$ and are compared to the normal LMP2-F12 results for $n \leq 14$. Prior to the accuracy check, we investigate the spatial distribution of the pair interactions that stabilize the 1,8-diperfluorobutyloctane hairpin. Figure 4.38a shows the 528 2-body increments that sum up to the 2-body only LMP2-F12 part of the relative hairpin correlation energy. About 400 of the 2-body increments do not contribute to ΔE at all. Most of these increments, involve pairs of 1-bodies that are distant and/or are located on the same arm, as depicted in figure 4.38g. Fluorine lone pair 1-bodies on the outside of the hairpin do not contribute to the correlation energy in a significant manner. The 9 largest 2-body increments are all associated with the octane unit, see figures 4.38b, c and e. This pattern is similar to the previous alkane and perfluoroalkylalkane examples. The stabilizing correlation energy between pairs of fluorine lone pair 1-bodies is larger than between σ_{CF_2} pairs, on different arms respectively. The majority of stabilizing (mostly dispersion) pair interactions is found in the range $-0.5 < \Delta\epsilon_{ij} < -0.1$ kJ/mol (figure 4.38d). The total 2-body only sum results in a stabilizing relative hairpin correlation energy of -28.3 kJ/mol. Therefore, we expect the stabilizing dispersion interactions to be lower in comparison to the n-alkane (-38.7 kJ/mol) and also the perfluoroalkylalkane (-31.8 kJ/mol) with $n = 16$. The number of important 2-body increments for the conformational correlation energy change for 1,8-diperfluorobutyloctane ($n = 16$) is 150, which is just about three times more than for the respective n-alkane, with 57 2-body increments. Hence, there is a huge optimization potential with respect to the brute force MoI approach used herein.



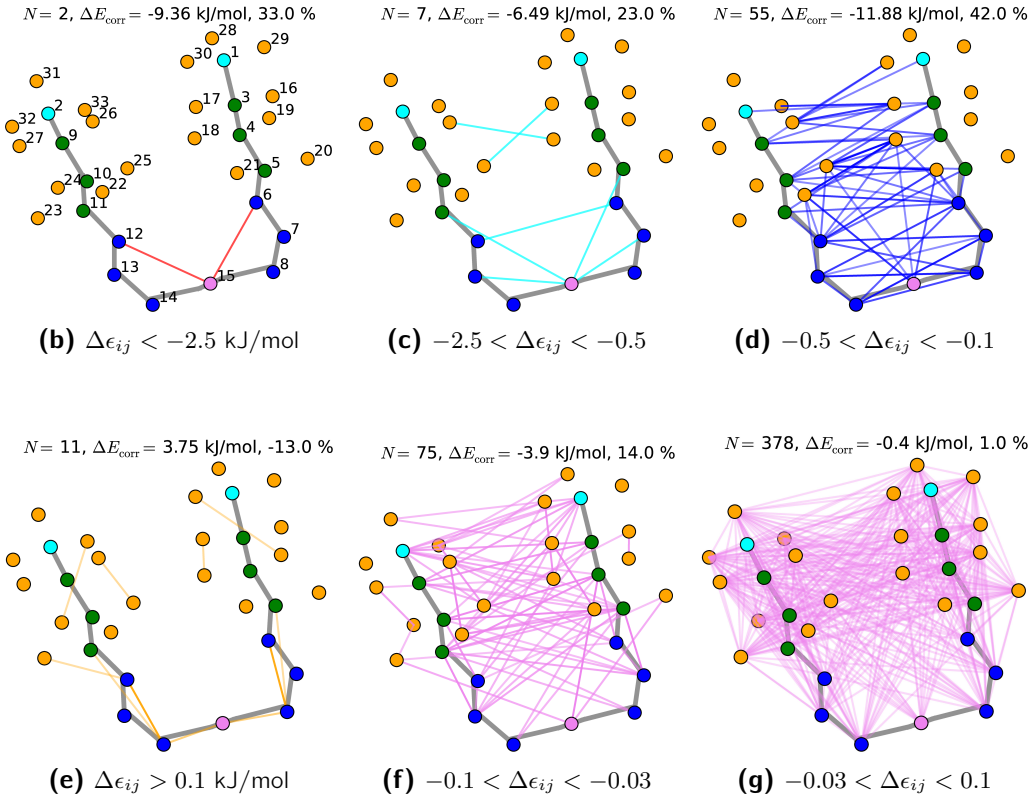
(a) 1,8-diperfluorobutyloctane hairpin conformer backbone (gray) superimposed by coloured circles representing the charge centres of the one-bodies.

(b) Like (a) but for the linear conformer of 1,8-diperfluorobutyloctane.

Figure 4.37: (a,b): Schematic depiction of one-bodies for 1,8-diperfluorobutyloctane.



(a) The 528 2-body increments for LMP2-F12/aVTZ [IEXT=1] ΔE_{corr} of 1,8-diperfluorobutyloctane sorted by their absolute value. Increments in certain energy ranges are color coded and depicted as lines of the same color superimposed on the 1,8-diperfluorobutyloctane hairpin structure in (b-f).

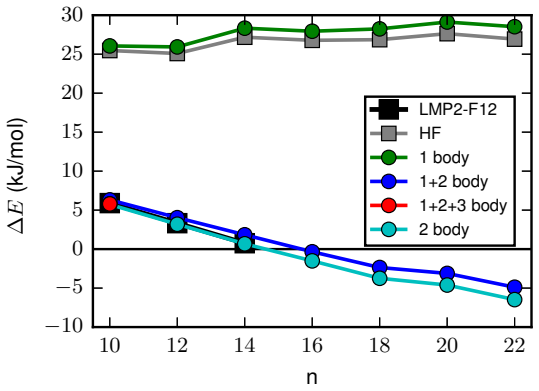


(b) $\Delta\epsilon_{ij} < -2.5$ kJ/mol **(c)** $-2.5 < \Delta\epsilon_{ij} < -0.5$ **(d)** $-0.5 < \Delta\epsilon_{ij} < -0.1$

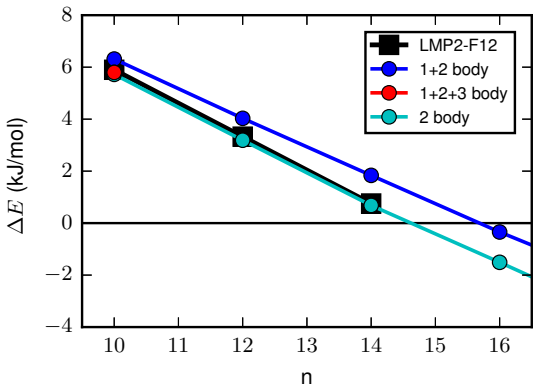
(e) $\Delta\epsilon_{ij} > 0.1$ kJ/mol **(f)** $-0.1 < \Delta\epsilon_{ij} < -0.03$ **(g)** $-0.03 < \Delta\epsilon_{ij} < 0.1$

Figure 4.38: Importance of individual 2-body increments for the relative LMP2-F12 correlation energy of the 1,8-diperfluorobutyloctane hairpin. (a) Overview of the increments ordered by magnitude. (b-f) The depiction of the location of 2-body increments within given energy ranges (kJ/mol) provides a map of the mainly dispersive interactions within the hairpin conformer. Also given are the number of increments (N), their sum (kJ/mol) and their fraction (%) relative to the sum over all 2-body increments.

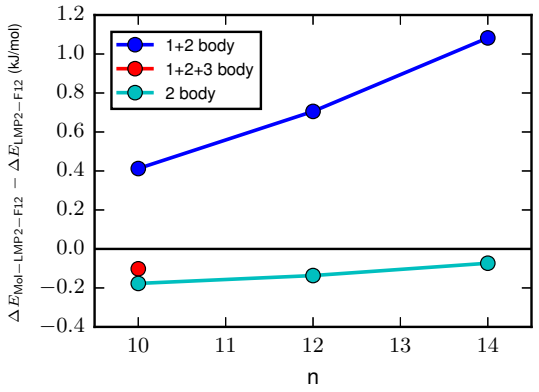
Next, we compare the chain length dependence of MoI-LMP2-F12 relative hairpin energy and determine the accuracy of the approximation for $n \leq 14$. Figure 4.39 provides an overview of ΔE obtained with the MoI. At the HF+[CABS singles] level 1,8-diperfluoroalkyloctane hairpins are about 25 kJ/mol less stable than linear conformers. This confirms again, that dispersion interactions are the main driving force of hairpin-stability. The correlation contribution to ΔE from 1-body increments is weakly destabilizing and ranges from 0.6 kJ/mol to 1.7 kJ/mol. At the 2-body level, the main part of the LMP2-F12 correlation energy contribution to ΔE is recovered in the MoI approximation. The 2-body LMP2-F12 relative hairpin correlation energy is stabilizing and ranges from -19 kJ/mol to -31 kJ/mol. This amounts to 96% of the LMP2-F12 relative hairpin correlation energy for $n = 14$, see figure 4.39d. At the 3-body level, the difference to the non-MoI result for $n = 10$ is -0.1 kJ/mol, while it is 0.4 kJ/mol at the 2-body level. Fortuitous cancellation of 1-body destabilizing and 3-body stabilizing correlation results in a good agreement of the 2-body only correlation energy approximation. The difference of the 2-body-only results to the non-MoI results for $n \leq 14$ ranges from -0.2 kJ/mol to -0.1 kJ/mol, see figure 4.39c. Hairpin stabilization due to the F12 part is overestimated at the 2-body level, as shown in figure 4.33b. Hence, the 2-body only correlation energy is closer to the converged full 3-body energy than the 2-body energy without F12. For 1,8-diperfluoroalkyloctane hairpins with $n > 14$, the 2-body-only approximation probably underestimates the hairpin stability given by the LMP2-F12 part slightly (≈ 0.2 kJ/mol). The errors due to the 2-body only MoI-treatment of the LMP2-F12 energy are included in our total error estimation (± 1.5 kJ/mol) of the LCC approach. Most importantly, they are certainly smaller than those introduced by other approximations that enter the LCC relative hairpin energy.



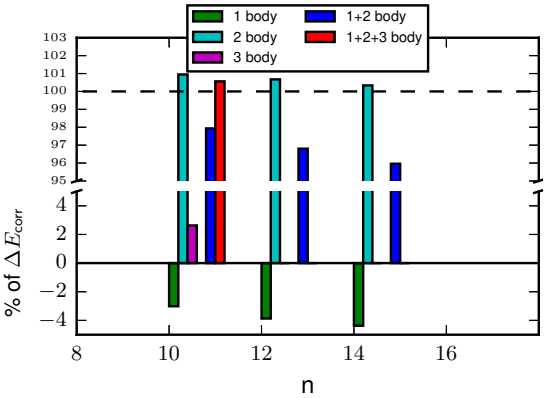
(a) Relative 1,8-diperfluoroalykloctane hairpin energy at the HF+[CABS singles] and LMP2-F12 level of theory and using the MoI-LMP2-F12 up to the 3-body level. The notations 1+2 body etc. refer to sums over all the respective n-body increments which approximate the LMP2-F12 correlation energy contribution to ΔE.



(b) Zoom-in of the plot in (a).



(c) Absolute error in the relative hairpin LMP2-F12 energy obtained for different MoI approximations.



(d) Fraction of the relative hairpin LMP2-F12 correlation energy obtained at different levels of the MoI.

Figure 4.39: Convergence of the relative hairpin energy of 1,8-diperfluoroalkyloctanes for the MoI-LMP2-F12/aVTZ [IEXT=1] level of theory with increasing incremental orders.

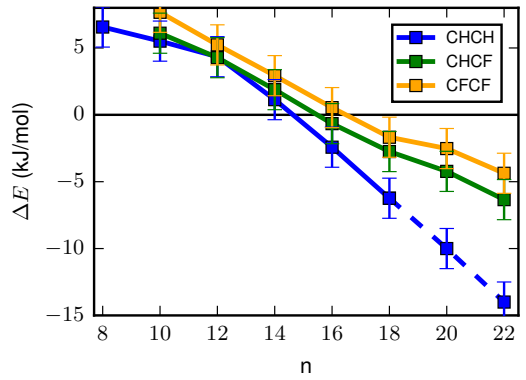
4.4 Influence of perfluoroalkyl groups on the stability of alkane hairpin conformers

In the preceding section the relative hairpin energies of alkane, perfluoroalkylalkane and 1,8-diperfluoroalkyloctane chain molecules were analysed in detail, with a focus on the specific systems and the local correlation methodology. Here a detailed comparison of the accurate ($\pm$1.5 kJ/mol) LCC results for the three systems is made. The question is, how do alkyl-alkyl, perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions compare in terms of hairpin stability? We begin to answer this question by comparing the relative hairpin energies obtained with the LCC method. As can be seen in figure 4.40, the relative hairpin energies for $n \leq 14$ are very similar, meaning they lie in a window of about 2 kJ/mol. Perfluoroalkyl-perfluoroalkyl interactions lead to generally higher relative hairpin energies, *i.e.* less stable hairpins, compared to the perfluoroalkyl-alkyl and alkyl-alkyl interactions. For longer chains ($n \geq 16$) the alkane hairpins are clearly more stabilized than perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpins. At a total chain length of $n = 22$ the relative hairpin energy of the n-alkane is about 8 kJ/mol lower than that of the respective perfluoroalkylalkane and 10 kJ/mol lower than that of the 1,8-diperfluoroalkyloctane. Throughout, the relative hairpin energies of 1,8-diperfluoroalkyloctanes are 1 to 2 kJ/mol higher than the energies of the respective perfluoroalkylalkanes. The stabilization per added pair of CH_2 units amounts to 3.5 kJ/mol, beginning from $n = 12$, while only about 2 kJ/mol is gained by adding two CF_2 units or one CH_2 and one CF_2 unit. In terms of n-octane hairpin stabilization perfluoroalkyl-perfluoroalkyl interactions and perfluoroalkyl-alkyl interactions are similar with respect to each other, but weaker in comparison to alkyl-alkyl interactions.

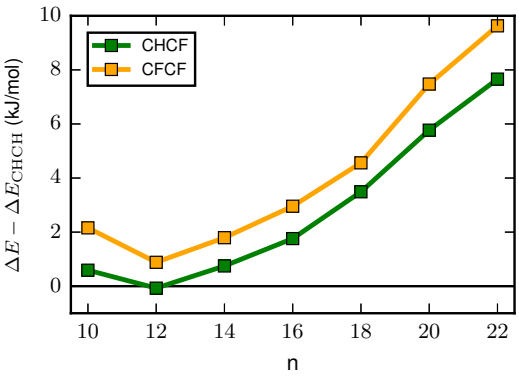
Why are the hairpins less stabilized by perfluoroalkyl-perfluoroalkyl and perfluoroalkyl-alkyl interactions compared to alkyl-alkyl interactions? To answer this question, we take a closer look at the HF and correlation energy components of the relative hairpin energies. The HF energy roughly represents electrostatic, induction and exchange repulsion interaction contributions, while the correlation energy mainly represents dispersion interactions.

Of course, these interactions depend on the mean C-C distance between the parallel groups within the hairpin. Beginning form $n = 12$, the mean C-C distances given in figure 4.40c show that alkyl groups can approach each other more closely than perfluoroalkyl groups, within the hairpin. The favourable alignment of alkyl groups has two origins. On one hand C-H bonds are shorter than C-F bonds and van der Waals radii of alkyl groups are smaller than for perfluoroalkylgroups. On the other hand, the all-trans structure of the alkyl groups facilitates their mutual alignment within the hairpin, while the helical structure of perfluoroalkyl groups (beginning from perfluorobutyl) hinders ideal alignment within the hairpin. This is clearly reflected in the n-dependence of the mean interarm C-C distance. For the alkane hairpins it is constant or even decreases [35]. For perfluoroalkylalkane hairpins and the 1,8-diperfluoroalkyloctane hairpins it increases with n.

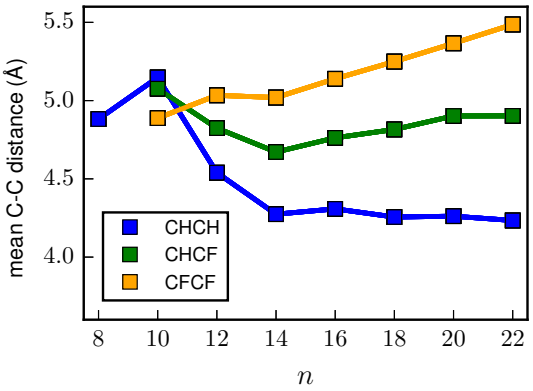
The HF part of ΔE for the longer alkane hairpins is more destabilizing compared to those of the respective perfluoroalkylalkanes and 1,8-diperfluoroalkyloctanes, see figure 4.40d. Possibly, this behaviour reflects increased exchange repulsion interactions in alkanes, due to the smaller distance between the alkyl groups. Another factor, to which we can only hint at this point, are differences in intra-molecular electrostatic and induction interactions. Figure 4.41 shows the molecular electrostatic potential for the $n = 18$ hairpins at the TPSS-D3/def2-TZVP level



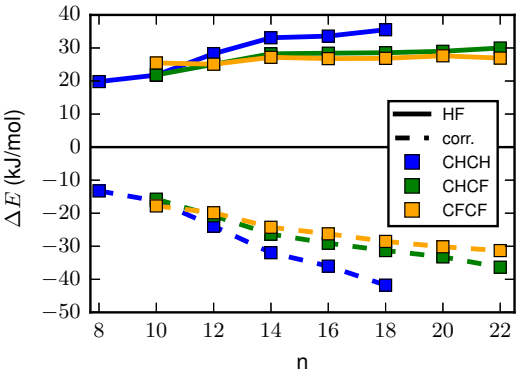
(a) Relative hairpin energy (ΔE) calculated using LCC.



(b) Difference in the LCC relative hairpin energies of perfluoroalkylalkynes and 1,8-diperfluoroalkyloctanes with respect to the n-alkanes.



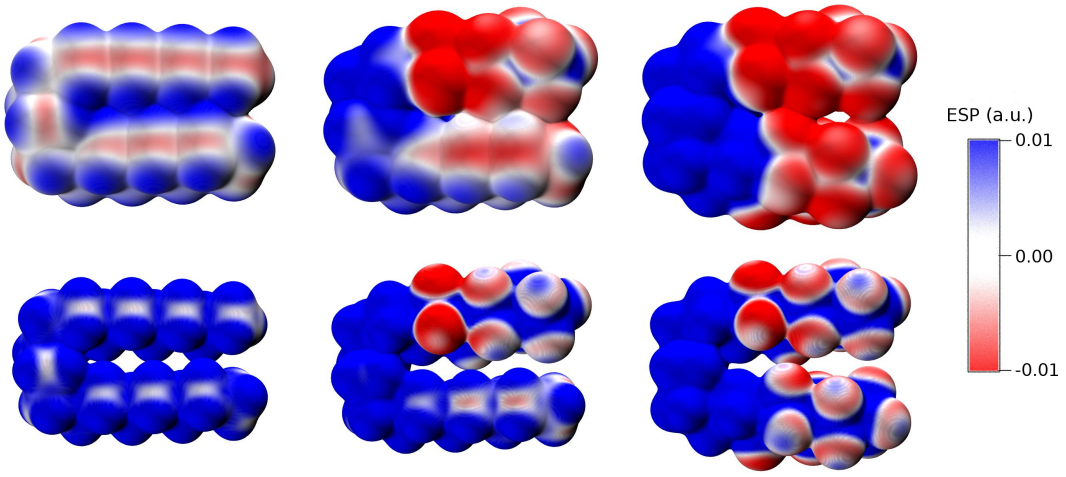
(c) Mean interarm C-C distances.



(d) HF+[CABS singles] and correlation energy parts of LCC relative hairpin energies.

Figure 4.40: Comparison of the chain length dependence of the LCC relative hairpin energies of n-alkanes (label: CHCH), perfluoroalkylalkanes (CHCF) and 1,8-diperfluoroalkanes (CFCF).

of theory. The values of the electrostatic potential are depicted at the 0.001 a.u. and 0.005 a.u. isosurfaces of the electron density. The introduction of perfluoroalkyl groups changes the electrostatic potential of the hairpin significantly. The opposite bond polarity of C-H and C-F bonds creates a molecular dipole moment at the boundary between alkyl and perfluoroalkyl groups. In the electrostatic potential map this is visible in form of a slightly more positive electrostatic potential at the bend of hairpins with perfluoroalkyl groups compared to the alkane hairpin. Likewise, the internal dipole is reflected by a slightly more negative potential at the fluorine atoms near the alkyl-perfluroalkyl boundary within the chain. Furthermore, σ-holes at each fluorine atom are visible, which hints to a complicated network of electrostatic intermolecular interactions. The electrostatic potential maps also nicely show the less optimal alignment of the perfluoroalkyl groups in comparison to the alkyl groups. It is therefore likely, that perfluoroalkylalkane and 1,8-diperfluoroalkylalkane hairpins benefit more from stabilizing electrostatic and induction interactions than alkane hairpins. The determination of these components is possible by state of the art intermolecular SAPT [181]. The weak n-dependence of the HF energy in figure 4.40d for $n \geq 14$ suggest, that the added exchange and electrostatic



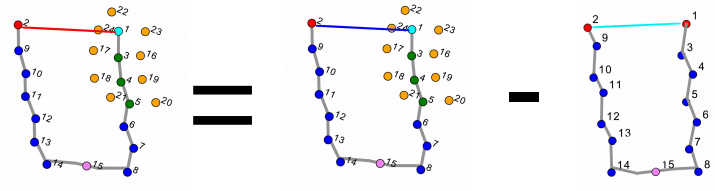
(a) n-Octadecane hairpin **(b)** Perfluoro-n-pentyl-n-tridecane hairpin **(c)** 1,8-Diperfluoro-n-pentyl-n-octane hairpin

Figure 4.41: Projection of the molecular electro static potential (ESP) in a.u. on the electron density isosurfaces at 0.001 e/a_0^3 (top row) and at 0.005 e/a_0^3 (bottom row), for the $n = 18$ n-alkane, perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpin conformers.

interactions between the arms of the hairpin are rather weak, and are most likely of greater importance within the bend of the hairpin. This is also supported by our previous HF-SAPT analyis of the methane and perfluoromethane dimers.

The main part of the intra-molecular interactions that leads to hairpin stabilization is the dispersion interaction. The correlation part of ΔE, mainly represents dispersion interactions and it increases in magnitude with chain elongation, see figure 4.40d. This increase in stabilizing correlation energy is distinctly stronger for the alkane hairpins than for the perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpins. This points to the following ordering of dispersion interactions in hairpins: alkyl-alkyl > perfluoroalkyl-alkyl ≥ perfluoroalkyl-perfluoroalkyl.

To visualize the hairpin stabilizing correlation energy differences within the differently substituted hairpins, we compare their two-body ΔE_{corr} increment maps for the $n = 16$ hairpins, see figures 4.42a–g. For this we caclulate the difference increments for the respective perfluoroalkylalkane and 1,8-diperfluoroalkyloctane with respect to the increments of the n-alkane as shown in figure 4.42a. Figures 4.42b-g show these individual difference increments for the respective perfluoroalkylalkane (b-f) and 1,8-diperfluoroalkyloctane (c-g). The largest two-body increments which arise due to correlation of one-bodies without an equivalent in the alkane, *i.e.* fluorine lone pairs, are depicted in figures 4.42f and g. In both cases, the 105 alkane 2-body increments outweigh their counterparts in the perfluoroalkane and 1,8-diperfluoroalkyloctane considerably. The largest differences can be attributed to the two-bodies between the perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl groups. In total the difference in correlation energy at the 2-body-only MoI-LMP2-F12 level with respect to the 105 two-body n-alkane increments with $n = 16$, amounts to 15.9 kJ/mol and 21.23 kJ/mol for the $n = 16$ perfluoroalkane and 1,8-diperfluoroalkyloctane respectively, which also correlates with the mean C-C distances in the hairpins. The additional two-body increments present in the partially fluorinated hairpins give rise to similar contributions to ΔE_{corr}, namely −9.0 kJ/mol



$$\Delta\epsilon_{12}^{\mathrm{CHCF}}(\Delta E_c - \Delta E_c(\mathrm{CHCH})) = \Delta\epsilon_{12}^{\mathrm{CHCF}}(\Delta E_c) - \Delta\epsilon_{ij}^{\mathrm{CHCH}}(\Delta E_c)$$
$$\Delta\epsilon_{12}^{\mathrm{CHCF}}(\Delta E_c) = \Delta\epsilon_{12}^{\mathrm{CHCF}}(\mathrm{HP}) - \Delta\epsilon_{12}^{\mathrm{CHCF}}(\mathrm{L})$$
$$\Delta\epsilon_{12}^{\mathrm{CHCH}}(\Delta E_c) = \Delta\epsilon_{12}^{\mathrm{CHCH}}(\mathrm{HP}) - \Delta\epsilon_{12}^{\mathrm{CHCH}}(\mathrm{L})$$

(a) Example for the definition of the backbone difference two-body increments with respect to n-alkane two-body increments for the relative hairpin correlation energy.

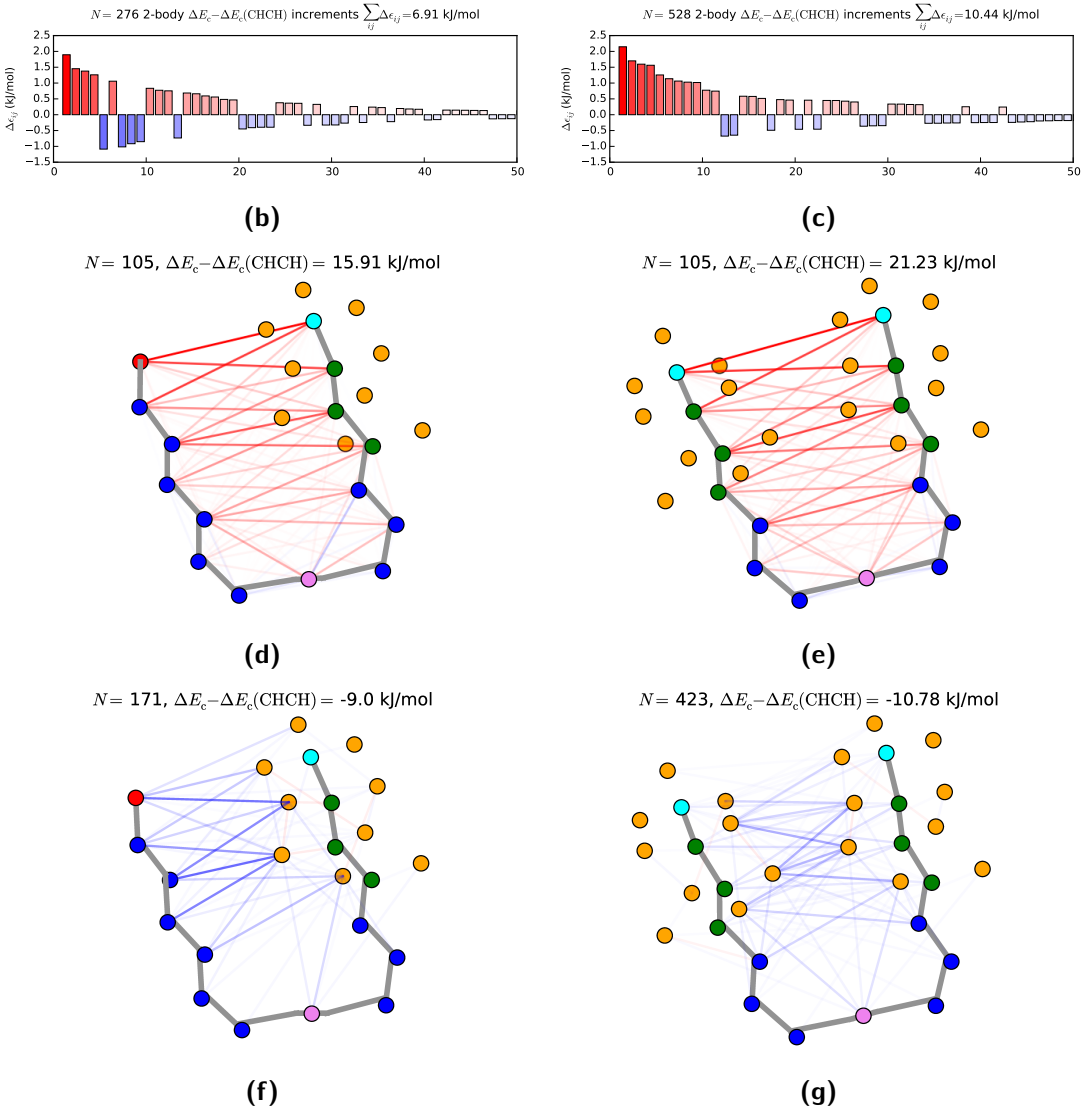
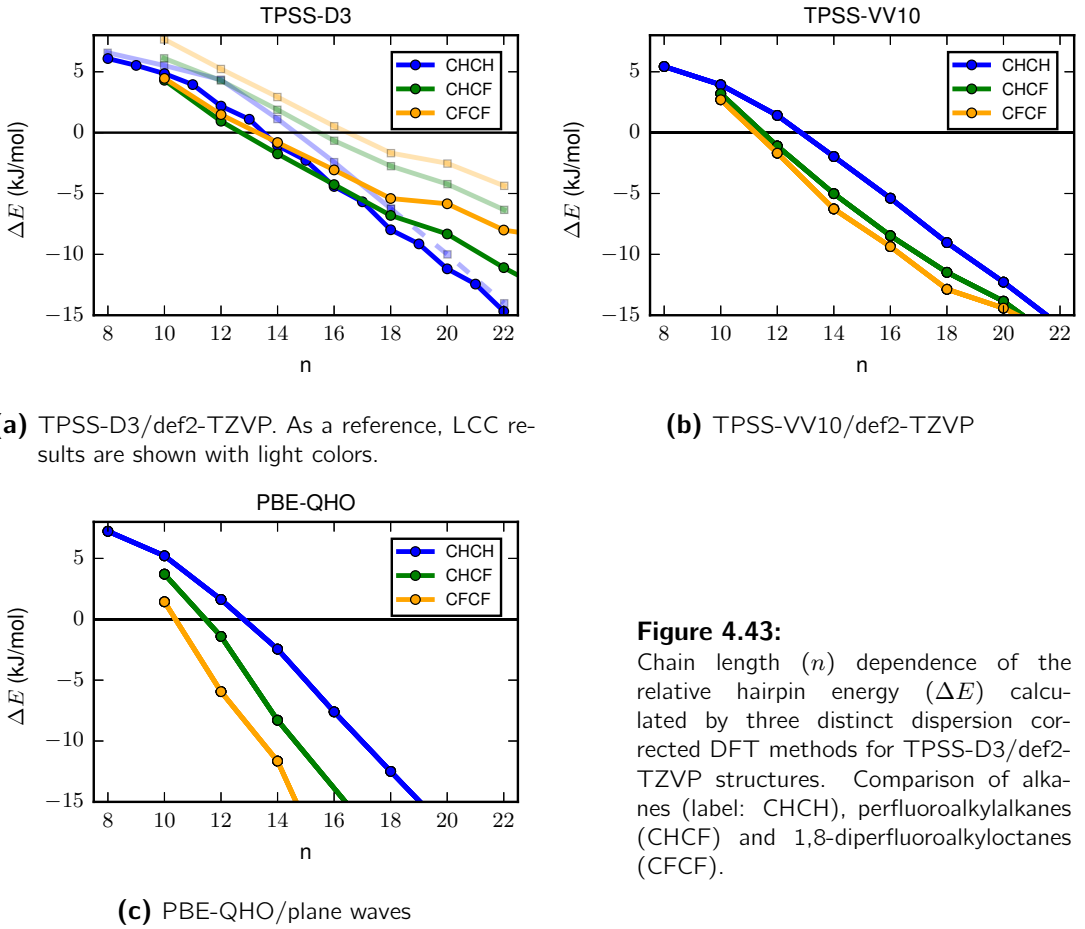


Figure 4.42: Difference 2-body increments of perfuorobutyl-n-dodecane (b,d,f) and 1,8-diperfluorobutyl-n-octane (c,e,g) with respect to the relative hairpin correlation energy of n-hexadecane ($n = 16$). The panels show: example (a), bar plots (b, c), maps for backbone increments (d,e) and increments involving fluorine lone pairs (f,g). The opacity of the lines reflect the relative magnitude of the increments.



(a) TPSS-D3/def2-TZVP. As a reference, LCC results are shown with light colors.

(b) TPSS-VV10/def2-TZVP

(c) PBE-QHO/plane waves

Figure 4.43:
Chain length (n) dependence of the relative hairpin energy (ΔE) calculated by three distinct dispersion corrected DFT methods for TPSS-D3/def2-TZVP structures. Comparison of alkanes (label: CHCH), perfluoroalkylalkanes (CHCF) and 1,8-diperfluoroalkyloctanes (CFCF).

and −10.8 kJ/mol. Since these cannot compensate the differences with respect to the alkane, the total relative hairpin correlation energy is less stabilizing in the case of perfluorobutyl-perfluorobutyl and butyl-perfluorobutyl interactions compared to butyl-butyl interactions.

The three tested dispersion corrected DFT methods differ strongly in their description of ΔE with respect to the partially fluorinated hairpins. As shown in figure 4.43a, the TPSS-D3 results are more similar to the LCC results, in contrast to the TPSS-VV10 and PBE-QHO results. In accordance to LCC, TPSS-D3 predicts similar hairpin stabilization for all three systems in the range $n = 10$ to $n = 16$. In this range, there is a tendency to slightly overstabilize partially fluorinated hairpins with respect to the n-alkane hairpins. The correct separation of perfluoroalkyl-perfluoroalkyl and alkyl-alkyl interactions is predicted in the range $n \geq 18$. Perfluoroalkyl-alkyl interactions in the TPSS-D3 model are however predicted to be less similar to perfluoroalkyl-perfluoroalkyl interactions than with LCC. For $n \geq 18$, TPSS-D3 gives the same overall order in hairpin stability for the three systems as LCC. TPSS-D3 also agrees with LCC on the gradients of the relative hairpin energies. TPSS-VV10 gives similar relative hairpin energies for perfluoroalkylalkanes and 1,8-diperfluoroalkyloctanes, which are both lower in comparison to the ones of the alkanes, see figure 4.43b. PBE-QHO extremely overestimates the stability of both perfluoroalkylalkane hairpins and 1,8-diperfluoroalkyloctane hairpins with respect to the n-alkane hairpins, see figure 4.43b. This shows, that interactions involving per-

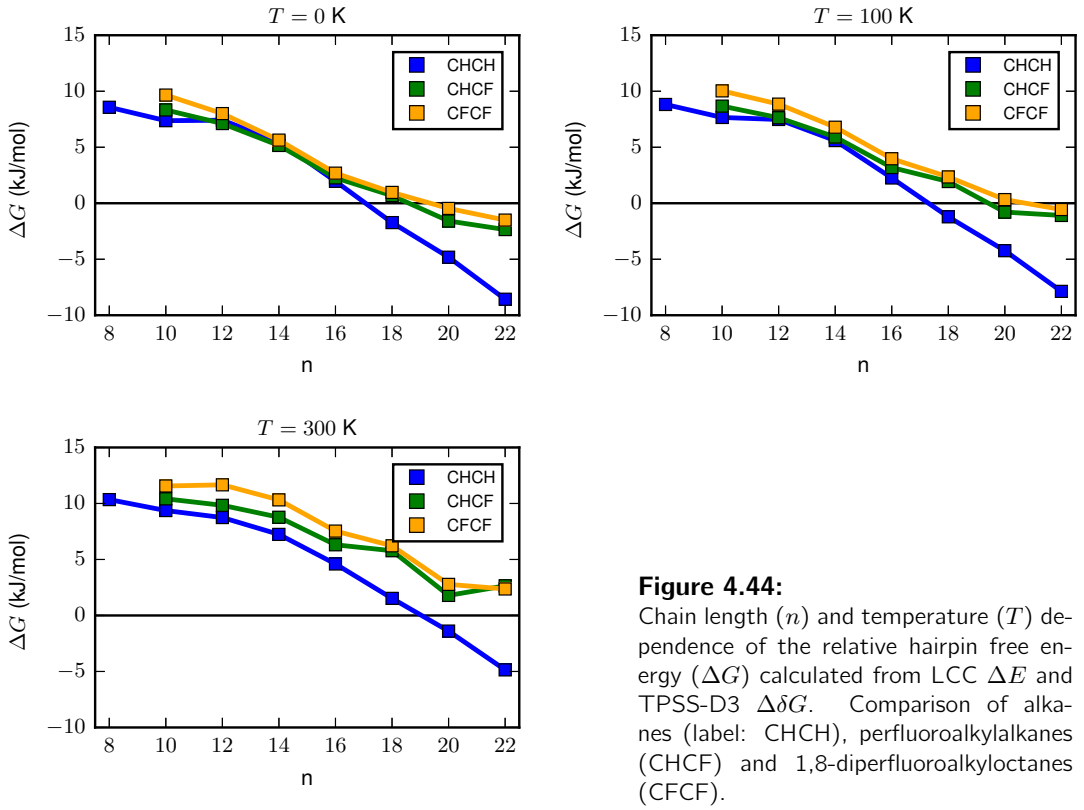


Figure 4.44:
Chain length (n) and temperature (T) dependence of the relative hairpin free energy (ΔG) calculated from LCC ΔE and TPSS-D3 $\Delta\delta G$. Comparison of alkanes (label: CHCH), perfluoroalkylalkanes (CHCF) and 1,8-diperfluoroalkyloctanes (CFCF).

fluoroalkyl groups pose a challenge for some dispersion corrected DFT methods and that interactions energies of perfluoroalkane dimers should be included in future re-parametrizations.

The relative hairpin free energy ΔG for all three hairpin systems is given in figure 4.44. At $T = 0$ K the relative hairpin free energy in the range $n = 12$ to $n = 16$ is virtually identical for all three systems. For chains with $n > 16$ the alkane hairpins are clearly more stabilized compared to the perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpins. The last globally stable linear (extended) conformers, with even chain length, at $T = 0$ K are predicted for $n = 16$ (alkanes) and $n = 18$ (perfluoroalkylalkanes and 1,8-diperfluoroalkylalkanes). Hairpin formation at $T = 0$ K is predicted by our results for $n = 18$ (alkanes) and $n = 20$ (perfluoroalkylalkanes and 1,8-diperfluoroalkylalkanes). The changes form $T = 0$ K to $T = 100$ K are subtle and result in a shift of the hairpin critical chain-lenght of the 1,8-diperfluoroalkylalkanes to $n = 22$, see figure 4.44c. However, considering the accuracy of the results (± 2 kJ/mol) and the similarity of the relative hairpin free energies, we can effectively only speak of a range $n = 20$ to $n = 24$ in which we expect the the turning point for the perfluoroalkylalkane and 1,8-diperfluoroalkylalkane hairpins at $T = 100$ K. We clearly expect perfluoroalkylalkane and 1,8-diperfluoroalkylalkane hairpin formation to occur for slightly ($\Delta n = 2$ to $\Delta n = 6$) longer chain lengths than in the case of n-alkanes. Thermodynamic corrections for the three systems are similar. Thus, the main reason for the lower hairpin stability is the lower amount of dispersion interactions present for the less well aligned helical perfluoroalkylgroups in comparison to all-trans alkyl groups. The calculated relative free energy energies of the stable perfluoroalkylalkane and 1,8-diperfluoroalkylalkane hairpins are very similar. Our $T = 100$ K predictions for the n-alkane hairpins are in agreement with the experiment [32] and other high level *ab*

initio studies [32–35]. Concerning the differences in relative hairpin free energy for $n \geq 18$, it should be possible to experimentally resolve the hairpin stability differences between the n-alkanes and the proposed partially fluorinated alkanes using the technique of Lüttschwager *et al.* [32, 33, 182]. At $T = 300$ K hairpin stability decreases further and only n-alkanes with $n \geq 20$ are predicted to form stable hairpins, see figure 4.44c. However at high temperatures additional gauche angles will be introduced along the chain and formed hairpins may open. Thus, the rigid hairpin model with straight arms looses its validity.

4.4.1 Hairpin conformers as solvation models

It is interesting, to find out what the low temperature gas phase hairpin model is predicting about the miscibility of alkanes and perfluoroalkanes. The miscibility model based on the hairpins is simple. We compare for each n the sum of the relative hairpin energies of the alkane and the 1,8-diperfluoroalkyloctane hairpin, which represent the separated phases, to the relative hairpin energies of two perfluoroalkylalkane hairpins, which represent the mixed phase. The mixing energy is given by

$$\Delta E_{\mathrm{mix}}(n) = 2\Delta E_{\mathrm{CHCF}}(n) - \Delta E_{\mathrm{CHCH}}(n) - \Delta E_{\mathrm{CFCF}}(n), \tag{4.17}$$

where the labels CHCF, CHCH and CFCF denote n-alkanes, perfluoroalkylalkanes and 1,8-diperfluoroalkyloctanes, respectively. In the same manner we can also compare the respective relative hairpin enthalpies. The free energy of mixing is given by

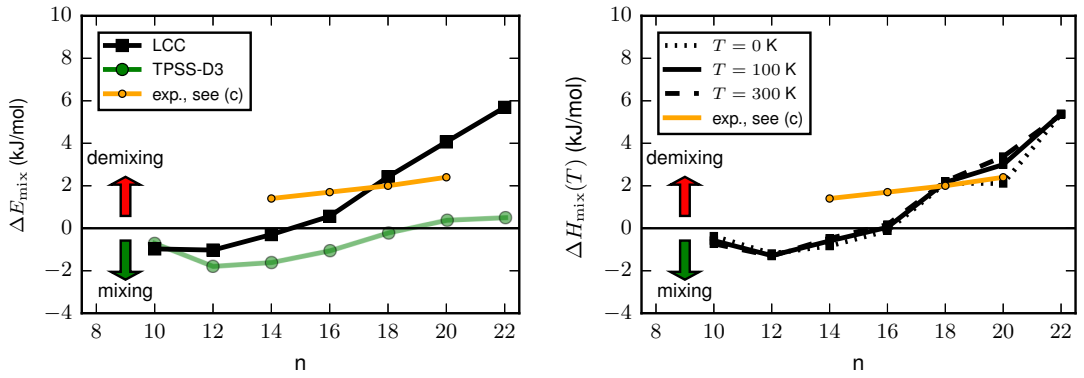
$$\Delta G_{\mathrm{mix}} = \Delta H_{\mathrm{mix}} - T\Delta S_{\mathrm{mix}}, \tag{4.18}$$

where the entropy of mixing is assumed to be always positive, since for an ideal mixture one can assume

$$\Delta S_{\mathrm{mix}} = -R(x_1 \ln(x_1) + x_2 \ln(x_2)) > 0, \tag{4.19}$$

where x_1 and x_2 are the mole fractions of a binary mixture [201]. Simply spoken, mixing is favoured entropically, due to the increased disorder in the mixed state. Hence, in our case, mixing will occur for $\Delta H_{\mathrm{mix}} \leq 0$ or $\Delta E_{\mathrm{mix}} \leq 0$, if we use $\Delta E_{\mathrm{mix}} \approx \Delta H_{\mathrm{mix}}$. If it is energetically more effective to fold two perfluoroalkylalkanes instead of folding one alkane and one 1,8-diperfluoroalkyloctane of the same length, then $\Delta E_{\mathrm{mix}}(n) < 0$ and the model would predict mixing. In the opposite case ($\Delta E_{\mathrm{mix}}(n) > 0$) it is energetically more favourable to fold the alkane and the 1,8-diperfluoroalkyloctane and demixing is predicted, under the assumption that the magnitude of $-T\Delta S_{\mathrm{mix}}$ is smaller than ΔE_{mix}. The hairpin model is very crude and no attempt is made to transfer its energies and enthalpies to an actual solvent model.

Figure 4.45a shows the predicted mixing or demixing behaviour predicted by the relative hairpin energies by the LCC and TPSS-D3 methodologies. LCC predicts mixing for $n = 10$ to $n = 14$ and demixing for $n \geq 16$. The reason for the demixing is given by the similar stability of both perfluoroalkylalkane and 1,8-diperfluoroalkyloctane hairpins, which are less stable in comparison to the respective alkane hairpins. TPSS-D3 predicts mostly mixing, with slight evidence for demixing for $n \geq 20$. Here the predicted behaviour occurs mainly due to the slight overestimation of perfluoroalkyl-alkyl interactions. The predictions based on the enthalpy are given in figure 4.45b and they follow the same trend as the LCC predictions. For comparison,



(a) Mixing energy as estimated from relative hairpin energies. Comparison of LCC and TPSS-D3 results.

(b) Enthalpy of mixing from hairpin model, based on LCC ΔE and TPSS-D3 $\Delta\delta H$.

m	T (K)	$\Delta H^T_{\mathrm{mix}}$ (kJ/mol)
3	214	1.4
4	245	1.7
5	285	2.0
6	308	2.4

(c) Experimental enthalpies of mixing $\Delta H^T_{\mathrm{mix}}$ at temperatures (T) of 1:1 liquid mixtures of n-alkanes and perfluoro-n-alkanes of chain length m, taken from refs. [202, 203].

Figure 4.45:
Predictions of the hairpin model for the mixing of n-alkanes and perfluoro-n-alkanes of same chain length given by $m = (n-8)/2$. For comparison, the actual enthalpies of mixing of the liquids [202, 203] are given in panel (c).

the measured enthalpies of mixing for 1:1 C_mH_{2m+2}/C_mF_{2m+2} mixtures [202, 203] are also given in figure 4.45.

It should be emphasized again, that our predictions are based on gas phase rigid hairpin formation low temperature enthalpies, and thus on a model that lacks many properties of actual liquids. First, the approach of the groups is constrained by the hairpin geometry. Secondly, interface effects like intramolecular dipole moments are absent in the pure alkane and perfluoroalkane solutions. Thirdly, in hairpins only a pair of groups interacts, while in solution many-body interactions are important. Fourthly, rigid hairpins are a low temperature phenomenon, while in solutions the flexibility of the molecules, *i.e.* their conformational distributions, at high temperatures should be taken into account. Nevertheless, the self solvating behaviour of cold chain molecules is by itself an interesting topic. It will be instructive to learn in the future, whether the findings on the hairpin systems can be confirmed experimentally and how they relate to the reasons for the observed fluorophilicity in solution.

4.5 Summary and conclusions

Herein the question of n-alkane hairpin stability [31] was extended to partially fluorinated n-alkanes, in order to compare intramolecular alkyl-alkyl, perfluoroalkyl-perfluoroalkyl and perfluoroalkyl-alkyl van der Waals interactions. Accurate relative hairpin energies were calculated using local correlation methods and in conjunction with resolution of the identity/density fitting approximations and the method of increments. The main part of the correlation energy

was calculated using explicitly correlated local MP2 methods with aVTZ basis sets. For the LMP2-F12 part extended domains were employed and all pair domains were included in the calculation. The method of increments was applied to obtain energies for the larger systems and to provide 2-body maps of the mainly dispersive intramolecular correlation interactions. Electron correlation effects beyond MP2 were treated via a small (aVDZ,VDZ) basis LCCSD(T0) correction. For larger fluorinated systems, we had to extrapolate this correction based on the values obtained for smaller systems. Calculations on CH_4/CH_4, CH_4/CF_4 and CF_4/CF_4 dimers confirmed the excellent performance of the method, which is effectively BSSE free and provides accurate interaction energies. The structures of the chain molecules were optimized using DFT in form of MARI-TPSS-D3/def2-TZVP, which was also used to calculate thermodynamic corrections in the rigid rotator, unscaled harmonic oscillator approximations. In the vibrational entropy term, low frequency vibrations where approximated by rigid rotors according to the scheme by Grimme. The composite local correlation method applied in this work agrees well on the relative hairpin energies of n-alkanes, obtained by others using high level results of CCSD(T)/CBS quality [34, 35]. DFT calculations with similar functionals (TPSS, PBE) and three different approaches to dispersion corrections show, that the results obtained with the D3 method agree best with our reference results, while the VV10 and the QHO methods overestimate hairpin stabilization due to perfluoroalkylgroups. Overall, we find that alkyl-alkyl interactions lead to more stable hairpin conformers in comparison to the weaker perfluoroalkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions. For groups with $m \geq 4$ (*i.e.* butyl) connected the central -$(CH_2)_8$- unit, the ordering of the interactions in terms of hairpin stabilization is alkyl-alkyl > perfluoroalkyl-alkyl $\gtrsim$ perfluoroalkyl-perfluoroalkyl. At $T = 100$ K, the calculated free energies predict hairpin conformers to be more stable than the linear conformers beginning from chain lengths of $n = 18$ for the n-alkanes, $n = 20$ for the perfluoroalkylalkanes and $n = 22$ for the 1,8-diperfluoroalkyloctanes.

4.6 Outlook

This study of partially fluorinated n-alkane folding is mainly motivated by the possibility of experimental studies similar to those on n-alkanes [32, 33] and perfluoro-n-alkanes [182]. We therefore hope that this theoretical investigation will motivate future experiments on the proposed partially fluorinated n-alkanes. For this the vibrational spectra, especially their Raman intensities for conformational ensembles, see ref. [33], need to be evaluated.

Furthermore, it would be very interesting to apply the highly efficient domain-based local pair natural orbital coupled cluster methodology [35, 204] to the problem of the hairpin stability of partially fluorinated n-alkanes. This would provide a reference to the LCC methodology used in this work. Double hybrid DFT-D3 appoaches should be tested for the partially fluorinated hairpins. Additional insight into the nature of the weak intramolecular interactions could be gained by the application of intramolecular SAPT [181] and a recent visualisation method for dispersion interactions [205].

In this work the method of increments was applied using a brute force approach. However, only a fraction of the increments is actually necessary to approximate the correlation energy of the investigated conformational change. An analysis of pairs of local orbitals in terms of changes of their mutual orientation and distance could provide a basis for methods which are extremely efficient for the calculation of conformational energies, where a large part of the

molecular structure remains similar. Future analysis of the data produced in this work may also help to develop an efficient treatment of post-MP2 corrections for conformational energies.

We found evidence, that pairs of helical perfluoroalkyl groups show less dispersion interactions than a pair of alkyl groups when confined to the hairpin structure. This raises the question: Which orientation supports the most favourable binding interaction of two helical perfluoro-*n*-alkanes? Another question that is related to the latter is: Do respective hypothetical non-helical perfluoro-*n*-alkanes interact more strongly compared to the actual helical conformers? For such a study, efficient local correlation methods and DFT-SAPT [109] should be applied. This would also help to improve dispersion corrected DFT methods with respect to their description of perfluoroalkyl interactions. It would also be interesting to apply recent force fields [206] for fluorinated *n*-alkanes, which can predict the observed demixing, to the herein investigated systems.

In case of the investigated hairpins, molecular dynamics and Monte Carlo simulations could be employed to investigate thier folding dynamics and statistics. In this regard, an interesting question is: how do the rigid perfluoroalkyl groups influence the folding pathways compared to the flexible alkyl groups?

In supramolecular chemistry ordering effects of perfluoroalkyl groups are often used in solution. Hence, solvent effects become very important. For studies of perfluoroalkyl and alkyl interactions in solution, molecular torsional balances are potentially better suited than hairpins. Molecular torsional balance are specially designed molecules, where a relatively unhindered rotation around a singe C-C bond can bring two groups in close contact. An impressive study in this regard was performed by Adam *et al.*, who measured the cohesive interaction energies between alkyl chains and between perfluoroalkyl chains in solution using molecular torsional balances [207]. It would certainly be interesting to determine, if it is possible to reproduce their findings with DFT-D3 calculations in conjunction with a continuum solvation model as in ref. [191].

The experimental and theoretical study of molecular interactions with respect to perfluoroalkyl groups is a fascinating subject with many challenging open questions that demand creative solutions.

5 Appendix

5.1 Bound state energy tables for $p\mathrm{H}_2$-X_2 dimers.

The tables 5.1, 5.2 and 5.3 contain the bound state energies of the $p\mathrm{H}_2$-$\mathrm{F}2_2$, $p\mathrm{H}_2$-Cl_2 and $p\mathrm{H}_2$-Br_2 dimers, based on the calculation described in section 3.4.

$p\mathrm{H}_2$-F_2, $n = 0$					
Nr.	E (kJ/mol)	location	L bending	T bending	stretching
0	−0.2094	L	0	-	0
1	−0.2070	L	0	-	0
2	−0.1691	T,(L)	(0)	0	0
3	−0.0797	T,L	0	1	0
4	−0.0124	T,L	0	2	0

Table 5.1: Bound state energies and assignment of localization and quantum numbers for the bending (θ) and stretching (R) modes of vibrational states of the $J_{\mathrm{tot}} = 0$ $p\mathrm{H}_2$-F_2 van der Waals dimer.

$p\mathrm{H_2}$-$\mathrm{Cl_2}$, $n = 0$					
Nr.	E (kJ/mol)	location	L bending	T bending	stretching
0	−0.8932	L	0	-	0
1	−0.8932	L	0	-	0
2	−0.7981	T	-	0	0
3	−0.6372	T	-	1	0
4	−0.5083	T	-	2	0
5	−0.4714	L	1	-	0
6	−0.4712	L	1	-	0
7	−0.4175	T	-	3	0
8	−0.3603	T,L	1	4	0
9	−0.3245	T,L	1	5	0
10	−0.2857	T,L	1	6	0
11	−0.2327	T,L			0
12	−0.1742	T,L			0
13	−0.1593	T	-	0	1
14	−0.1462	L	0	-	1
15	−0.1423	L	0	-	1
16	−0.1041	T,L			0
17	−0.0644	T	-	1	1
18	−0.0363	T,L			0
19	−0.0291	T,L			1
20	−0.0159	L	1		1

Table 5.2: Bound state energies and assignment of localization and quantum numbers for the bending (θ) and stretching (R) modes of vibrational states of the $J_{\mathrm{tot}} = 0$ $p\mathrm{H_2}$-$\mathrm{Cl_2}$ van der Waals dimer.

pH$_2$-Br$_2$, $n = 0$					
Nr.	E (kJ/mol)	location	L bending	T bending	stretching
0	−1.3402	L	0	-	0
1	−1.3402	L	0	-	0
2	−0.9608	T	-	0	0
3	−0.9213	L	1	-	0
4	−0.9213	L	1	-	0
5	−0.8527	T	-	1	0
6	−0.7490	T	-	2	0
7	−0.6574	T	-	3	0
8	−0.6277	L	2	-	0
9	−0.6277	L	2	-	0
10	−0.5806	T	-	4	0
11	−0.5160	T	-	5	0
12	−0.4613	T,L	3	6	0
13	−0.4544	L	3	-	0
14	−0.4532	L,T	3	6	0
15	−0.4173	L	0	-	1
16	−0.4172	L	0	-	1
17	−0.4166	T,L			0
18	−0.3911	L,T			0
19	−0.3705	L,T			0
20	−0.3427	L,T			0
21	−0.3113	L,T			0
22	−0.2777	L,T			0
23	−0.2745	T	-	0	1
24	−0.2411	L,T			0
25	−0.2021	L,T			0
26	−0.1967	T	-	1	1
27	−0.1608	L,T			0
28	−0.1573	L	1	-	1
29	−0.1572	L	1	-	1
30	−0.1334	T	-	2	1
31	−0.1171	L,T			0
32	−0.0886	T	-	3	1
33	−0.0712	L,T			0
34	−0.0602	T,L	1	4	1
35	−0.0533	L	2	-	1
36	−0.0472	L,T	2	4	1
37	−0.0288	L.T			1
38	−0.0233	L,T			0
39	−0.0131	T	-	0	2
40	−0.0109	L,T			0
41	−0.0071	L	0	-	2
42	−0.0071	L	0	-	2

Table 5.3: Bound state energies and assignment of localization and quantum numbers for the bending (θ) and stretching (R) modes of vibrational states of the $J_{\text{tot}} = 0$ pH$_2$-Br$_2$ van der Waals dimer.

Bibliography

[1]F. London, "Zur Theorie und Systematik der Molekularkräfte", Z. Phys. **63**, 245–279 (1930).

[2]F. London, "The general theory of molecular forces", Trans. Faraday Soc. **33**, 8b–26 (1937).

[3]A. Stone, *The Theory of Intermolecular Forces*, International Series of Monographs on Chemistry (Clarendon Press, Oxford, Great Britain, 2000).

[4]T. Steiner, "The Hydrogen Bond in the Solid State", Angew. Chem. Int. Ed. **41**, 48–76 (2002).

[5]M. H. Kolář and P. Hobza, "Computer Modeling of Halogen Bonds and Other σ-Hole Interactions", Chem. Rev. **116**, 5155–5187 (2016).

[6]G. Maitland, M. Rigby, E. Smith, and W. Wakeham, *Intermolecular Forces*, International Series of Monographs on Chemistry (Clarendon Press, Oxford, Great Britain, 1981).

[7]K. Szalewicz and B. Jeziorski, "Symmetry-adapted Perturbation Theory of Intermolecular Interactions", in *Molecular interactions*, edited by S. Scheiner (John Wiley & Sons, Inc., 1997), pp. 3–43.

[8]J. P. Wagner and P. R. Schreiner, "London dispersion in molecular chemistry reconsidering steric effects", Angew. Chem. Int. Ed. **54**, 12274–12296 (2015).

[9]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu", J. Chem. Phys. **132**, 154104 (2010).

[10]W. Nils, *Lehrbuch der Anorganischen Chemie, 102. Auflage* (De Gruyter, Berlin, Boston, 2008).

[11]D. R. Lide, ed., *CRC Handbook of Chemistry and Physics, 87th Edition* (CRC Press, Boca Raton, FL, 2006).

[12]P. Kirsch, *Modern fluoroorganic chemistry synthesis, reactivity, applications* (Wiley-VCH Verlag GmbH, Weinheim, Germany, 2013).

[13]H. H. Anderson, "Boiling Points and Boiling Point Numbers of Perfluoroalkanes and Perfluoroalkenes.", J. Chem. Eng. Data **10**, 156–159 (1965).

[14]J. H. Hildebrand and D. R. F. Cochran, "Liquid-Liquid Solubility of Perfluoromethylcyclohexane with Benzene, Carbon Tetrachloride, Chlorobenzene, Chloroform and Toluene", J. Am. Chem. Soc. **71**, 22–25 (1949).

[15]J. H. Clark and J. M. Duncan, eds., *Handbook of Green Chemistry and Technology* (Blackwell Science Ltd., Oxford, England, 2002).

[16]J. G. Riess, "Fluorous micro- and nanophases with a biomedical perspective", Tetrahedron **58**, 4113–4131 (2002).

[17]E. K. Watkins, and W. L. Jorgensen, "Perfluoroalkanes: Conformational Analysis and Liquid-State Properties from ab Initio and Monte Carlo Calculations", J. Phys. Chem. A **105**, 4118–4125 (2001).

[18]J. Chattoraj, T. Risthaus, O. Rubner, A. Heuer, and S. Grimme, "A multi-scale approach to characterize pure CH_4, CF_4, and CH_4/CF_4 mixtures", J. Chem. Phys. **142**, 164508 (2015).

[19]S. Tsuzuki, T. Uchimaru, M. Mikami, and S. Urata, "Magnitude and orientation dependence of intermolecular interaction between perfluoroalkanes: High level ab initio calculations of CF_4 and C_2F_6 dimers", J. Chem. Phys. **116**, 3309–3315 (2002).

[20]S. Tsuzuki, T. Uchimaru, M. Mikami, and S. Urata, "Magnitude and orientation dependence of intermolecular interaction of perfluoropropane dimer studied by high-level ab initio calculations: Comparison with propane dimer", J. Chem. Phys. **121**, 9917–9924 (2004).

[21]R. Mahlanen, J.-P. Jalkanen, and T. A. Pakkanen, "Potential energy surfaces of CF_4, CCl_4 and CBr_4 dimers", Chem. Phys. **313**, 271–277 (2005).

[22]W. Zhang, "Fluorous synthesis of heterocyclic systems", Chem. Rev. **104**, 2531–2556 (2004).

[23]M. Zieringer, M. Wyszogrodzka, K. Biskup, and R. Haag, "Supramolecular behavior of fluorous polyglycerol dendrons and polyglycerol dendrimers with perfluorinated shells in water", New J. Chem. **36**, 402–406 (2012).

[24]O. Wagner, B. N. S. Thota, B. Schade, F. Neumann, J. L. Cuellar, C. Bottcher, and R. Haag, "Perfluoroalkylated linear polyglycerols and their supramolecular assemblies in aqueous solution", Polym. Chem. **7**, 2222–2229 (2016).

[25]D. M. Lemal, "Perspective on Fluorocarbon Chemistry", J. Org. Chem. **69**, 1–11 (2004).

[26]H. Li, P.-N. Roy, and R. J. Le Roy, "An "Adiabatic-Hindered-Rotor" Treatment Allows *para*-H_2 to be Treated as if it were Spherical", J. Chem. Phys. **133**, 104305–14 (2010).

[27]A. Accardi, "Quantum Mechanical Studies of Small Molecular Systems in Cryogenic Environments", PhD thesis (Department of Biology, Chemistry and Pharmacy - Freie Universität Berlin, 2012).

[28]J. Beswick, N. Halberstadt, and K. Janda, "Structure and dynamics of noble gas-halogen and noble gas ionic clusters: When theory meets experiment", Chem. Phys. **399**, 4–16 (2012).

[29]S. C. Kettwich, P. L. Raston, and D. T. Anderson, "The Cl+H_2 → HCl+H Reaction Induced by IR+UV Irradiation of Cl_2 in Solid para-H_2: Experiment", J. Phys. Chem. A **113**, 7621–9 (2009).

[30]M. V. Korolkov, J. Manz, and A. Schild, "The Cl + H_2 → HCl + H Reaction Induced by IR + UV Irradiation of Cl_2 in Solid para-H_2: Quantum Model Simulation", J. Phys. Chem. A **113**, 7630–46 (2009).

[31]J. M. Goodman, "What Is the Longest Unbranched Alkane with a Linear Global Minimum Conformation?", Journal of Chemical Information and Computer Sciences **37**, 876–878 (1997).

[32]N. O. B. Lüttschwager, T. N. Wassermann, R. A. Mata, and M. A. Suhm, "The Last Globally Stable Extended Alkane", Angew. Chem. Int. Ed. **52**, 463–466 (2013).

[33]N. O. B. Lüttschwager and M. A. Suhm, "Stretching and folding of 2-nanometer hydrocarbon rods", Soft Matter **10**, 4885–4901 (2014).

[34]J. N. Byrd, R. J. Bartlett, and J. A. Montgomery, "At What Chain Length Do Unbranched Alkanes Prefer Folded Conformations?", J. Phys. Chem. A **118**, 1706–1712 (2014).

[35]D. G. Liakos and F. Neese, "Domain Based Pair Natural Orbital Coupled Cluster Studies on Linear and Folded Alkane Chains", J. Chem. Theory Comput. **11**, 2137–2143 (2015).

[36]S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory", J. Comput. Chem. **32**, 1456–1465 (2011).

[37]S. Grimme, "Supramolecular binding thermodynamics by dispersion-corrected density functional theory", Chem. Eur. J. **18**, 9955–9964 (2012).

[38]O. A. Vydrov and T. Van Voorhis, "Nonlocal van der Waals density functional: The simpler the better", J. Chem. Phys. **133**, 244103 (2010).

[39]P. Partovi-Azar, M. Berg, S. Sanna, and T. D. Kühne, "Improved parameterization of the quantum harmonic oscillator model based on localized Wannier functions to describe Van der Waals interactions in density functional theory", Int. J. Quantum Chem. **116**, 1160–1165 (2016).

[40]C. Hampel and H.-J. Werner, "Local treatment of electron correlation in coupled cluster theory", J. Chem. Phys. **104**, 6286–6297 (1996).

[41]H.-J. Werner and M. Schütz, "An efficient local coupled cluster method for accurate thermochemistry of large systems", J. Chem. Phys. **135**, 144116 (2011).

[42]G. Stollhoff and P. Fulde, "A local approach to the computation of correlation energies of molecules", Z. Phys. B **26**, 257–262 (1977).

[43]G. Stollhoff and P. Fulde, "Description of intraatomic correlations by the Local Approach", Z. Phys. B **29**, 231–237 (1978).

[44]B. Paulus, "The method of increments a wavefunction-based ab initio correlation method for solids", Phys. Rep **428**, 1–52 (2006).

[45]C. Müller and B. Paulus, "Wavefunction-based electron correlation methods for solids", Phys. Chem. Chem. Phys. **14**, 7605–7614 (2012).

[46]J. Friedrich and J. Hänchen, "Incremental CCSD(T)(F12)—MP2: A Black Box Method To Obtain Highly Accurate Reaction Energies", J. Chem. Theory Comput. **9**, 5381–5394 (2013).

[47]A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover Publications Inc., Mineola, N.Y., 1996).

[48]S. Chelkowski, T. Zuo, O. Atabek, and A. D. Bandrauk, "Dissociation, ionization, and Coulomb explosion of H_2^+ in an intense laser field by numerical integration of the time-dependent Schrödinger equation", Phys. Rev. A **52**, 2977–2983 (1995).

[49]M. Baer, *Beyond Born-Oppenheimer: Electronic Nonadiabatic Coupling Terms and Conical Intersections* (John Wiley & sons, Inc. Hoboken, New Jersey, 2006).

[50]M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln", Ann. Phys. **389**, 457–484 (1927).

[51]M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, International series of monographs on physics (Clarendon Press, Oxford, 1998).

[52]F. Jensen, *Introduction to Computational Chemistry* (John Wiley & Sons Ltd, Chichester, England, 1999).

[53]T. Helgaker, P. Jørgensen, and O. Jeppe, *Molecular Electronic-Structure Theory* (John Wiley & Sons Ltd, Chichester, England, 2000).

[54]W. Kutzelnigg and W. Klopper, "Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory", J. Chem. Phys. **94**, 1985–2001 (1991).

[55]H.-J. Werner, T. B. Adler, and F. R. Manby, "General orbital invariant MP2-F12 theory", The Journal of Chemical Physics **126**, 164102 (2007).

[56]T. B. Adler, G. Knizia, and H.-J. Werner, "A simple and efficient CCSD(T)-F12 approximation", The Journal of Chemical Physics **127**, 221106 (2007).

[57]T. B. Adler, H.-J. Werner, and F. R. Manby, "Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules", J. Chem. Phys. **130**, 054106 (2009).

[58]T. B. Adler and H.-J. Werner, "Local explicitly correlated coupled-cluster methods: Efficient removal of the basis set incompleteness and domain errors", The Journal of Chemical Physics **130**, 241101 (2009).

[59]D. P. Tew and W. Klopper, "A comparison of linear and nonlinear correlation factors for basis set limit Møller-Plesset second order binding energies and structures of He_2, Be_2, and Ne_2", J. Chem. Phys. **125**, 094302 (2006).

[60]P. Pulay, "Localizability of dynamic electron correlation", Chem. Phys. Lett. **100**, 151–154 (1983).

[61]S. Saebo and P. Pulay, "Local Treatment of Electron Correlation", Annu. Rev. Phys. Chem. **44**, 213–236 (1993).

[62]J. Pipek and P. G. Mezey, "A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions", J. Chem. Phys. **90**, 4916–4926 (1989).

[63]J. M. Foster and S. F. Boys, "Canonical Configurational Interaction Procedure", Rev. Mod. Phys. **32**, 300–302 (1960).

[64]J. W. Boughton and P. Pulay, "Comparison of the Boys and Pipek-Mezey localizations in the local correlation approach and automatic virtual basis selection", J. Comput. Chem. **14**, 736–740 (1993).

[65]M. Schütz and H.-J. Werner, "Local perturbative triples correction (T) with linear cost scaling", Chem. Phys. Lett. **318**, 370–378 (2000).

[66]M. Schütz, "Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T)", J. Chem. Phys. **113**, 9986–10001 (2000).

[67]F. Neese, F. Wennmohs, and A. Hansen, "Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method", J. Chem. Phys. **130**, 114108 (2009).

[68]J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, and H.-J. Werner, "The orbital-specific-virtual local coupled cluster singles and doubles method", J. Chem. Phys. **136**, 144105 (2012).

[69]C. Krause and H.-J. Werner, "Comparison of explicitly correlated local coupled-cluster methods with various choices of virtual orbitals", Phys. Chem. Chem. Phys. **14**, 7591–7604 (2012).

[70]H. Stoll, "Correlation energy of diamond", Phys. Rev. B **46**, 6700–6704 (1992).

[71]H. Stoll, "The correlation energy of crystalline silicon", Chem. Phys. Lett. **191**, 548–552 (1992).

[72]H. Stoll, "On the correlation energy of graphite", J. Chem. Phys. **97**, 8449–8454 (1992).

[73]L. Faddeev, "Scattering Theory for a System of Three Particles", Engl. Transl. Sov. Phys. Jetp **12**, 1014–1019 (1961).

[74]R. K. Nesbet, "Atomic Bethe-Goldstone Equations", in *Advances in chem. phys.* (John Wiley & Sons, Inc., 1969), pp. 1–34.

[75]H. Stoll, B. Paulus, and P. Fulde, "On the accuracy of correlation-energy expansions in terms of local increments", J. Chem. Phys. **123**, 144108 (2005).

[76]K. G. Steenbergen, N. Gaston, C. Müller, and B. Paulus, "Method of increments for the halogen molecular crystals: Cl, Br, and I", J. Chem. Phys. **141**, 124707 (2014).

[77]E. Voloshina and B. Paulus, "First Multireference Correlation Treatment of Bulk Metals", J. Chem. Theory Comput. **10**, 1698–1706 (2014).

[78]E. Fertitta, B. Paulus, G. Barcza, and Ö. Legeza, "On the calculation of complete dissociation curves of closed-shell pseudo-onedimensional systems via the complete active space method of increments", J. Chem. Phys. **143**, 114108 (2015).

[79]P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", Phys. Rev. **136**, B864–B871 (1964).

[80]W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Phys. Rev. **140**, A1133–A1138 (1965).

[81]S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis", Can. J. Phys. **58**, 1200–1211 (1980).

[82]J. P. Perdew and W. Yue, "Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation", Phys. Rev. B **33**, 8800–8802 (1986).

[83]J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", Phys. Rev. B **45**, 13244–13249 (1992).

[84]C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", Phys. Rev. B **37**, 785–789 (1988).

[85]J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

[86]J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the Density Functional Ladder: Nonempirical Meta Generalized Gradient Approximation Designed for Molecules and Solids", Phys. Rev. Lett. **91**, 146401 (2003).

[87]A. D. Becke, "A new mixing of Hartree-Fock and local density-functional theories", J. Chem. Phys. **98**, 1372–1377 (1993).

[88]A. D. Becke, "Density functional thermochemistry. III. The role of exact exchange", J. Chem. Phys. **98**, 5648–5652 (1993).

[89]V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, "Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes", J. Chem. Phys. **119**, 12129–12137 (2003).

[90]S. Grimme, "Semiempirical hybrid density functional with perturbative second-order correlation", J. Chem. Phys. **124**, 034108 (2006).

[91]T. Schwabe and S. Grimme, "Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability", Phys. Chem. Chem. Phys. **9**, 3397–3406 (2007).

[92]P. L. Silvestrelli, "Van der Waals interactions in density functional theory by combining the quantum harmonic oscillator-model with localized Wannier functions", J. Chem. Phys. **139**, 054106 (2013).

[93]J. Cao and B. J. Berne, "Many-body dispersion forces of polarizable clusters and liquids", J. Chem. Phys. **97**, 8628–8636 (1992).

[94]O. A. Vydrov and T. V. Voorhis, "Benchmark Assessment of the Accuracy of Several van der Waals Density Functionals", J. Chem. Theory Comput. **8**, 1929–1934 (2012).

[95]W. Hujo and S. Grimme, "Performance of Non-Local and Atom-Pairwise Dispersion Corrections to DFT for Structural Parameters of Molecules with Noncovalent Interactions", J. Chem. Theory Comput. **9**, 308–315 (2013).

[96]Y. Qiu and Z. Bačić, "Exact six-dimensional quantum calculations of the rovibrational levels of $(HCl)_2$", J. Chem. Phys. **106**, 2158–2170 (1997).

[97]K. A. Peterson and T. H. Dunning, "Benchmark calculations with correlated molecular wave functions. VII. Binding energy and structure of the HF dimer", J. Chem. Phys. **102**, 2032–2041 (1995).

[98]R. J. Donnelly, "Fritz London: A Scientific Biography (Book Review)", American Journal of Physics **64**, 827–828 (1996).

[99]K. Gavroglu, *Fritz London: A Scientific Biography*, Sescelades: Biografia (Cambridge University Press, 2005).

[100]H. Hettema, *Quantum Chemistry: Classic Scientific Papers*, World Scientific series in 20th century chemistry (World Scientific Publishing Company Pte Limited, 2000).

[101]F. London, "On some Properties and Applications of Molecular Forces", Z. Phys. Chem. B **11**, 222 (1930).

[102]H. B. G. Casimir and D. Polder, "The Influence of Retardation on the London-van der Waals Forces", Phys. Rev. **73**, 360–372 (1948).

[103]S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors", Mol. Phys. **19**, 553–566 (1970).

[104]F. van Duijneveldt, "Basis Set Superposition Error", in *Molecular interactions*, edited by S. Scheiner (John Wiley & Sons, Inc., 1997), pp. 81–104.

[105]M. Schütz, G. Rauhut, and H.-J. Werner, "Local Treatment of Electron Correlation in Molecular Clusters: Structures and Stabilities of $(H_2O)_n$, n = 2-4", J. Phys. Chem. A **102**, 5997–6003 (1998).

[106]B. Jeziorski, K. Szalewicz, and G. Chałasiński, "Symmetry forcing and convergence properties of perturbation expansions for molecular interaction energies", Int. J. Quantum Chem. **14**, 271–287 (1978).

[107]E. G. Hohenstein and C. D. Sherrill, "Wavefunction methods for noncovalent interactions", Wiley Interdiscip. Rev.-Comput. Mol. Sci. **2**, 304–326 (2012).

[108]K. Szalewicz, "Symmetry-adapted perturbation theory of intermolecular forces", Wiley Interdiscip. Rev.-Comput. Mol. Sci. **2**, 254–272 (2012).

[109]G. Jansen, "Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions", Wiley Interdiscip. Rev.-Comput. Mol. Sci. **4**, 127–144 (2014).

[110]M. Korolkov and J. Manz, "Resonant versus off-resonant quantum reaction dynamics in quantum solids: Model simulations for $Cl + H_2 \rightarrow ClH_2^* \rightarrow HCl + H$ in solid para-hydrogen", Chem. Phys. **370**, 159–167 (2010).

[111]L. Andrews and M. Moskovits, eds., *Chemistry and Physics of Matrix-Isolated Species* (Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1989).

[112]I. F. Silvera, "The solid Molecular Hydrogen in the Condensed Phase: Fundamentals and Static Properties", Rev. Mod. Phys. **52**, 393–452 (1980).

[113]P. L. Raston and D. T. Anderson, "Infrared-induced reaction of Cl atoms trapped in solid parahydrogen", Phys. Chem. Chem. Phys. **8**, 3124–3129 (2006).

[114]M. S. Anderson and C. A. Swenson, "Experimental Compressions for Normal Hydrogen and Normal Deuterium to 25 kbar at 4.2 K", Phys. Rev. B **10**, 5184–91 (1974).

[115]M. Rama Krishna and K. Whaley, "Structure of small molecular hydrogen clusters", Z. Phys. D Atom Mol. Cl. **20**, 223–226 (1991).

[116]M. A. McMahon and K. B. Whaley, "Variational and diffusion Monte Carlo studies of $(H_2)_N$ clusters", Chem. Phys. **182**, 119–130 (1994).

[117]F. Mezzacapo and M. Boninsegni, "Superfluidity and Quantum Melting of p-H_2 Clusters", Phys. Rev. Lett. **97**, 045301 (2006).

[118]M. B. Sevryuk, J. P. Toennies, and D. M. Ceperley, "Why are para-hydrogen clusters superfluid? A quantum theorem of corresponding states study", J. Chem. Phys. **133**, 064505 (2010).

[119]O. Kühn, J. Manz, and A. Schild, "Quantum Effects on Translational Motions in Solid para-Hydrogen and ortho-Deuterium: Anharmonic Extension of the Einstein Model", J. Phys.: Cond. Matt. **22**, 135401–8 (2010).

[120]K. Rościszewski and B. Paulus, "The Zero-Point Energy in the Molecular Hydrogen Crystal", Mol. Phys. **108**, 2147–52 (2010).

[121]A. Einstein, "Die Plancksche Theorie der Strahlung und die Theorie der Spezifischen Wärme", Ann. Phys. (Berlin) **22**, 180 (1907).

[122]H. Li, P.-N. Roy, and R. J. Le Roy, "An "Adiabatic-Hindered-Rotor" Treatment Allows *para*-H_2 to be Treated as if it were Spherical", J. Chem. Phys. **133**, 104305–14 (2010).

[123]T. C. Lillestolen and R. J. Hinde, "Vibrationally Averaged Isotropic Dispersion Energy Coefficients of the Parahydrogen Dimer", J. Chem. Phys. **136**, 204303–06 (2012).

[124]J. P. Darr, R. A. Loomis, S. E. Ray-Helmus, and A. B. McCoy, "Probing the Dependence of Long-Range, Four-Atom Interactions on Intermolecular Orientation: 3. Hydrogen and Iodine", J. Phys. Chem. A **115**, 7368–7377 (2011).

[125]J. E. Kenny, T. D. Russell, and D. H. Levy, "Van der Waals complexes of iodine with hydrogen and deuterium: Intermolecular potentials and laser-induced photodissociation studies", J. Chem. Phys. **73**, 3607–3616 (1980).

[126]W. Göepel and H.-D. Wiemhöfer, *Statistische Thermodynamik* (Spektrum Akademischer Verlag GmbH, Heidelberg, Germany, 2000).

[127]K. Yoshioka, P. L. Raston, and D. T. Anderson, "Infrared spectroscopy of chemically doped solid parahydrogen", Int. Rev. Phys. Chem. **25**, 469–496 (2006).

[128]P. L. Raston and D. T. Anderson, "Infrared-active vibron bands associated with rare gas atom dopants isolated in solid parahydrogen", Low Temp. Phys. **33**, 487–492 (2007).

[129]P. L. Raston, S. C. Kettwich, and D. T. Anderson, "Infrared studies of ortho-para conversion at Cl-atom and H-atom impurity centers in cryogenic solid hydrogen", Low Temp. Phys. **36**, 392–399 (2010).

[130]L. Pauling, "The Rotational Motion of Molecules in Crystals", Phys. Rev. **36**, 430–43 (1930).

[131]A. F. Devonshire, "The Rotation of Molecules in Fields of Octahedral Symmetry", Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **153**, 601–621 (1936).

[132]B. Schmidt and P. Jungwirth, "Vibrational line shifts of hydrogen halides in a rare gas environment: HF/DF and HCl/DCl in Ar matrices and clusters", Chem. Phys. Lett. **259**, 62–68 (1996).

[133]P. Jungwirth, P. Zdanska, and B. Schmidt, "Librational Control of Photochemical Reactions in Small Clusters", J. Phys. Chem. A **102**, 7241–44 (1998).

[134]T. Kiljunen, M. Bargheer, M. Guhr, N. Schwentner, and B. Schmidt, "Photodynamics and ground state librational states of ClF molecule in solid Ar. Comparison of experiment and theory", Phys. Chem. Chem. Phys. **6**, 2932–2939 (2004).

[135]T. Kiljunen, B. Schmidt, and N. Schwentner, "Intense-Field Alignment of Molecules Confined in Octahedral Fields", Phys. Rev. Lett. **94**, 123003–7 (2005).

[136]T. Kiljunen, B. Schmidt, and N. Schwentner, "Aligning and Orientating Molecules Trapped in Octahedral Crystal Fields", Phys. Rev. A **72**, 053415–26 (2005).

[137]T. Kiljunen, B. Schmidt, and N. Schwentner, "Time-dependent Alignment of Molecules Trapped in Octahedral Crystal Fields", J. Chem. Phys. **124**, 164502–15 (2006).

[138]T. Pham Van, "Ab initio calculation of intermolecular potentials, prediction of second virial coefficients for dimers H_2-H_2, H_2-O_2, F_2-F_2 and H_2-F_2, and Monte Carlo simulations of the vapor-liquid equilibria for hydrogen and fluorine", PhD thesis (Universität zu Köln, 2006).

[139]M. Berg, A. Accardi, B. Paulus, and B. Schmidt, "Rotationally adiabatic pair interactions of para- and ortho-hydrogen with the halogen molecules F_2, Cl_2, and Br_2", J. Chem. Phys. **141**, 074303 (2014).

[140]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, *MOLPRO, version 2012.1, a package of ab initio programs*, Cardiff, UK, 2012.

[141]C. Hampel, K. A. Peterson, and H.-J. Werner, "A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods", Chem. Phys. Lett. **190**, 1–12 (1992).

[142]M. J. Deegan and P. J. Knowles, "Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories", Chem. Phys. Lett. **227**, 321–326 (1994).

[143]T. H. Dunning, "Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen", J. Chem. Phys. **90**, 1007 (1989).

[144]K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg, "Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16-18 elements", J. Chem. Phys. **119**, 11113–11123 (2003).

[145]A. Bondi, "Van der Waals Volumes and Radii", J. Chem. Phys. **68**, 441–51 (1964).

[146]R. Ahlrichs, R. Penco, and G. Scoles, "Intermolecular Forces in Simple Systems", Chem. Phys. **19**, 119–30 (1977).

[147]K. W. Chan, T. D. Power, J. Jai-nhuknan, and S. M. Cybulski, "An ab initio study of He-F_2, Ne-F_2, and Ar-F_2 van der Waals complexes", J. Chem. Phys. **110**, 860–869 (1999).

[148]J. Williams, A. Rohrbacher, D. Djahandideh, K. C. Janda, A. Jamka, F.-M. Tao, and N. Halberstadt, "Are rare-gas Cl_2 van der Waals molecules linear or T-shaped?", Mol. Phys. **91**, 573–588 (1997).

[149]R. Prosmiti, C. Cunha, P. Villarreal, and G. Delgado-Barrio, "Ab initio ground state potential energy surfaces for Rg-Br_2 (Rg=He,Ne,Ar) complexes", J. Chem. Phys. **116**, 9249–9254 (2002).

[150]K. Patkowski, W. Cencek, P. Jankowski, K. Szalewicz, J. B. Mehl, G. Garberoglio, and A. H. Harvey, "Potential energy surface for interactions between two hydrogen molecules", J. Chem. Phys. **129**, 094304 (2008).

[151]P. Jankowski and K. Szalewicz, "A new ab initio interaction energy surface and high-resolution spectra of the H_2-CO van der Waals complex", J. Chem. Phys. **123**, 104301 (2005).

[152]H. Li, P.-N. Roy, and R. J. Le Roy, "Analytic Morse/long-range potential energy surfaces and predicted infrared spectra for CO_2-H_2", J. Chem. Phys. **132**, 214309 (2010).

[153]L. Wang, D. Xie, R. J. Le Roy, and P.-N. Roy, "A new six-dimensional potential energy surface for H_2-N_2O and its adiabatic-hindered-rotor treatment", J. Chem. Phys. **139**, 034312 (2013).

[154]D. B. Abdallah, F Najar, N. Jaidane, F. Dumouchel, and F. Lique, "Hyperfine excitation of HCN by H_2 at low temperature", Mon. Not. R. Astron. Soc. **419**, 2441–2447 (2012).

[155]O. Denis-Alpizar, Y. Kalugina, T. Stoecklin, M. H. Vera, and F. Lique, "A new abinitio potential energy surface for the collisional excitation of HCN by para- and ortho-H_2", J. Chem. Phys. **139**, 224301 (2013).

[156]K. J. Higgins, Z. Yu, and W. Klemperer, "Paper ra02", 57th Ohio State University International Symposium on Molecular Spectroscopy, Columbus, OH, June 17-21, Paper RA02 (unpublished), 2002.

[157]F. Paesani, R. E. Zillich, and K. B. Whaley, "OCS in small para-hydrogen clusters: Energetics and structure with N=1-8 complexed hydrogen molecules", J. Chem. Phys. **119**, 11682–11694 (2003).

[158]P. Valiron, M. Wernli, A. Faure, L. Wiesenfeld, C. Rist, S. Kedžuch, and J. Noga, "R12-calibrated H_2O-H_2 interaction: Full dimensional and vibrationally averaged potential energy surfaces", J. Chem. Phys. **129**, 134306 (2008).

[159]T. Zeng, H. Li, R. J. Le Roy, and P.-N. Roy, "Adiabatic-hindered-rotor treatment of the parahydrogen-water complex", J. Chem. Phys. **135**, 094304 (2011).

[160]B. Schmidt and U. Lorenz, *WavePacket 5.2.1: A program package for quantum-mechanical wavepacket propagation and time-dependent spectroscopy*,
Available via http://sourceforge.net/p/wavepacket, 2016.

[161]B. Schmidt and U. Lorenz, "WavePacket: A Matlab package for numerical quantum dynamics. I: Closed quantum systems and discrete variable representations",
https://arxiv.org/abs/1607.01714v2 (2016).

[162]D. S. Boucher, D. B. Strasfeld, R. A. Loomis, J. M. Herbert, S. E. Ray, and A. B. McCoy, "Stabilization and rovibronic spectra of the T-shaped and linear ground-state conformers of a weakly bound rare-gas homonuclear dihalogen complex: He-Br_2", J. Chem. Phys. **123**, 104312 (2005).

[163]S. E. Ray, A. B. McCoy, J. J. Glennon, J. P. Darr, E. J. Fesser, J. R. Lancaster, and R. A. Loomis, "Experimental and theoretical investigations of the He-I_2 rovibronic spectra in the I_2B-X, 20-0 region", J. Chem. Phys. **125**, 164314 (2006).

[164]Á. Valdés and R. Prosmiti, "Vibrational Calculations of Higher-Order Weakly Bound Complexes: The $He_{3,4}I_2$ Cases", J. Phys. Chem. A **119**, 12736–12741 (2015).

[165]F. Paesani and K. B. Whaley, "Vibrational shifts of OCS in mixed clusters of parahydrogen and helium", J. Chem. Phys. **124**, 234310 (2006).

[166]F. Paesani and K. B. Whaley, "Potential Energy Surface and Infrared Spectra of OCS-Hydrogen Complexes", Mol. Phys. **104**, 61–72 (2006).

[167]H. Zhu and D. Xie, "N_2O in Small para-Hydrogen Clusters: Structures and Energetics", J. Comput. Chem. **30**, 841–6 (2009).

[168]S. Moroni, M. Botti, S. D. Paolo, and A. R. W. McKellar, "Small para-Hydrogen Clusters Doped with Carbon Monoxide: Quantum Monte Carlo Simulations and Observed Infrared Spectra", J. Chem. Phys. **122**, 94314 (2005).

[169]J. Berg, B. Häcker, A. Held, J. Tymoczko, L. Stryer, C. Lange, K. Mahlke, G. Maxam, L. Seidler, and N. Zellerhoff, *Biochemie* (Spektrum Akademischer Verlag, Heidelberg, 2003).

[170]M. Fox and J. Whitesell, *Organische chemie* (Spektrum Akademischer Verlag, Heidelberg, 2004).

[171]H. Jahangiri, J. Bennett, P. Mahjoubi, K. Wilson, and S. Gu, "A review of advanced catalyst development for Fischer-Tropsch synthesis of hydrocarbons from biomass derived syn-gas", Catal. Sci. Technol. **4**, 2210–2229 (2014).

[172]R. T. Morrison and R. N. Boyd, *Lehrbuch der Organische Chemie* (Verlag Chemie, Weinheim, 1983).

[173]J. J. H. B. Sattler, J. Ruiz-Martinez, E. Santillan-Jimenez, and B. M. Weckhuysen, "Catalytic Dehydrogenation of Light Alkanes on Metals and Metal Oxides", Chemical Reviews **114**, 10613–10653 (2014).

[174]F. D. Rossini, K. S. Pitzer, R. L. Arnett, R. M. Braun, and G. C. Pimentel, *Selected values of physical and thermodynamic properties of hydrocarbons and realted compounds* (Carnegie Press, Pittsburgh, USA, 1953).

[175]M. J. Biller and S. Mecozzi, "A high level computational study of the CH_4/CF_4 dimer: how does it compare with the CH_4/CH_4 and CF_4/CF_4 dimers?", Mol. Phys. **110**, 377–387 (2012).

[176]A. H.-T. Li and S. D. Chao, "Interaction energies of dispersion-bound methane dimer from coupled cluster method at complete basis set limit", J. Mol. Struct. THEOCHEM **897**, 90–94 (2009).

[177]S. Tsuzuki, K. Honda, T. Uchimaru, and M. Mikami, "Estimated MP2 and CCSD(T) interaction energies of n-alkane dimers at the basis set limit: Comparison of the methods of Helgaker et al. and Feller", J. Chem. Phys. **124**, 114304 (2006).

[178]T. Takatani and C. David Sherrill, "Performance of spin-component-scaled Moller-Plesset theory (SCS-MP2) for potential energy curves of noncovalent interactions", Phys. Chem. Chem. Phys. **9**, 6106–6114 (2007).

[179]G. D. Smith and R. L. Jaffe, "Quantum Chemistry Study of Conformational Energies and Rotational Energy Barriers in n-Alkanes", The Journal of Physical Chemistry **100**, 18718–18724 (1996).

[180]L. L. Thomas, T. J. Christakis, and W. L. Jorgensen, "Conformation of Alkanes in the Gas Phase and Pure Liquids", J. Phys. Chem. B **110**, 21198–21204 (2006).

[181]E. Pastorczak, A. Prlj, J. F. Gonthier, and C. Corminboeuf, "Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction", J. Chem. Phys. **143**, 224107 (2015).

[182]P. Drawe, N. O. B. Lüttschwager, and M. A. Suhm, "The elastic modulus of isolated polytetrafluoroethylene filaments", ScienceOpen Research (2014) `10.14293/S2199-1006.1.SOR-MATSCI.AKA0J6.v2`.

[183]D. A. Dixon, "Torsional potential about the central carbon-carbon bond in perfluoro-n-butane", J. Chem. Phys. **96**, 3698–3701 (1992).

[184]L. Conte, A. Zaggia, A. Sassi, and R. Seraglia, "Synthesis and characterization of tri-block fluorinated-n-alkanes", J. Fluorine Chem. **128**, 493 –499 (2007).

[185]F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, "RI-MP2: optimized auxiliary basis sets and demonstration of efficiency", Chem. Phys. Lett. **294**, 143–152 (1998).

[186] *TURBOMOLE V6.5 2013, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from* `http://www.turbomole.com`.

[187] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "CP2K: atomistic simulations of condensed matter systems", Wiley Interdiscip. Rev.-Comput. Mol. Sci. **4**, 15–25 (2014).

[188] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach", Comput. Phys. Commun. **167**, 103–128 (2005).

[189] S. Lei and B. Paulus, "Incremental DF-LCCSD(T) Calculations for a Water Molecule Inside and Outside Armchair Carbon Nanotubes", Z. Phys. Chem. **230**, 651 (2016).

[190] A. Karton, D. Gruzman, and J. M. L. Martin, "Benchmark Thermochemistry of the C_nH_{2n+2} Alkane Isomers (n = 2-8) and Performance of DFT and Composite Ab Initio Methods for Dispersion-Driven Isomeric Equilibria", J. Phys. Chem. A **113**, 8434–8447 (2009).

[191] A. J. Achazi, L. K. S. von Krbek, C. A. Schalley, and B. Paulus, "Theoretical and experimental investigation of crown/ammonium complexes in solution", J. Comput. Chem. **37**, 18–24 (2016).

[192] M. Meot-Ner, "The Ionic Hydrogen Bond. 3. Multiple $NH^+ \cdots O$ and $CH^{\delta+} \cdots O$) Bonds. Complexes of Ammonium Ions with Polyethers and Crown Ethers", J. Am. Chem. Soc. **105**, 4912–4915 (1983).

[193] L. Hammerschmidt, L. Maschio, C. Müller, and B. Paulus, "Electron Correlation at the MgF2(110) Surface: A Comparison of Incremental and Local Correlation Methods", Journal of Chemical Theory and Computation **11**, 252–259 (2015).

[194] B. M. Axilrod and E. Teller, "Interaction of the van der Waals Type Between Three Atoms", J. Chem. Phys. **11**, 299–300 (1943).

[195] P. Schwerdtfeger, B. Assadollahzadeh, and A. Hermann, "Convergence of the Møller-Plesset perturbation series for the fcc lattices of neon and argon", Phys. Rev. B **82**, 205111 (2010).

[196] C. W. Bunn and E. R. Howells, "Structures of Molecules and Crystals of Fluoro-Carbons", Nature **174**, 549–551 (1954).

[197] S. S. Jang, M. Blanco, W. A. Goddard, G. Caldwell, and R. B. Ross, "The Source of Helicity in Perfluorinated N-Alkanes", Macromolecules **36**, 5331–5341 (2003).

[198] G. D. Smith, R. L. Jaffe, and D. Y. Yoon, "Conformational Characteristics of Poly(tetrafluoroethylene) Chains Based upon ab Initio Electronic Structure Calculations on Model Molecules", Macromolecules **27**, 3166–3173 (1994).

[199] J. A. Fournier, R. K. Bohn, J. A. Montgomery, Jr., and M. Onda, "Helical C_2 Structure of Perfluoropentane and the C_{2v} Structure of Perfluoropropane", J. Phys. Chem. A **114**, 1118–1122 (2010).

[200] M. Schütz, O. Masur, and D. Usvyat, "Efficient and accurate treatment of weak pairs in local CCSD(T) calculations. II. Beyond the ring approximation", J. Chem. Phys. **140**, 244107 (2014).

[201] J. Murrell and A. Jenkins, *Properties of Liquids and Solutions* (John Wiley & Sons, Chichester, England, 1997).

[202]P. L. Nostro, "Phase separation properties of fluorocarbons, hydrocarbons and their copolymers", Advances in Colloid and Interface Science **56**, 245–287 (1995).

[203]J. B. Gilmour, J. O. Zwicker, J. Katz, and R. L. Scott, "Fluorocarbon solutions at low temperatures. V. The liquid mixtures C2H6 + C2F6, C3H8 + C2F6, CH4 + C3F8, C2H6 + C3F8, C3H8 + C3F8, C4C10 + C3F8, iso-C4H10 + C3F8, C3H8 + C4F10, n-C6H14 + C4F10, n-C7H16 + C4F10, n-C,H20 + n-C4F10, and n-C10H22 + n-C4F10", J. Phys. Chem. **71**, 3259–3270 (1967).

[204]D. G. Liakos and F. Neese, "Is It Possible To Obtain Coupled Cluster Quality Energies at near Density Functional Theory Cost? Domain-Based Local Pair Natural Orbital Coupled Cluster vs Modern Density Functional Theory", J. Chem. Theo. Comp. **11**, 4054–4063 (2015).

[205]A. Wuttke and R. A. Mata, "Visualizing dispersion interactions through the use of local orbital spaces", J. Comput. Chem. **38**, 15–23 (2017).

[206]A. Nikitin, Y. Milchevskiy, and A. Lyubartsev, "AMBER-ii: New Combining Rules and Force Field for Perfluoroalkanes", J. Phys. Chem. B **119**, 14563–14573 (2015).

[207]C. Adam, L. Yang, and S. L. Cockroft, "Partitioning Solvophobic and Dispersion Forces in Alkyl and Perfluoroalkyl Cohesion", Angew. Chem. Int. Ed. **54**, 1164–1167 (2015).

Acknowledgements

I want to deeply thank my supervisor Prof. Dr. Beate Paulus, for her guidance, many discussions, for being always supportive, for her patience and for organizing her group in such a way, that one really enjoys working there. Especially, I want to thank her for giving me the opportunity to become a member of the research training group "Fluorine as a Key Element", which has helped me tremendously in my development. PD Dr. Burkhard Schmidt supervised much of my work on the van der Waals dimers and skilfully guided me through this part of the research, I thank him a lot for that. This part of the work is directly linked to the work of Dr. Antonio Accardi. It was his nicely written PhD thesis, that facilitated my adoption of the methods. Many thanks go to Dr. Carsten Müller for many helpful comments and ideas regarding my work and for organizing the activities of the research training group, which let to many fruitful discussions with the guests and the other members. Furthermore, I thank Dr. Pouya Partovi for his interest in my work, which lead to a mutual collaboration. I like to thank Andreas Achazi, for sharing his experience on DFT calculations with Turbomole with me. Likewise, I like to thank Dr. Edoardo Fertitta, with whom I had many discussions about the method of increments and many other topics. Furthermore I like to thank PD Dr. Denis Usvyat, Prof. Dr. Georg Jansen, Prof. Dr. David Dixon and Dr. Johannes Floß for brief but very helpful discussions. Many thanks go to all the past and present members of the theory groups (AG Paulus, AG Manz, AG Leibscher, AG Tremblay, AG Keller, AG Kaghazchi and AG Bande), that I got to know during my years at our institute. In that regard, I want to acknowledge the excellent teaching skills of PD Dr. Monika Leibscher, Dr. Thomas Grohmann, PD Dr. Dirk Andrae, Dr. Axel Schild and Dr. Timm Bredtmann. I'm also happy that I could meet so many new friends: Edoardo Fertitta, Francesca Vitalini, Lidice Cruz, Stevan Aleksić, Jhon Frédy Perez Torres, Sara Panahian Jand, Pouya Partovi-Azar and many more. I like to especially thank Francesca for bringing together so many people and inventing the "healthy Wednesday". I further like to thank my friends Tim Baumert and Johannes Floß for always staying in touch. I acknowledge financial support through the DFG research training group 1582 "Fluorine as a Key Element". The HPC services of the ZEDAT are thanked for providing valuable computational resources.

Finally, I like to thank my parents and my brother as well as my partner and best friend Barbara Mlody.

Publications

- in relation to this thesis:
  - Matthias Berg, Antonio Accardi, Beate Paulus, and Burkhard Schmidt, "Rotationally adiabatic pair interactions of para- and ortho-hydrogen with the halogen molecules F_2, Cl_2, and Br_2", *Journal of Chemical Physics*, **2014**, 141, 074303. http://dx.doi.org/10.1063/1.4892599

- others:
  - Matthias Berg, Beate Paulus, and Timm Bredtmann, "Electronic quantum fluxes in vibrating symmetric and polar single, double and triple bonds". *Molecular Physics*, **2016**, 114, 1356-1364. http://dx.doi.org/10.1080/00268976.2015.1122843
  - Pouya Partovi-Azar, Matthias Berg, Simone Sanna, and Thomas D. Kühne, "Improved parameterization of the quantum harmonic oscillator model based on localized Wannier functions to describe van der Waals interactions in density functional theory.", *International Journal of Quantum Chemistry*, **2016**, 116, 1160-1165. http://dx.doi.org/10.1002/qua.25150
  - Steffen Belz, Shmuel Zilberg, Matthias Berg, Thomas Grohmann and Monika Leibscher. "Pyridinylidene-phenoxide in strong electric fields: controlling orientation, conical intersection, and radiation-less decay", *The Journal of Physical Chemistry A*, **2012**, 116, 11189-11198. http://dx.doi.org/10.1021/jp305090b,

Curriculum vitæ

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.
For reasons of information privacy, the curriculum vitae is not included in the online version of this document.