

# Computational Study of Dispersion Interactions through Local Orbital Analysis

Dissertation

for the award of the degree

"Doctor rerum naturalium" (Dr.rer.nat.)

of the Georg-August-Universität Göttingen

within the doctoral program Chemistry

of the Georg-August University School of Science (GAUSS)

submitted by

**Axel Wuttke**

from Hameln

Göttingen, 2018

Thesis Committee

Prof. Dr. Ricardo Mata

Institute for Physical Chemistry,  
Georg-August University Göttingen

Prof. Dr. Guido Clever

Institute for Inorganic Chemistry,  
Technical University of Dortmund

Members of the Examination Board

Reviewer:

Prof. Dr. Ricardo Mata

Institute for Physical Chemistry,  
Georg-August University Göttingen

Second Reviewer:

Prof. Dr. Guido Clever

Institute for Inorganic Chemistry,  
Technical University of Dortmund

Further members of the Examination Board

Prof. Dr. Thomas Waitz

Institute for Inorganic Chemistry,  
Georg-August University Göttingen

Dr. Sebastian Kruss

Institute for Physical Chemistry,  
Georg-August University Göttingen

Prof. Dr. Jörg Behler

Institute for Physical Chemistry,  
Georg-August University Göttingen

PD Dr. Thomas Zeuch

Institute for Physical Chemistry,  
Georg-August University Göttingen

Date of the oral examination:

25.01.2019



# Acknowledgements

First of all, I would like to thank Prof. Dr. Ricardo Mata for the excellent supervision during my PhD thesis. Since my Master's thesis in 2014 I have been part of the Computational Chemistry and Biochemistry (CCB) Group. Especially in the first years, I was excellently instructed and got to know a variety of calculation methods and computer applications in the field of theoretical chemistry. In the last phase of my doctoral studies, I got time and freedom to further explore my personal interests. This mix of guidance and independent work was superb for my personal development.

Furthermore, I would like to thank Prof. Dr. Guido Clever and his group for the great cooperation within the SPP1807. The provided supramolecular host-guest complexes were very difficult for theoretical investigations because of their size and complexity. In the end, however, we were able to find solutions together and publish three reports. I am very grateful for that.

All previous and current CCB group members should be also mentioned here. The topics are very diversified in our group and you could always learn something new at the Tuesday seminars. Special thanks go to Mirko Paulikat, who read parts of my dissertation and was always up for game, fun and closing time beer. Also Karin Schoppe deserves special thanks for the daily care of my dogs. Not only did I save time but also many feed costs.

Last but not least, my girlfriend Lydia, who supported me during the whole time, should not miss. Especially in the last few weeks of writing, she was a real support and foresaw every of my wishes. At this point I could pronounce my acknowledgments to many more people, but I will personally catch up for this.



# Abstract

Non-covalent interactions are responsible for molecular aggregation and thus indispensable in the study of condensed matter. In chemical research, such information plays an important role for example in creating new drugs with tailored properties through rational design. The tools of computational chemistry are invaluable to understand such events. Two popular methods are electrostatic potential maps (EPM) and noncovalent interactions (NCI) plots. They are both easy in application and give an intuitively interpretable graphical representation. However, in order to quantify and consistently study intermolecular interactions, higher level methods are required. In this thesis, an extension to local correlation methods that allows to visualize dispersion interactions is presented. The so-called Dispersion Interaction Density (DID) is obtained by scaling the local molecular orbitals (LMOs) closed-shell densities with their corresponding contributions to the dispersion interaction. From the DID-matrix a 3-dimensional grid can be calculated and afterwards visualized in the form of contour plots, color projections on the electronic density or Voxel graphics. The latter proves to be the method with the greatest breadth of information. Studies on selected systems such as the well-known benzene dimer demonstrate their usefulness. Furthermore, it is also shown how intramolecular effects can be easily investigated by means of local orbital analysis. Having exhibited that the latter in conjunction with DIDs are well suited to study such effects, more complex structures have been examined.

The Clever group has made important advances in metal mediated self assembly of supramolecular compounds. By different substitutions on a bis-monodentate pyridyl backbone two supramolecular structures which can serve as a host system for different guest molecules were obtained. The first host examined has adamantyl residues showing in the cages interior and is able to bind different ionic guest molecules. Local orbital analysis revealed that the main driving forces are given by dispersion and Pauli repulsion. Motivated by these results the Clever group is currently extending the scope of dispersion energy donors (DEDs) that can be implemented

into the ligand backbones. The second host system discussed is an interpenetrated double cage structure. It is able to take up various neutral guest molecules in its middle pocket. Local correlation calculations revealed that this is mainly achieved by dispersion interactions. Further, a dependence between the dispersion contributions and experimental free binding energies was found. In order to build a bridge between theory and experiment a protocol to compute free binding energies was applied. For small guest molecules the scheme performed rather well while it could be shown that larger guest molecules require a relaxation of the pockets geometry. For this purpose three more sophisticated computing protocols are proposed.

The study of non-covalent interactions is continued by the investigation of metallophilic contacts. The previous work of our group was extended by the investigation of experimental crystal structures and mixed metallophilic contacts. Through local orbital analysis, fragments were built and the interaction energies were decomposed. The outcome revealed that there are good reasons to highlight the strength of metallophilic contacts. However, mistaking such an interaction for the main driving force in molecular aggregation should be avoided. The calculations carried out for the  $\text{ClAuR}_2\text{bimy}$  ( $\text{R}=\text{Me}, \text{Et}$ ) complexes show that simple changes in the ligands can lead to crystal structures where even no aurophilic contacts are observed. The  $d^{10}$  cations can contribute in stabilizing a molecular crystal, but the ligand composition seems to be the dominating factor.

In all previous calculations the DF-PAO-LMP2 code implemented in Molpro was used. Especially regarding the supramolecular host-guest structures it was difficult to make the calculations possible at all. This was at least achieved by applying approximations such as multipoles for the calculation of distant orbital pairs. Nevertheless, the calculations were very time intensive. By rediscovering PNO as virtual space the groups of Neese and Werner made great progress in the recent years. The computational cost of the latest generation local correlation methods scales linear with regard to the molecular size. Combined with massive parallelization over many processors large molecular clusters can nowadays be computed within minutes. To benefit from and inspired by recent development in local orbital analysis,<sup>[1]</sup> an EDA scheme for PNO-LMP2 has been implemented in Molpro. It is shown that strict spatial localization of the PNOs and a reliable population analysis are necessary to obtain stable results, especially for the charge transfer terms. For the latter, a visualization related to DIDs is probed. The improvement of the EDAs stability is, however, also purchased with an increased computational cost of the method. In the future, calculations of large molecular systems must show whether a temporal advantage is obtained in comparison to the initial PAO-based methods.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theoretical Background</b>	<b>7</b>
2.1 Hartree-Fock – Independent Particle Model . . . . .	9
2.2 Electron Correlation Methods . . . . .	10
2.2.1 Møller-Plesset Perturbation Theory . . . . .	11
2.2.2 Coupled-Cluster Theory . . . . .	13
2.3 Local Correlation Methods . . . . .	14
2.3.1 Orbital Invariant Formulation . . . . .	14
2.3.2 Orbital Localization . . . . .	16
2.3.3 Virtual Space Definition . . . . .	17
2.4 Population Analysis . . . . .	19
2.5 Energy Decomposition Analysis (EDA) . . . . .	21
<b>3 Local orbital analysis</b>	<b>27</b>
3.1 Fragmentation Using Local Orbital Subspaces . . . . .	27
3.2 Dispersion Interaction Density (DID) . . . . .	28
3.2.1 Method . . . . .	30
3.2.2 Visualization of DID . . . . .	32
<b>4 Applications of Dispersion Interaction Densities</b>	<b>35</b>
4.1 Computational Details . . . . .	36
4.2 Benzene Dimer . . . . .	36
4.3 Substituted Benzene Dimer . . . . .	40
4.4 Methylated Anisole-Methanol Dimers . . . . .	43
4.5 Coupled Diamondoids . . . . .	47

4.6	Conclusions . . . . .	49
<b>5</b>	<b>Supramolecular Coordination Cages</b>	<b>51</b>
5.1	A Sterically Overcrowded Cage . . . . .	53
5.1.1	Summary of Experimental Findings . . . . .	54
5.1.2	Theoretical Investigations . . . . .	55
5.2	An Interpenetrated Double-Cage . . . . .	58
5.2.1	Summary of Experimental Findings . . . . .	58
5.2.2	Theoretical Investigations . . . . .	61
5.2.3	Outlook . . . . .	65
5.3	Conclusions . . . . .	67
<b>6</b>	<b>Metallophilic interactions</b>	<b>69</b>
6.1	Computational Details . . . . .	71
6.2	Results and Discussion . . . . .	72
6.2.1	Cl-Au-Me <sub>2</sub> bimy . . . . .	72
6.2.2	Cl-Au-Et <sub>2</sub> bimy . . . . .	77
6.2.3	AuHg <sub>2</sub> (o-C <sub>6</sub> F <sub>4</sub> ) <sub>3</sub> – Hg <sub>3</sub> (o-C <sub>6</sub> F <sub>4</sub> ) <sub>3</sub> . . . . .	80
6.3	Conclusions . . . . .	84
<b>7</b>	<b>Energy Decomposition Analysis for PNO-LMP2</b>	<b>85</b>
7.1	Implementation . . . . .	86
7.2	Results for Intermolecular Interactions . . . . .	90
7.3	Time Scaling . . . . .	99
7.4	Conclusions . . . . .	100
<b>8</b>	<b>Summary</b>	<b>103</b>
	<b>List of Abbreviations</b>	<b>109</b>
	<b>Bibliography</b>	<b>111</b>

# 1 Introduction

In the course of chemistry undergraduate studies, each student learns to know different models of chemical bonding. The rough classification is usually provided as metallic, ionic, covalent and non-covalent bonds. The covalent bond is determined by shared electrons between atoms that form a molecule.<sup>[2]</sup> Ionic bonds, for example in salt crystals, arise from electrostatic interactions<sup>[3]</sup> and metals are described by a delocalized electron gas.<sup>[4]</sup> The non-covalent bonds are often limited to intermolecular effects in the literature, which is insufficient, as will be shown later in this thesis.

One of the best known examples of non-covalent interactions is the hydrogen bond.<sup>[5]</sup> Due to the strong difference in the electronegativity between hydrogen and, for example, oxygen or nitrogen, strongly polarized regions are formed. The latter build strong electrostatic interacting bridges which explain the relatively high boiling point of water and ammonia.<sup>[6]</sup> The effect of bond polarization always occurs when there is a significant difference in the electronegativity of covalently bonded atoms.<sup>[7]</sup> The latter leads to a dipole which can then interact electrostatically with the dipole of another molecule.<sup>[8]</sup> Furthermore, higher order multipoles are possible, e.g. the quadrupole moment of benzene.<sup>[9]</sup> The  $\pi$ -system results in a negatively charged region above and below the ring surface while the ring surface itself is positively charged. Based on this, the T-shaped benzene conformation<sup>[10,11]</sup> or the cation benzene interaction<sup>[12]</sup> can be explained. Turning to non-polar molecules such as methane, or the noble gases, their dimerization cannot be described by electrostatic interactions. Their intermolecular bonds are established by dispersion forces. The latter arise from the non-uniform distribution of the electrons around the atoms. As a result, there is a charge imbalance that induces instantaneous dipoles in a nearby molecule (or atom in case of the noble gases).<sup>[13]</sup> Thereby, a weak attraction arises. Dispersion interactions are always present when atoms or molecules approach each other. However, when strong dipoles or multipoles are present, they are far surpassed

by electrostatic interactions. This feature led to a neglect of the importance of these interactions in theory for a long time.

In the Hartree-Fock method,<sup>[14,15]</sup> dispersion interactions are not included by the mean field approach. The latter does not allow for the description of instantaneous dipoles. By performing a post Hartree-Fock method such as Møller-Plesset perturbation theory (MPPT),<sup>[16]</sup> Coupled-Cluster (CC)<sup>[17,18]</sup> or Configuration Interaction (CI)<sup>[19]</sup> these energy terms are recovered. In the late seventies, however, these methods were far too costly for the existing computers. This led Ahlrichs in 1977 to introduce the HFD, Hartree-Fock Dispersion method.<sup>[20]</sup> Thereby, the dispersion was fitted with an empirical functional using C6 and C8 coefficients for the interactions of specific elements. Although this approach initially did not have much use, it was taken up again years later in the density-functional theory (DFT).<sup>[21,22]</sup> Due to the development of generalized gradient approximation (GGA) functionals<sup>[23]</sup> and their comparatively low demands on the computing power it became possible to calculate medium sized molecular systems in a reasonable time. However, Pulay reported in 1994 that the conventional DFT functionals are not able to describe rare gas dimers.<sup>[24]</sup> For a few more years, these problems continued to be neglected until in 2000 further inconsistencies in the calculation of thermodynamic parameters<sup>[25]</sup> and in 2004 errors for isodesmic reactions<sup>[26]</sup> (DFT is for example unable to correctly reflect the energetic order of the butane conformers) were reported. In the same year Stefan Grimme published the first empirical dispersion correction for DFT, DFT-D.<sup>[27]</sup> Over the years the D correction has been steadily developed and is now available in the 4th generation.<sup>[28]</sup> DFT calculations without dispersion corrections have become uncommon today and illustrate the importance of these interactions.

The dispersion contributions obtained by a DFT-D calculation may serve as a qualitative estimate, but should not be used as quantitative results, as unfortunately often done. Since DFT functionals already implicitly capture dispersion effects at short range to a certain extent, the dispersion contributions obtained by the D correction depend strongly on the functional used. On the basis of wave-function methods, there are different approaches to obtain accurate and consistent values. One of the most popular methods is Symmetry-Adapted Perturbation Theory (SAPT).<sup>[29]</sup> Over the last years many approaches and extensions have been published.<sup>[30]</sup> The Sherrill group developed schemes to partition molecular systems atomically and fragment-wise using local orbital spaces.<sup>[31,32]</sup> This also allows the investigation of weak intramolecular interactions,<sup>[33]</sup> which can influence e.g. catalysis or the molecular



---

conformation. In addition, visualization techniques of the interaction contributions were presented. Often a picture is worth a thousand words and facilitates the interpretation of computed results. The big disadvantage of SAPT methods is their computational effort.

Another option to determine dispersion interactions is provided by the local versions of the conventional post Hartree-Fock procedures, MPPT and CC. Their computational effort is drastically reduced compared to the initial canonical methods and by far lower than SAPT approaches. In this development process especially the pioneering work of Pulay should be mentioned, who presented first the definition of the virtual space by projected atomic orbitals.<sup>[34]</sup> The latest generation of local correlation methods uses pair natural orbitals (PNOs), originally proposed by Mayer,<sup>[35,36]</sup> as virtual space.<sup>[37-46]</sup> These methods scale linearly with the molecular size, as well as with the processors used. By massive parallelization on large computational clusters, molecular systems containing hundreds of atoms can be calculated with quantum mechanical accuracy in short time. The advantages and possibilities of local correlation methods have been demonstrated in many publications. By simple modifications the interaction of specific fragments or orbital spaces can be analyzed,<sup>[47-53]</sup> and the ad hoc definition of PAOs allows a simple and computationally cost-effective determination of dispersion interactions.<sup>[47]</sup>

In this thesis the aforementioned possibilities of local correlation methods are used for the theoretical investigation of different molecular systems, ranging from the benzene dimer up to large supramolecular clusters. The theoretical basics are discussed in detail in Chapter 2. Wave-function based methods are first presented in a fundamental way and more in-depth discussion on the use of local orbital spaces. Since the thesis often focuses on dispersion interactions, their determination using different energy decomposition analyses is evaluated and local correlation methods are compared with other approaches, such as SAPT.

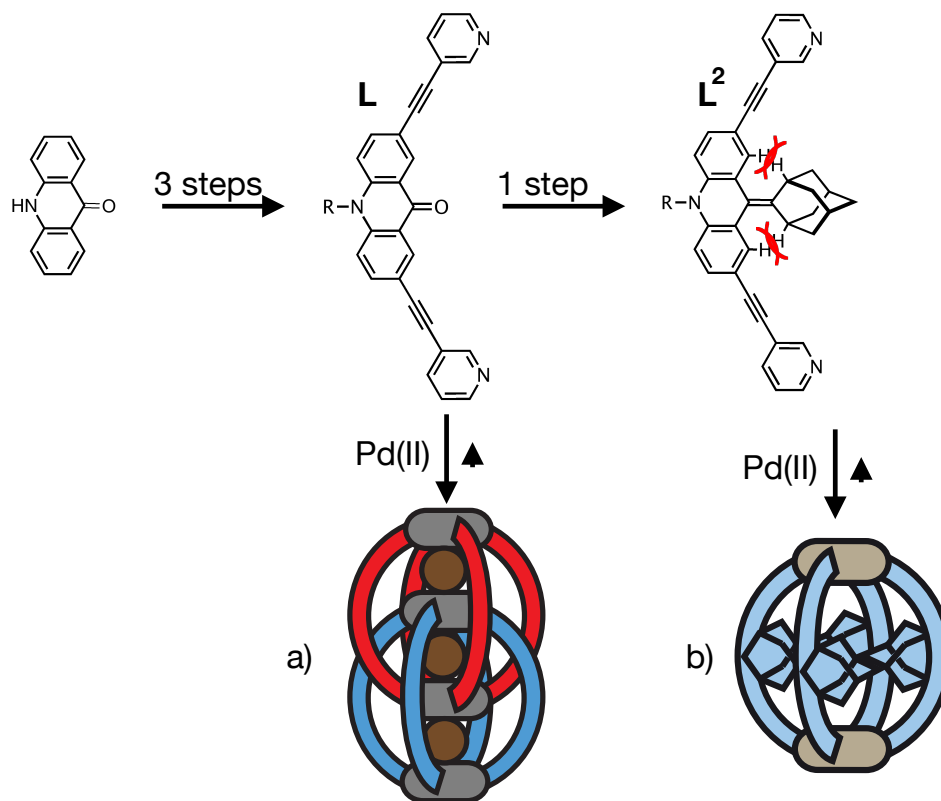
Chapter 3 includes techniques in the context of local orbital analysis. A natural population analysis based robust fragmentation scheme is presented. It provides the possibility to divide molecules into groups of atoms and therefore, permits an analysis of the interaction of clearly defined areas. This allowed, among other tasks, the study of aurophilic<sup>[51,54]</sup> and intramolecular interactions<sup>[55,56]</sup> in our group. Furthermore, the dispersion interaction density (DID) is introduced.<sup>[55]</sup> It can be deduced from the closed-shell densities of orbitals and their contribution to the dispersion

energy. Based on the DID, three dimensional grids can be generated. The latter can be visualized as contour plots, color projections on the electronic density or Voxel graphics by suitable software.

Chapter 4 features an analysis of the DID plots. Thereby, the benzene dimer, substituted benzene dimers, methylated anisole methanol complexes and a coupled diamondoid are studied. The different visualization techniques presented in Chapter 3 are compared. It is discussed whether the plots can facilitate the interpretation of the calculated results or not. Furthermore, the possibility to compare spin component scaled local second order Møller-Plesset (SCS-LMP2)<sup>[57]</sup> calculations with highly accurate CC and CI results at the complete basis limit is taken for the benzene dimer. In the case of the substituted benzene dimers, the SCS-LMP2 dispersion is compared to results obtained by fragment-based partitioning within the F-SAPT approach.<sup>[58]</sup> A coupled diamondoid system calculation illustrates how easily the energy decomposition scheme can be applied to intramolecular interactions.

Chapter 5 discusses the results of the cooperation with the Clever Group within the SPP1807.<sup>[59-61]</sup> Through the substitution of two 3-ethynylpyridines to an acridone they were able to synthesize the banana shaped ligand **L** shown in Figure 1.1 a). By a self-assembly process in the presence of  $[\text{Pd}(\text{CH}_3\text{CN})_4]-(\text{BF}_4)_2$  an interpenetrated double cage with three pockets filled by  $\text{BF}_4^-$  ions is obtained. This system can be activated through the exchange of the outer pocket ions by  $\text{Cl}^-$  or  $\text{Br}^-$  to bind neutral guest molecules in the middle pocket. A further modification of **L** by the substitution of adamantyl groups in the carbonyl position, lead to a sterically overcrowded coordination cage as shown in Figure 5.1 b). Different ionic guest molecules are able to bind in its pocket and it is of interest if dispersion has a determining role, since one would suspect that electrostatics are the dominant driving force. Local correlation calculations are carried out to determine the strength of dispersion interactions in the pocket regions of both systems. In the case of the interpenetrated double cage, an approach for the determination of free binding energies is presented.

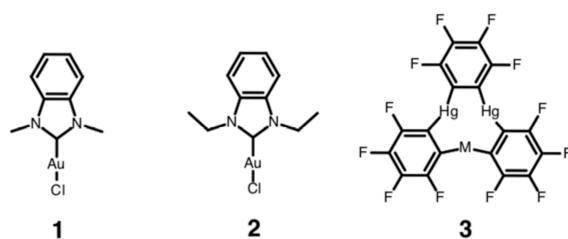
Chapter 6 continues the previous work of our group on metallophilic interactions.<sup>[51]</sup> Thereby, the relative weight of such interactions was studied by the crystal structure of two closely related complexes being the  $[\text{Au}(\text{Me}_2\text{bimy})\text{Cl}]$  and  $[\text{Au}(\text{Et}_2\text{bimy})\text{Cl}]$  complexes (**1** and **2** in Figure 1.2)(bimy = benzimidazol-2-ylidene).<sup>[62]</sup> Furthermore, Au(I)-Hg(II) complexes aroused interest as examples of mixed metallophilic interac-



**Figure 1.1:** Supramolecular Complexes derived through the substitution of acridone.

tions. It has been stated that the forces between those metals are much stronger than the aurophilic ones.<sup>[63,64]</sup> To investigate the effect a trinuclear complex  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  containing one short  $\text{Au(I)} \cdots \text{Hg(II)}$  contact (**3** see Fig. 1.2, M can be either Au or Hg) is considered.<sup>[64]</sup> The major question is how the competing forces, such as  $\pi$ - $\pi$  or Au-Au, contribute to the final crystal structures. For this purpose, use is made of a fragment orbital analysis of the electron correlation energy.

Chapter 7 introduces an energy decomposition analysis in the framework of PNO-LMP2. In comparison to PAO-LMP2 the spatial location of PNOs is not inherently known. To correct this issue a sophisticated localization procedure is presented. A similar scheme has already been reported by the Neese group for DLPNO-CCSD(T),<sup>[1]</sup> which, however, lead to unstable charge transfer energies. It is shown that a strong spatial restriction of the PNOs and their local determination by a reliable population analysis are necessary to obtain a robust scheme. To analyze the scheme, the S22 benchmark set of non-covalent interactions<sup>[65]</sup> served as a guideline. Furthermore, a visualization technique similar to DIDs is presented for the charge transfer interactions.



**Figure 1.2:** Lewis structures of **1** [Au(Me<sub>2</sub>bimy)Cl] **2** [Au(Et<sub>2</sub>bimy)Cl] and **3** MHg<sub>2</sub>(o-C<sub>6</sub>F<sub>4</sub>)<sub>3</sub> where M can be either Au or Hg. Reproduced from Reference [54] with the permission of the American Chemical Society.

## 2 Theoretical Background

It is well-known that non-covalent interactions play a decisive role in molecular aggregation. Despite their variety, the discussion is often dominated by electrostatic effects. However, several publications revealed that London dispersion forces are essential to properly describe such event.<sup>[66–69]</sup> Targeting molecular interactions is of huge interest for a broad variety of systems, from small sized dimers up to large supramolecular clusters. For this reason different demands on the methods are made. Supramolecular host-guest complexes contain up to hundreds of atoms. Accordingly, the method of choice has to be computationally manageable at such sizes. Molecular scales provide different binding sites which often only differ by less than a  $\text{kJ}\cdot\text{mol}^{-1}$ .<sup>[70,71]</sup> Therefore, a highly accurate computational treatment is required. Finding the balance between computational effort and accuracy is an important task. In this Chapter the fundamentals to derive the physical nature of weak interactions via wave-function theory are discussed.

Wave-function theory in general builds upon solving the Schrödinger equation<sup>[72]</sup>

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

obtaining the energy  $E$  of a system with its associated wave-function  $\Psi$ . The work featured in this thesis will deal with time-independent problems, accordingly  $\Psi$  is only a function of spatial and spin coordinates with its corresponding Hamiltonian. The equation reads as an eigenvalue problem, in which the wave-function  $\Psi$  is an eigenfunction of the Hamiltonian  $\hat{H}$  with the corresponding eigenvalue  $E$ . The non-relativistic Hamiltonian in atomic units for a system containing  $n$  electrons and  $M$

## 2 Theoretical Background

---

nuclei is given by

$$\begin{aligned}
 \hat{H} = & \underbrace{-\frac{1}{2} \sum_{i=1}^n \nabla_i^2}_{T_e} - \underbrace{\frac{1}{2} \sum_{a=1}^M \frac{1}{m_a} \nabla_a^2}_{T_n} + \underbrace{\sum_{i=1}^n \sum_{j>i}^n \frac{1}{r_{ij}}}_{V_{ee}} \\
 & + \underbrace{\sum_{a=1}^M \sum_{b>a}^M \frac{Z_a Z_b}{r_{ab}}}_{V_{nn}} - \underbrace{\sum_{a=1}^M \sum_{i=1}^n \frac{Z_a}{r_{ia}}}_{V_{ne}}
 \end{aligned} \tag{2.2}$$

with  $i$  and  $j$  as electron indices and  $a$  and  $b$  for the nuclei, respectively.  $m_a$  and  $Z_b$  denote the mass and charge of a nuclei. The distance between two particles is declared by  $r_{xy}$ . The energy terms are split up to kinetic contributions  $T_e, T_n$  and repulsive potentials  $V_{ee}, V_{nn}$  for electrons and nuclei, respectively. The last term  $V_{ne}$  accounts for the attractive potential between electrons and nuclei.

The application of this Hamiltonian is very limited, since it can be only solved analytically for two particle systems. In order to reduce the complexity, the Born-Oppenheimer approximation<sup>[73]</sup> is a powerful approach. Hereby, the coupling between the motion of electrons and nuclei is neglected. This is explained by the huge mass difference between those particles and the resulting higher velocity of electrons. Therefore, the Schrödinger equation is split into an electronic and a nuclear part, whereby the electronic Schrödinger equation

$$\hat{H}_{\text{el}} \Psi = E_{\text{el}} \Psi \tag{2.3}$$

is solved for a fixed nuclei configuration. The electronic Hamiltonian is given by

$$\hat{H}_{\text{el}} = -\frac{1}{2} \sum_{i=1}^n \nabla_i^2 - \sum_{a=1}^M \sum_{i=1}^n \frac{Z_a}{r_{ia}} + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{r_{ij}} . \tag{2.4}$$

The nuclear repulsion potential is added *a posteriori* to the electronic energy. In the following Sections different methods to compute the electronic Schrödinger equation are discussed.

## 2.1 Hartree-Fock – Independent Particle Model

Within the Hartree-Fock method<sup>[14,15]</sup> the wave-function  $\Psi$  is treated by a single Slater determinant.<sup>[74]</sup> Through this definition the resulting  $\Psi^{\text{HF}}$  is antisymmetric by construction and fulfills the Pauli principle.<sup>[75]</sup> Since in this thesis only closed-shell systems are under investigation, the electron spin can be integrated out and doubly occupied spatial orbitals  $\phi_i$  are obtained. They only depend on the location  $\mathbf{r}_i$ . The corresponding Slater determinant for an  $2n$  electron system is determined by

$$\Psi^{\text{HF}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \dots & \phi_n(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \phi_1(\mathbf{r}_n) & \dots & \dots & \phi_n(\mathbf{r}_n) \end{vmatrix} \quad (2.5)$$

where all  $n$  orbitals are doubly occupied. This treatment is commonly referred to as restricted Hartree-Fock (RHF). The Hartree-Fock procedure is by construction variational and orbital invariant relative to unitary transformations within the occupied space.<sup>[76]</sup> Correspondingly, the best energy obtained with the method is always an upper bound to the exact solution. Further the spatial orbitals are determined to be orthonormal to each other. The energy is then evaluated through a set of pseudo eigenvalue equations

$$\hat{f} |\phi_i\rangle = \epsilon_i |\phi_i\rangle \quad (2.6)$$

with  $\hat{f}$  being the Fock operator, which is defined as

$$\hat{f}(i) = \hat{h}(i) + \sum_{j=1}^{n/2} 2\hat{j}_j(i) - \hat{k}_j(i) = \hat{h}(i) + \hat{g}(i) . \quad (2.7)$$

Hereby,  $\hat{j}$  and  $\hat{k}$  denote the Coulomb and exchange operators, respectively. The derived orbitals  $\phi_i$  are the so called canonical orbitals with their corresponding orbital energies  $\epsilon_i$ . By the introduction of the Fock operator, the exact electron-electron operator in Equation 2.4 has been replaced by the operator  $g(i)$ . The latter treats an electron in the averaged field of all other electrons, thereby deriving the name independent particle model.

To represent the orbitals and finally compute the Hartree-Fock equations, the linear combination of atomic orbitals (LCAO) approach<sup>[77]</sup> is applied. Under this approximation a set of basis functions is defined for each atom. An orbital  $\phi_i$  is then

constructed by

$$\phi_i = \sum_{\mu=1}^{N_{\text{bas}}} C_{\mu i} \chi_{\mu}(\mathbf{r}) . \quad (2.8)$$

Here,  $N_{\text{bas}}$  is the total number of basis functions,  $\chi_{\mu}$  is an atom centred basis function.  $C_{\mu i}$  is the coefficient of basis function  $\chi_{\mu}$  for orbital  $i$ . If the basis set size is approaching infinity, the so called Hartree-Fock limit is reached and the correct wave-function under the Hartree-Fock methodology is derived. However, taking infinite basis sets is computationally not feasible and the number of basis functions is restricted to a certain size. Incorporating the LCAO approach to the canonical Hartree-Fock formalism leads to the so called Roothaan-Hall equations, which can then be expressed in matrix notation

$$\mathbf{FC} = \mathbf{SCE} . \quad (2.9)$$

Here  $\mathbf{F}$  is the Fock-Matrix,  $\mathbf{C}$  the orbital coefficient matrix and  $\mathbf{S}$  the overlap matrix between atomic basis functions. To solve the equation the orbital coefficients are varied to obtain the lowest energy. Through the variational ansatz it is an upper bound to the exact energy. If two electrons get close together they are avoiding each other, thereby reducing the repulsive interactions. This effect is not captured by the Hartree-Fock method. In the methodology of wave-function based methods this type of interactions is described by the electron correlation energy, which will be discussed in the following Section.

## 2.2 Electron Correlation Methods

In this Section two approaches to correct Hartree-Fock for the missing electron correlation are discussed, namely Møller-Plesset perturbation theory (MPPT) and Coupled-Cluster (CC) theory. As it will be discussed in Section 2.5 both methods give the possibility to quantify dispersion interactions. Here a brief overview of both approaches is given.



### 2.2.1 Møller-Plesset Perturbation Theory

The Møller-Plesset formalism<sup>[16]</sup> is a special case of the Rayleigh-Schrödinger perturbation theory using the Hartree-Fock wave-function as the reference.<sup>[19]</sup> Thereby, a perturbation is added to the unperturbed Hamiltonian, as

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{V} \quad (2.10)$$

with  $\hat{H}^{(0)}$  being the unperturbed Hamiltonian and therefore, the sum over one electron Fock operators. The perturbation operator  $\hat{V}$  accounts for the electron-electron repulsion with  $\lambda$  as a scaling parameter defining its strength. The wave-function  $\Psi$  and energy expressions are expanded in a series of Taylor expansions

$$\Psi = \Psi^{(0)} + \lambda \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \dots \quad (2.11)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \quad (2.12)$$

By substituting  $\hat{H}$ ,  $\Psi$  and  $E$  in the electronic Schrödinger equation and subsequently ordering in terms of the power of  $\lambda$  one obtains the Møller-Plesset equations with  $n$  being the order of the perturbation (MP $n$ ).

$$\begin{aligned} \lambda^0 : \quad & \hat{H}^{(0)} \Psi^{(0)} = E^{(0)} \Psi^{(0)} \\ \lambda^1 : \quad & \hat{H}^{(0)} \Psi^{(1)} + \hat{V} \Psi^{(0)} = E^{(0)} \Psi^{(1)} + E^{(1)} \Psi^{(0)} \\ \lambda^2 : \quad & \hat{H}^{(0)} \Psi^{(2)} + \hat{V} \Psi^{(1)} = E^{(0)} \Psi^{(2)} + E^{(1)} \Psi^{(1)} + E^{(2)} \Psi^{(0)} \\ & \vdots \\ \lambda^n : \quad & \hat{H}^{(0)} \Psi^{(n)} + \hat{V} \Psi^{(n-1)} = \sum_{i=0}^n E^{(i)} \Psi^{(n-i)} \end{aligned}$$

In this series of Møller-Plesset expansions the most commonly and also in this thesis used method is the second-order MP2 procedure. Hereby, the zeroth- the first- and the second-order terms are summed up. The zeroth-order term is derived with the HF determinant and therefore, the zeroth-order energy is the sum of the occupied orbital energies. The first-order term can be understood as the mean-field correction, which describes one electron in the averaged field of all electrons.<sup>[76]</sup> The sum of the zeroth- and first-order energies exactly yields the Hartree-Fock result. Conversely this means, the second-order term gives the first correction caused by electron correlation to the Hartree-Fock result.

## 2 Theoretical Background

---

For that, the first-order wave function is built by excitations from occupied orbitals  $i, j$  to virtual orbitals  $a, b$ . The couplings between singly excited determinants and the reference wave-function vanishes due to Brillouins-Theorem and the orthonormality conditions between the orbitals. Therefore, the second-order energy is derived by the coupling between the Hartree-Fock reference and doubly excited states. The closed-shell second-order perturbation energy is finally expressed by

$$E_{\text{MP2}} = \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} \frac{(2 \langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle)}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \quad (2.13)$$

The canonical MP2 methods computational cost formally scales with the power  $N_{\text{bas}}^5$  and is therefore, computationally cheap if compared to Coupled-Cluster or Configuration Interaction methods including double excitations.<sup>[76]</sup> The method is by construction size extensive, which made it popular to use for weakly interacting dimer systems. However, there are also some disadvantages when using Møller-Plesset perturbation theory (MPPT). The procedure is not variational, so that the calculated energy can be lower than the exact energy. Moreover, poor HF wave-functions lead to larger perturbation corrections which can end up in poor convergence patterns. Another problem arises if systems with a significant multi reference character are considered, since the theory is built upon single Slater determinants. For this work it is important to mention that molecular systems including larger dispersion and/or charge transfer interactions tend to be overestimated at the MP2 level. However, this problem can be reduced by the spin-component scaling (SCS) approximation.<sup>[57]</sup> Hereby, one distinguishes between anti-parallel (triplet) and parallel (singlet) spin-electron pairs. It was proved that the triplets are commonly overestimated, while the singlets tend to be underestimated. Through a set of benchmark calculations Grimme postulated to compute the correlation energy as

$$E_{\text{corr}} = p_s E_s^{(2)} + p_t E_t^{(2)} \quad (2.14)$$

with  $p_s = \frac{6}{5}$  and  $p_t = \frac{1}{3}$ . Also in our group it has been shown that SCS-MP2 provides accurate results for a variety of systems, such as the benzene dimer<sup>[55]</sup> or aurophilic interactions<sup>[54]</sup> featured in Chapters 4 and 6, respectively.

### 2.2.2 Coupled-Cluster Theory

In Coupled-Cluster theory<sup>[17,18]</sup> the Hartree-Fock wave-function also serves as the reference. Excitations to virtual orbitals are constructed with an exponential ansatz

$$e^{\hat{T}} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + \dots \quad (2.15)$$

Applying the excitation operators  $T$  to the reference wave-function  $\Phi_0$  yields

$$\begin{aligned} \hat{T}_1\Phi_0 &= \sum_i^{\text{occ}} \sum_a^{\text{vir}} t_a^i \Phi_i^a \\ \hat{T}_2\Phi_0 &= \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} t_{ab}^{ij} \Phi_{ij}^{ab} \\ &\dots \end{aligned} \quad (2.16)$$

$t_a^i$  and  $t_{ab}^{ij}$  are cluster amplitudes describing the probability of an excitation from occupied orbitals  $i, j$  to virtual orbitals  $a, b$ . Since a treatment of all excitations is computationally not feasible the excitation operator is truncated at a certain point. Often already the triples are cut off, as the doubles provide the largest contribution to the correlation energy. This method is then referred to as CCSD. The CCSD exponential excitation operator is expanded by

$$e^{\hat{T}_1 + \hat{T}_2} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + (\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1^3) + \dots \quad (2.17)$$

Through the exponential ansatz also higher excitations are included by connecting single and double excitations. This ensures the size extensive behaviour of CC methods. If the amplitudes and two-electron integrals are derived the closed-shell CCSD energy is computed as

$$E_{\text{CCSD}} = \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} (t_{ab}^{ij} + t_a^i t_b^j - t_b^i t_a^j) (2 \langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle) . \quad (2.18)$$

CC methods suffer, as well as MPPT, of being non variational.<sup>[76]</sup> The computational cost of canonical CCSD scales with the power of  $N_{\text{bas}}^6$  and is therefore, more demanding than MP2. Commonly, a correction to account for the triple excitations is included by applying MP4. To denote the perturbative treatment of the triples the prefix is written in brackets. Accordingly, the method is referred to as

CCSD(T).<sup>[78]</sup> It is known as the *gold standard* in theoretical chemistry. With triple zeta or even larger basis sets it is used to set up benchmarks and calculate highly accurate results. Its computational cost scales with the power of  $N_{\text{bas}}^7$ .

### 2.3 Local Correlation Methods

In the previous Section a short overview of Møller-Plesset and Coupled-Cluster theory was given. Both methods suffer of their steep scaling regarding the molecular system size. Considering truncations to double excitations, MP2 and CCSD scale with the formal powers of  $N_{\text{bas}}^5$  and  $N_{\text{bas}}^6$ , respectively. This can be overcome by using local orbital spaces together with a truncated virtual space. In the past years several publications dealt with methods to achieve a linear scaling of the computation time regarding the molecular size, starting with the pioneering work of Pulay et al.<sup>[34,79]</sup> up to the latest pair natural orbital approaches as discussed by the groups of Neese and Werner.<sup>[37-46]</sup> In the following Subsections it will be first shown that MP2 and CCSD can be formulated in an orbital invariant fashion. With this requirement different techniques to describe the occupied and virtual space are discussed.

#### 2.3.1 Orbital Invariant Formulation

The orbital invariant formulation provides the basis and the possibility to use different orbital representations for the computation of MP2 or CCSD energies and wave-functions.<sup>[80,81]</sup> The notations explained and applied here will be used throughout the whole thesis. For the first order MP2 wave-function we now use an expression containing MP2 amplitudes

$$\Psi^{(1)} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} |\Phi_{ij}^{ab}\rangle T_{ab}^{ij} . \quad (2.19)$$

To derive the amplitudes the Hylleraas functional is introduced

$$E_2 = 2 \langle \Psi^{(1)} | \hat{V} | \Phi_0 \rangle + \langle \Psi^{(1)} | \hat{H}^{(0)} - E^{(0)} | \Psi^{(1)} \rangle . \quad (2.20)$$

It has to be minimized in an iterative procedure. For this purpose, residuals are computed in first order amplitude equations. When the amplitudes are converged

the residuals must vanish

$$R_{ab}^{ij} = \langle \tilde{\Phi}_{ij}^{ab} | \hat{V} | \Phi_0 \rangle + \langle \tilde{\Phi}_{ij}^{ab} | \hat{H}^{(0)} - E^{(0)} | \Psi^{(1)} \rangle \stackrel{!}{=} 0 . \quad (2.21)$$

At this point it is convenient to use contravariant configurations. The configurations  $|\Phi_{ij}^{ab}\rangle$  are non-orthogonal for pairs  $ij$  and  $ji$  and the corresponding contravariant is derived by

$$|\tilde{\Phi}_{ij}^{ab}\rangle = \frac{1}{6} \left( \frac{1}{2} |\Phi_{ij}^{ab}\rangle + |\Phi_{ij}^{ba}\rangle \right) . \quad (2.22)$$

The projections of the contravariant configurations on the first order wave-function yield the amplitudes, with their corresponding contravariants defined as

$$\begin{aligned} T_{ab}^{ij} &= \langle \tilde{\Phi}_{ij}^{ab} | \Psi^{(1)} \rangle \\ \tilde{T}_{ab}^{ij} &= 2T_{ab}^{ij} - T_{ba}^{ij} . \end{aligned} \quad (2.23)$$

To finally calculate the energy with use of the Hylleraas functional the coupling between the first order wave-function and the reference has to be derived

$$\langle \tilde{\Phi}_{ij}^{ab} | \hat{V} | \Phi_0 \rangle = \langle \phi_i \phi_j | \phi_a \phi_b \rangle = K_{ab}^{ij} . \quad (2.24)$$

With the residuals  $R_{ab}^{ij}$ , the two electron integrals  $K_{ab}^{ij}$  and the contravariant amplitudes  $\tilde{T}_{ab}^{ij}$  the Hylleraas functional can than be computed as

$$E_2 = \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} (K_{ab}^{ij} + R_{ab}^{ij}) \tilde{T}_{ab}^{ij} . \quad (2.25)$$

For the optimized amplitudes the residuals vanish, so that the final MP2 energy expression is given by

$$E_{\text{MP2}} = \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} K_{ab}^{ij} \tilde{T}_{ab}^{ij} . \quad (2.26)$$

It is convenient to further simplify the equations by introducing the matrix notation

$$E_{\text{MP2}} = \sum_{ij}^{\text{occ}} = \text{tr}[\mathbf{K}^{ij} \tilde{\mathbf{T}}^{ij}] \quad (2.27)$$

with  $[\mathbf{K}^{ij}]_{ab} = K_{ab}^{ij}$ , the same applies correspondingly for  $\mathbf{T}^{ij}$  and  $\mathbf{R}^{ij}$ . If the orbital invariant CCSD approach is considered, the formalism is much more complex due to the coupling of the different excitations classes. The derivation of the equations is discussed in great detail in Reference [81]. In the end, this leads to equivalent energy expressions for the double excitations compared to MP2.

### 2.3.2 Orbital Localization

Orbital localization schemes transform canonical orbitals to local molecular orbitals (LMOs) and restrict their spatial extent. For this purpose, one has to find a unitary transformation which holds

$$|\phi_i^{loc}\rangle = \sum_k^{\text{occ}} |\phi_k^{\text{can}}\rangle U_{ki}^{\text{loc}} \quad (2.28)$$

Two of the most popular approaches to find  $U^{\text{loc}}$  are the Forster-Boys (FB)<sup>[82]</sup> and Pipek-Mezey (PM)<sup>[83]</sup> methods. In the FB method the space spanned by an orbital is minimized. This is achieved through minimizing

$$P^{\text{FB}} = \sum_p^{[x,y,z]} \sum_i^{\text{occ}} \langle \phi_i | (\hat{p} - \langle \phi_i | \hat{p} | \phi_i \rangle) | \phi_i \rangle . \quad (2.29)$$

The PM scheme maximizes the sum of the squared Mulliken partial charges (will be discussed in Section 2.4) as

$$P^{\text{PM}} = \sum_A^{\text{atoms}} \sum_i^{\text{occ}} q_{iA}^2 . \quad (2.30)$$

Effectively, the number of atoms over which an orbital is spanned is minimized.

The FB method yields well localized orbitals but suffers of so called banana like bonds.<sup>[84]</sup> Considering a double bond between two carbons the FB methods describes it by two equal banana shaped orbitals resulting from a  $sp^2$  hybridization of the carbons. The PM scheme keeps  $\sigma$ - $\pi$  separation, so that the double bond consists of one  $\sigma$  and one  $\pi$  orbital. Therefore, PM localization allows for a chemical intuitive interpretation. However, PM suffers of poor localization, if large basis sets with diffuse functions are used. In order to overcome this problem it is convenient to neglect the diffuse basis functions of each angular momentum during the localization procedure.<sup>[85]</sup>

Another ansatz is the application of intrinsic bond orbitals (IBOs).<sup>[86]</sup> Thereby, a minimal basis is used to generate intrinsic atomic orbitals (IAOs). The occupied space is expanded in this basis and IAO charges are computed. The orbitals in IAO basis are PM localized yielding IBOs through back transformation to the full basis set. They are quite insensitive to basis set variations and the IAO charges

converge smoothly with the basis set size.<sup>[86]</sup> This behaviour is also advantageous in the domain definition which will be discussed in the next Subsection.

### 2.3.3 Virtual Space Definition

The efficiency of local correlation treatments is achieved by truncating the virtual space. In order to accurately reproduce the canonical correlation energy, the orbital domains have to span a large enough space. However, they should be as small as possible to increase the speed of the methods. Over the past years several methods to automatically generate orbital domains have been presented. In this work use is made of projected atomic orbitals (PAOs -  $r, s$ ), orbital specific virtuals (OSVs) and pair natural orbitals (PNOs -  $p, q$ ) approaches.

The application of PAOs to span the virtual space was proposed by Pulay. PAOs are constructed by projecting out the occupied space from the full set of AOs.<sup>[34]</sup> This is achieved by

$$|\phi_r^{\text{PAO}}\rangle = (1 - \sum_i^{\text{occ}} |\phi_i\rangle \langle \phi_i|) |\phi_r^{\text{AO}}\rangle = \hat{P} |\phi_r^{\text{AO}}\rangle . \quad (2.31)$$

For an orbital  $\phi_i$  a set of PAOs spanning its virtual space is defined. It is referred to as the orbital domain  $[i]_{\text{PAO}}$ . The PAOs belonging to the domains are commonly determined by applying a population analysis (will be discussed in Section 2.4) and a completeness criterion, like Boughton-Pulay (BP),<sup>[84]</sup> to all occupied orbitals. The outcome identifies the atoms which each localized MO spans. Since PAOs are center based, the domain is constructed as the union of all PAOs located on those atoms. The PAO pair domains, which are necessary to derive pair energies, are then constructed by taking the union of the orbital domains ( $[ij]_{\text{PAO}} = [i]_{\text{PAO}} \cup [j]_{\text{PAO}}$ ).

Compared to the canonical treatment of the virtual space, in case of PAOs only a fraction of the integrals and amplitudes have to be computed to accurately reproduce the canonical correlation energy. This is further reduced by applying pair specific virtual orbital approaches.

PNOs are constructed by approximate amplitudes to build pair density matrices for

## 2 Theoretical Background

---

each possible orbital pair  $ij$ <sup>[36,87]</sup>

$$D_{ab}^{ij} = \frac{1}{1 + \delta_{ij}} [\tilde{T}^{ij\dagger} T^{ij} + \tilde{T}^{ij} T^{ij\dagger}]_{ab} . \quad (2.32)$$

In the original formulation the approximate amplitudes were derived using the canonical virtual space. Latest linear scaling methods, regarding the molecular size, take the advantages of PAOs or OSVs to already reduce in this first step the computational cost.<sup>[39,42]</sup> Diagonalizing the pair density matrices yields pair-specific transformation matrices

$$[Q^{ij\dagger} D^{ij} Q^{ij}]_{pq} = n_p^{ij} \delta_{pq} , \quad (2.33)$$

with natural occupation numbers  $n_p^{ij}$  for each PNO  $p$ . For each orbital pair  $ij$  the PNO pair domains are then constructed using the diagonalization matrices. PNOs with smaller natural occupation numbers than a certain cut-off are neglected.<sup>[39,40]</sup> If canonical virtual MOs are used to build the initial pair densities the PNOs are obtained by

$$|p^{ij}\rangle = \sum_a^{\text{vir}} |a\rangle Q_{ap}^{ij} . \quad (2.34)$$

For a pair function  $ij$  the PNOs are by construction orthonormal. However, for different orbital pairs the PNOs are not orthogonal. Commonly, this is handled by pseudocanonicalizing the PNOs by a series of transformations.<sup>[81]</sup> Compared to PAOs the number of virtual orbitals needed to accurately recover the canonical correlation energy is reduced by at least one order of magnitude.<sup>[42]</sup>

OSV domains are constructed by building pseudocanonical PNOs for each diagonal pair  $ii$ .<sup>[88–90]</sup> The obtained OSV domains  $[i]_{\text{OSV}}$  are then merged for pairs  $ij$  by taking their union similar to PAOs. The total numbers of OSVs and PNOs scales linearly with the molecular size. In case of OSVs less virtual orbitals have to be build compared to PNOs. However, the pair domains have to be 3-4 times larger for the OSV approach to obtain the same accuracy.<sup>[42]</sup>

The latest PNO-LMP2 and PNO-LCCSD approaches of the Werner group combine PAOs, OSVs and PNOs in a series of transformations.<sup>[42,44]</sup> First PAO domains are built and approximate diagonal pair amplitudes  $T^{ii}$  are computed. From these amplitudes OSVs are derived by their diagonalization. Further, pair domains for all orbital pairs  $ij$  are built and the integrals and amplitudes are evaluated in this basis. In the final step the PNOs are constructed using the OSV amplitudes. All necessary integrals and amplitudes are obtained by transformation from the OSV to the PNO



basis. For further information the interested reader should refer to References [42–46].

## 2.4 Population Analysis

Population analysis aims to decompose the electron density into atomic contributions. Two of the most known and popular approaches are the Mulliken<sup>[91]</sup> and Löwdin<sup>[92]</sup> analysis. In case of the Mulliken analysis the electron population of an atom  $A$  is computed by

$$\begin{aligned}\rho_A &= \sum_{\mu \in A} \sum_{\nu}^{N_{\text{bas}}} D_{\mu\nu} S_{\mu\nu} \\ D_{\mu\nu} &= \sum_i^{\text{occ}} \sum_{\mu\nu}^{N_{\text{bas}}} n_i C_{\mu i} C_{\nu i} .\end{aligned}\tag{2.35}$$

Here  $D_{\mu\nu}$  is the density matrix and  $n_i$  the occupation number of an orbital. The Löwdin approach applies a transformation matrix to the density, making use of an orthogonal AO-space

$$\rho_A = \sum_{\mu \in A} \sum_{\nu}^{N_{\text{bas}}} S_{\mu\nu}^{1/2} D_{\mu\nu} S_{\mu\nu}^{-1/2} .\tag{2.36}$$

There are other methods using similar transformations to partition the wavefunction. They lack in several issues.<sup>[76]</sup> If the Mulliken scheme is employed, the Pauli principle can be violated, such that an orbital is occupied by more than two electrons. Furthermore, a negative count of electrons between two basis functions is possible. Additionally, the method simply splits off diagonal contributions without regard to polarization. Those three problems are solved when applying the Löwdin transformation. However, this scheme suffers from other problems. Considering diffuse basis functions on an atom, they describe the electron density far away, but are however counted to the population on the atoms where they are centred. Moreover, higher angular momenta like dipoles, quadrupoles and so on are not reproduced.

A main problem of Mulliken and Löwdin population analysis is the basis set sensitivity.<sup>[76]</sup> There is no convergence when increasing the basis set size and for large basis sets with diffuse functions the electron population becomes unreliable. To avoid this behaviour intrinsic atomic orbital charges can be used.<sup>[86]</sup> Thereby,

## 2 Theoretical Background

---

a minimal basis set is introduced to span the occupied space and the population analysis is evaluated in this basis.

Another approach, showing convergence behaviour with the basis set size, is the natural population analysis as proposed by Weinholdt.<sup>[93,94]</sup> The AOs are transformed and evaluated in a natural AO basis (NAOs). Thereby, the density is transformed in a nonorthogonal basis  $\mathbf{\Gamma} = \mathbf{SDS}$  and atom wise ( $A, B, C$ , and so on) blocks are build

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Gamma}^{(AA)} & \mathbf{\Gamma}^{(AB)} & \mathbf{\Gamma}^{(AC)} & \dots \\ \mathbf{\Gamma}^{(BA)} & \mathbf{\Gamma}^{(BB)} & \mathbf{\Gamma}^{(BC)} & \dots \\ \mathbf{\Gamma}^{(CA)} & \mathbf{\Gamma}^{(CB)} & \mathbf{\Gamma}^{(CC)} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (2.37)$$

For elements  $\mathbf{\Gamma}_{\mu\nu}^{(AB)}$  it applies  $\mu \in A$  and  $\nu \in B$ . The same separation is done for the overlap matrix in the AO basis. The generalized eigenvalue problem

$$\mathbf{\Gamma}^{(AA)}\mathbf{X} = \mathbf{S}^{(AA)}\mathbf{X}\mathbf{W} \quad (2.38)$$

is solved for all diagonal sub-matrices. With  $\mathbf{W}$  as diagonal matrix containing the eigenvalues, and  $\mathbf{X}$  as pre-NAOs. Through a series of diagonalizations and orthogonalizations, the transformation matrix  $\mathbf{T}^{\text{NAO}}$  is derived. Applying  $\mathbf{T}^{\text{NAO}}$  to  $\mathbf{\Gamma}$  a block diagonal density matrix is obtained

$$\tilde{\mathbf{D}} = \mathbf{T}^{\text{NAO}\dagger}\mathbf{\Gamma}\mathbf{T}^{\text{NAO}} \quad (2.39)$$

Summing up the diagonal elements for NAOs belonging to a certain center one obtains the electron population in the NAO basis

$$\rho_A = \sum_{r \in A} \tilde{D}_{rr} \quad (2.40)$$

This method is computationally much more demanding, if compared to the simple formulations of Mulliken and Löwdin analysis. However, for large basis sets the application is recommended since Mulliken as well as Löwdin analysis provide unreliable electron populations.<sup>[76]</sup> In Sections 3.1 fragmentation techniques based on population analysis are discussed. It will be shown that the method of choice has to be robust (e.g. relative to the basis set size) and the electron population has to be physically meaningful.

## 2.5 Energy Decomposition Analysis (EDA)

Understanding intermolecular interactions has always been a central topic to theoretical chemistry, from the smallest systems to supramolecular chemistry. This knowledge is vital in guiding the design and synthesis of new molecular constructs or in the interpretation of complex spectra, may it be in the gas phase, solution or condensed phase. Over the years, several schemes have been presented to dissect the different forces at play, including the seminal work by Kitaura and Morokuma<sup>[95,96]</sup> or later the powerful family of Symmetry-Adapted Perturbation Theory (SAPT) approaches.<sup>[29,30]</sup> For some time now, in parallel to these developments, use has been made of local orbital based analysis to separate the different energy contributions.<sup>[31–33]</sup> No matter what scheme is picked, a few properties should be kept. 1) the separation of energy terms should sum to the total interaction energy, 2) the analysis should converge with an increasing level of theory (may it be the method or the basis set), 3) generally applicable to any molecular aggregate.

The original Kitaura-Morokuma decomposition scheme of the Hartree-Fock interaction energy suffered of numerically unstable energy components and its basis set sensitivity. Solutions to these problems have been presented in several later publications and plenty decomposition schemes like natural EDA,<sup>[97]</sup> absolutely localized molecular orbital EDA<sup>[98]</sup> or reduced variational space EDA<sup>[99]</sup> sprung up the ground. Here, a brief overview of the energy decomposition analysis scheme proposed by Su<sup>[100]</sup> is discussed. In the latter, the HF interaction for a supermolecule  $X$ , consisting of monomers  $A$  and  $B$ , is decomposed in 4 energy terms, allowing for a physical interpretation of the binding nature.

$$\Delta E^{\text{HF}} = E_X^{\text{HF}} - E_A^{\text{HF}} - E_B^{\text{HF}} = \Delta E^{\text{ele}} + \Delta E^{\text{ex}} + \Delta E^{\text{rep}} + \Delta E^{\text{pol}} \quad (2.41)$$

Firstly, HF calculations of the supermolecule  $X$  and the monomers  $A, B$  are performed. Through merging the obtained monomer canonical orbitals the coefficient matrix  $C_{AB}$  is derived

$$\mathbf{C}_{AB} = \begin{pmatrix} \mathbf{C}_A & 0 \\ 0 & \mathbf{C}_B \end{pmatrix} \quad (2.42)$$

With this basis two intermediates are calculated

$$E_X^{(1)} = 2\text{Tr}[\mathbf{C}_{AB}^T \mathbf{h}_X \mathbf{C}_{AB}] + 2\text{Tr}[\mathbf{C}_{AB}^T \mathbf{J}_{AB} \mathbf{C}_{AB}] - \sum_A^{\text{mon}} \text{Tr}[\mathbf{C}_A^T \mathbf{K}_A \mathbf{C}_A] + E_X^{\text{nuc}}, \quad (2.43)$$

## 2 Theoretical Background

---

$$E_X^{(2)} = 2\text{Tr}[\mathbf{C}_{AB}^T \mathbf{h}_X \mathbf{C}_{AB}] + 2\text{Tr}[\mathbf{C}_{AB}^T \mathbf{J}_{AB} \mathbf{C}_{AB}] - \text{Tr}[\mathbf{C}_{AB}^T \mathbf{K}_{AB} \mathbf{C}_{AB}] + E_X^{\text{nuc}} , \quad (2.44)$$

and electrostatic and exchange terms are obtained with

$$\Delta E^{\text{ele}} = E_X^{(1)} - \sum_A^{\text{mon}} E_A^{\text{HF}} , \quad (2.45)$$

$$\Delta E^{\text{ex}} = E_X^{(2)} - E_X^{(1)} . \quad (2.46)$$

As it can be easily observed the components display the sum of the exchange and coulomb integrals between the unperturbed monomer orbitals, respectively. In order to calculate the repulsion and polarization terms, the monomer orbitals are orthonormalized ( $\mathbf{C}_{\overline{AB}}$ ) and a third intermediate is constructed

$$\mathbf{S}_{AB}^{-1/2} = (\mathbf{C}_{AB}^T \mathbf{S}_X \mathbf{C}_{AB})^{-1/2} , \quad (2.47)$$

$$\mathbf{C}_{\overline{AB}} = \mathbf{C}_{AB} \mathbf{S}_{AB}^{-1/2} , \quad (2.48)$$

$$E_X^{(3)} = 2\text{Tr}[\mathbf{C}_{\overline{AB}}^T \mathbf{h}_X \mathbf{C}_{\overline{AB}}] + 2\text{Tr}[\mathbf{C}_{\overline{AB}}^T \mathbf{J}_{\overline{AB}} \mathbf{C}_{\overline{AB}}] - \text{Tr}[\mathbf{C}_{\overline{AB}}^T \mathbf{K}_{\overline{AB}} \mathbf{C}_{\overline{AB}}] + E_X^{\text{nuc}} . \quad (2.49)$$

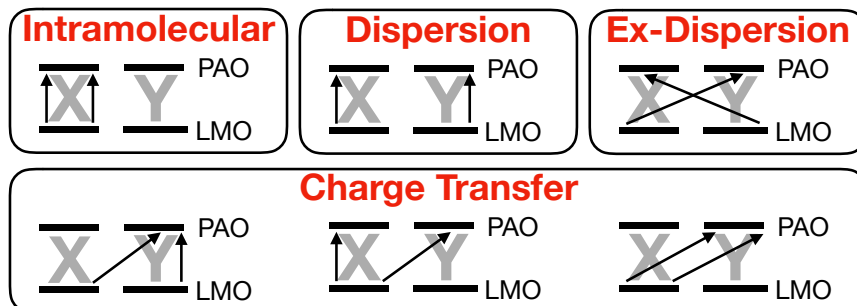
The repulsion energy is expressed as the difference between the HF energy of the supermolecule in the orthonormalized basis and the primary orbitals. The polarization contribution is extracted by the discrepancy between the HF result of the supermolecule and the HF result with the orthonormalized monomer orbitals

$$\Delta E^{\text{rep}} = E_X^{(3)} - E_X^{(2)} , \quad (2.50)$$

$$\Delta E^{\text{pol}} = E_X^{\text{HF}} - E_X^{(3)} . \quad (2.51)$$

In their publication they also considered electron correlation effects using CCSD and MP2. The dispersion contribution was estimated by the difference between the MP2/CCSD and HF energies. However, since dispersion is only one contribution to the dynamical electron correlation this should be avoided. Schütz et al. suggested a scheme to partition the electron correlation on the basis of local correlation methods using PAOs as virtual space.<sup>[47]</sup> As described in Section 2.2.1 the electron correlation, in case of LMP2, is derived by coupling the reference and the first-order wavefunction. Thereby, double excitations from LMOs to PAOs are considered. Their locality allows to assign both to fragments. Regarding this assumption the double excitations are decomposed into different classes as shown in Figure 2.1.

Double dynamic and intramolecular ionic excitations are excluded to correct the



**Figure 2.1:** Schematic representation of the decomposition of double excitations in local CCSD and MP2 approaches. The charge transfer terms are divided in intermolecular ionic, intramolecular ionic and double dynamic excitations in order.

basis set superposition error (BSSE). Considering two interacting monomers their basis functions overlap at short ranges.<sup>[76]</sup> Effectively, this involves an increased basis set size regarding each monomer itself. Since in most cases one is interested in forming relative energies, a mismatch arises compared to the separated monomers. However, there is still an ongoing debate if those excitations should be excluded, since for largely polarized systems it is observed that at least up to 15% of the correlation energy is lost when applying this scheme.<sup>[47]</sup>

After the exclusion of the upper excitations four contributions are left forming the total electron correlation energy, namely dispersion, exchange-dispersion, ionic and intramolecular contributions

$$E_{\text{corr}} = E_{\text{disp}} + E_{\text{exdisp}} + E_{\text{ionic}} + E_{\text{intra}} \quad (2.52)$$

In terms of LCCSD the double excitations are treated in the same way. In the scheme proposed by Schütz *et al.* all single excitations are concluded to be intramolecular contributions, as well as the disconnected single excitations obtained by the  $T_1^2$  operator. They are a by-product of the exponential approach and ensure size consistency. Moreover, the singles contributions to the interaction energy is only a fraction of the doubles contribution, so that in fact their treatment does not significantly affect the energy partitioning.

Neese *et al.* presented in a recent publication their Local Energy Decomposition (LED) scheme.<sup>[1]</sup> The partitioning of the HF interaction energy is closely related to the approach of Su. For the electrostatic and exchange terms use is made of local

## 2 Theoretical Background

---

orbital spaces regarding these contributions directly by coupling orbitals located on different fragments. Instead of polarization and repulsion terms an electronic preparation contribution is introduced, describing the energy needed to deform the electronic structure for an optimum interaction between fragments. In the local correlation treatment instead of PAOs the virtual space is represented by PNOs. The location of PAOs is known by their construction using an *ad hoc* criterion. In case of PNOs their locality has to be determined to apply a decomposition similar to Figure 2.1. Since PNOs are by nature delocalized they need to be localized in a first step. This is done in the same fashion as for occupied orbitals. Then in a second step, a population analysis has to be applied to assign the PNOs to a certain fragment. As will be shown in Chapter 7 this needs to be carefully evaluated.

Another ansatz to decompose the interaction energy in physical contributions is the Symmetry-Adapted Perturbation Theory (SAPT).<sup>[29]</sup> Over the past years several different treatments to derive intermolecular potentials using SAPT were published and discussed.<sup>[30,101]</sup> Here a brief overview to the wave-function based ansatz is given.

Considering molecular dimer systems the SAPT interaction energy is expanded as a perturbation series

$$\Delta E^{\text{SAPT}} = \sum_n \sum_k \sum_l (E_{\text{pol}}^{(nkl)} + E_{\text{exch}}^{(nkl)}) \quad (2.53)$$

with  $n$  as the order of the intermolecular potential and  $k, l$  as the order of the intramonomer electron correlation. The energy terms derived are then designated with  $E^{(vw)}$ , where  $v = n$  and  $w = k, l$ . One of the earliest and most common approaches is SAPT0. Hereby, the intermolecular potential is treated at second order, while the intramolecular perturbation is of zeroth order HF-level. Following this scheme the interaction energy is provided by

$$\begin{aligned} \Delta E^{\text{SAPT0}} &= [E_{\text{elst}}^{(10)}]_{\text{elst}} + [E_{\text{exch}}^{(10)}]_{\text{exch}} + [E_{\text{ind,r}}^{(20)} + E_{\text{exch-ind,r}}^{(20)} + \delta E_{\text{HF}}^{(2)}]_{\text{ind}} \\ &\quad + [E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}]_{\text{disp}} \\ &= \Delta E^{\text{HF}} + [E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}]_{\text{disp}} \end{aligned} \quad (2.54)$$

with  $\delta E_{\text{HF}}^{(2)}$  as a correction to account for third order terms.<sup>[30]</sup> Further, it establishes an equality to the HF interaction energy. The subscripts reveal that the interaction energy is split up in four physical components, namely electrostatic, exchange,

induction and dispersion. With increasing order of the inter- or intramolecular perturbation more energy terms are derived and are successively assigned to those components. Here one should at least mention SAPT2 since it is also commonly used. Compared to SAPT0 the intramolecular electron correlation is treated by second order perturbation theory and the performance of the method is quite comparable to MP2.<sup>[30]</sup> The number of SAPT approaches is numerous and for further reading, reviews by Sherrill, Parker, Szalewicz and Jeziorski give insight in great detail.<sup>[29,30,101]</sup>

SAPT implementations on the basis of DFT are also commonly used. It has been shown that they are in good agreement with wave function based SAPT<sup>[102,103]</sup> and furthermore, the computational cost is dramatically reduced. In the past years, there was also great progress in using local orbitals spaces in the framework of SAPT to partition interaction energies. The Sherrill group implemented different SAPT approaches including atomic,<sup>[31]</sup> fragment<sup>[32]</sup> and intramolecular<sup>[33]</sup> partitioning.

Finally, the relationships between the combined HF-EDA/local correlation energy analysis and SAPT should be discussed. Regarding HF-EDA and SAPT0 clear relationships can be drawn. The electrostatic terms are derived in a similar fashion and therefore, show the same behaviour. The SAPT0 exchange is comparable to the exchange plus repulsion in HF-EDA and lastly the polarization term in HF-EDA is accounted by the induction term in SAPT0. When applying intramonomer correlation treatments of higher order in SAPT, more terms are added and the relation breaks to a certain extent.

Considering the dispersion interactions obtained via local orbital subspaces and SAPT, the contributions are often larger in SAPT approaches. Since dispersion is no measurable observable, the discussion which method gives the correct value is somewhat arbitrary. It has been shown that the dispersion forces predicted by both methods are at least in the same order of magnitude.<sup>[1]</sup> The nature of the charge transfer contributions obtained in local orbital approaches is still an ongoing debate. This will be discussed in Chapter 7.

Both methods have their advantages and disadvantages. SAPT is by construction BSSE free and different interaction components are directly defined through operators. Furthermore, the energy contributions are well defined adding higher order terms to all components. However, the computational cost of higher order SAPT

## 2 Theoretical Background

---

approaches increases rapidly and therefore, are only possible for mid-sized molecular systems. Considering the combined approach, the HF-EDA suffers from the single Slater determinant reference wave-function including no higher order terms. In several weakly bound molecular systems the major interest is often the dynamic electron correlation and therefore, it is often useless to decompose the HF interaction. Through local approximations computations for large clusters are possible and a qualitative insight in the nature of binding is given. Moreover, the local treatment allows for fragmentation and visualization techniques as presented in Chapter 3. The result whether a molecular system is bound electrostatically or dispersive should be the same for both approaches.



## 3 Local orbital analysis

The local treatment of dynamic electron correlation reduces the computational effort compared to the initial canonical methods by far. However, it is not the only benefit. Local orbital subspaces are able to provide a deeper insight to physical forces and the influence of different functional groups

In the following Sections the methods are discussed in the framework of local second-order Møller-Plesset perturbation theory (LMP2). It is also possible to apply the same principles to LCCSD. Hereby, the single excitations as well as the coupled single excitations ( $T_1^2$  operator) are classified as intramolecular contributions. Once having derived the reference wave-function via HF the canonical orbitals are localized using a conventional procedure as outlined in Section 2.3.2. The virtual space is represented by PAOs. The application of PNOs or OSVs will be discussed in Chapter 7. For both, at first glance their locality is not known and has to be determined.

### 3.1 Fragmentation Using Local Orbital Subspaces

Consider a molecular dimer consisting of the monomers  $A$  and  $B$ . Often, it is of interest to consider a particular region of a monomer separately, such as a functional group or varying substituents. This can be realized by the assignment of atoms to fragments. The number of fragments is arbitrary and depends on the system. If both monomers are divided into two fragments  $A_1 - A_2$  and  $B_1 - B_2$ , the total intermolecular interaction is calculated from the possible combinations

$$\Delta E_{\text{inter}} = \Delta E(A_1 \cdots B_1) + \Delta E(A_1 \cdots B_2) + \Delta E(A_2 \cdots B_1) + \Delta E(A_2 \cdots B_2) . \quad (3.1)$$

To obtain these contributions on the basis of wave-function methods, orbitals must be assigned to atoms. In local correlation treatments this can be justified through

population analysis. NPA charges<sup>[93,94]</sup> are calculated for all LMOs. Orbital spanned over atoms within a single fragment can be clearly assigned. For orbitals shared between fragments, the resulting energy is weighted in percentages based on the NPA charges<sup>[51]</sup>

$$\Delta E_i^A = \frac{\rho_A(i)}{\rho_A(i) + \rho_R(i)} \Delta E_i, \quad (3.2)$$

$\rho_A(i)$  denotes the charge of orbital  $i$  on fragment  $A$  and  $\rho_R(i)$  the NPA charges on other fragments. The correlation energy is calculated from all pair energies. If the orbitals are assigned to the fragments, the individual energy terms, as in Equation 3.1, can be calculated as follows

$$\Delta E(A \cdots B) = \sum_{i \in A} \sum_{j \in B} \frac{\rho_A(i)}{\rho_A(i) + \rho_R(i)} \frac{\rho_B(j)}{\rho_B(j) + \rho_R(j)} \epsilon_{ij}. \quad (3.3)$$

Instead of NPA charges also IAO charges<sup>[86]</sup> can be used. The application of the Mulliken or Löwdin analysis should be disregarded since they show no convergence with the basis set size and therefore, the fragmentation scheme would also not converge.

## 3.2 Dispersion Interaction Density (DID)

This Section is reproduced from the introduction and methodology sections of Reference:

[55] A. Wuttke, R. A. Mata, Visualizing dispersion interactions through the use of local orbital spaces, *Journal of computational chemistry*, **2017**, *38*, 15-23 (DOI: 10.1002/jcc.24508),

with the permission of John Wiley & Sons, Inc. In Subsection 3.2.2 the visualization possibilities are expanded by the Voxel representation.

Considering weakly interacting molecular systems for decades electrostatics were determined as major driving force. They can already reach a considerable size for small molecular systems. Since calculations for large systems have only been available for a few years, it was only gradually apparent how important dispersion forces are.<sup>[67]</sup> Nowadays it is well-known that dispersion interactions determine many chemical processes such as protein folding, molecular recognition and crystal formation.<sup>[68,69,104]</sup> In recent publications the functionalization of so called dispersion

energy donors (DEDs) is argued. Thereby, mainly the introduction of organic bulky groups is discussed.<sup>[105,106]</sup> All atoms in a system contribute to the attraction (as small as this contribution may be). The total sum can become as significant as the strong electrostatic terms. The latest generation of local correlation methods allows calculations with quantum mechanical accuracy in an unprecedented size scale of molecular systems. To take advantage of these advancements in theory and in order to analyze them, the development of new tools is of great interest. Instead of simply looking at numbers, a pictorial representation allows an intuitive interpretation.

Recently, Parrish and Sherrill introduced an atomic partitioning of Symmetry-Adapted Perturbation Theory (A-SAPT).<sup>[31]</sup> This procedure is general for all types of molecular interactions. An extension to intramolecular interactions has also been put forward.<sup>[33]</sup> The latter is achieved through the use of an embedding procedure in the description of covalent bound fragments. An alternative in the context of SAPT has also been presented, making use of the Chemical Hamiltonian<sup>[107]</sup> to capture intramolecular interactions.<sup>[108]</sup>

A fully different method, but one which has found some resonance in the community is the non-covalent interaction (NCI) approach of Yang and coworkers.<sup>[109]</sup> In this case, the electronic density, its reduced gradient as well as the Laplacian are used to identify and characterize regions of weak interactions. The NCI method can be applied to both DFT and wave-function methods. However, several issues could be raised about the method when considering dispersion. First of all, it has been proposed that low density and reduced density gradient regions, with a  $\lambda^2$  component of the Laplacian close to zero should indicate van der Waals interactions. However, this provides no criterium to distinguish between Keesom, Debye and London forces. The criteria itself draws only a thin line between weak repulsion and interactions. Secondly, gradient isosurfaces, which ultimately define an interaction volume between two interacting molecules is not coincident with the physical picture of London dispersion forces. These come about as electronic densities interact through instantaneously induced dipoles, which can be significantly far away in space. A contact region where density change takes place (as small as this may be) is not a requirement. In short, without acknowledging the value of the analysis, its application to the specific case of London interactions should be considered with caution.

The here presented approach to visualize dispersion interactions is based on the

well-established correlation energy decomposition used in the context of local correlation methods.<sup>[47,50]</sup> Although closely related to the A-SAPT approach, the use of local orbital spaces with correlated wave-function methods holds several advantages. Foremost, it allows for a decomposition directly from regular single point calculations. With very little modifications, intramolecular interactions can also be considered.

#### 3.2.1 Method

The DID method can be applied to any number of fragments. For sake of simplicity the example of two interacting systems  $A$  and  $B$  is considered. The excitation space of a LMO  $i$  is restricted to a group of PAOs located in close vicinity of the latter. This constitutes the orbital domain  $[i]$ . In the case of LMP2, the amplitude matrix elements  $\mathbf{T}^{ij}$  are limited to  $r, s \in [ij]$ , a pair domain formed from the union of the orbital domains  $[i]$  and  $[j]$ . The pair energies  $\epsilon_{ij}$  can be decomposed according to the spatial location of  $i, j, r$  and  $s$ .<sup>[47]</sup> In the specific case of dispersion interactions, the energy is given by the excitations where  $i, r$  are located in fragment  $A$ , while  $j, s$  are contained in  $B$ . This corresponds to a simultaneous excitation of two electrons within separate interacting fragments, and is thoroughly consistent with the corresponding SAPT operator definition. By considering only these elements in the amplitude matrices, the dispersion pair energies  $\epsilon_{ij}^d$  are obtained. A matrix  $\mathbf{D}^A$  in the AO space for a fragment  $A$  is computed as

$$D_{\mu\nu}^A = \sum_{i \in A} \frac{1}{N_i} \left( \sum_{j \in B} \epsilon_{ij}^d \right) P_{\mu\nu}^i . \quad (3.4)$$

Hereby,  $\mathbf{P}^i$  corresponds to the closed-shell orbital density of orbital  $i$ , obtained from the orbital indices as

$$P_{\mu\nu}^i = 2C_{\mu i}C_{\nu i} . \quad (3.5)$$

LMOs  $i$  are located on fragment  $A$ , while  $j$  is restricted to the interaction partner  $B$ .  $N_i$  represents the number of electrons for orbital  $i$  and equals 2 for closed-shell systems.

In order to visualize the interaction, a numerical grid is generated and the value for

a point at a given coordinate  $\mathbf{r}$  is computed as

$$\Gamma^A(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}^A \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}) . \quad (3.6)$$

By integrating  $\Gamma^A(\mathbf{r})$  over the whole space the dispersion interaction between the fragment  $A$  and  $B$  is recovered. For this reason,  $\Gamma^A(\mathbf{r})$  is interpreted as a Dispersion Interaction Density (DID) with  $\mathbf{D}^A$  being this quantity in AO matrix representation. If there are several fragments of interest, the scheme can be extended by including all fragments to the sum over  $j$  in Equation 3.4.

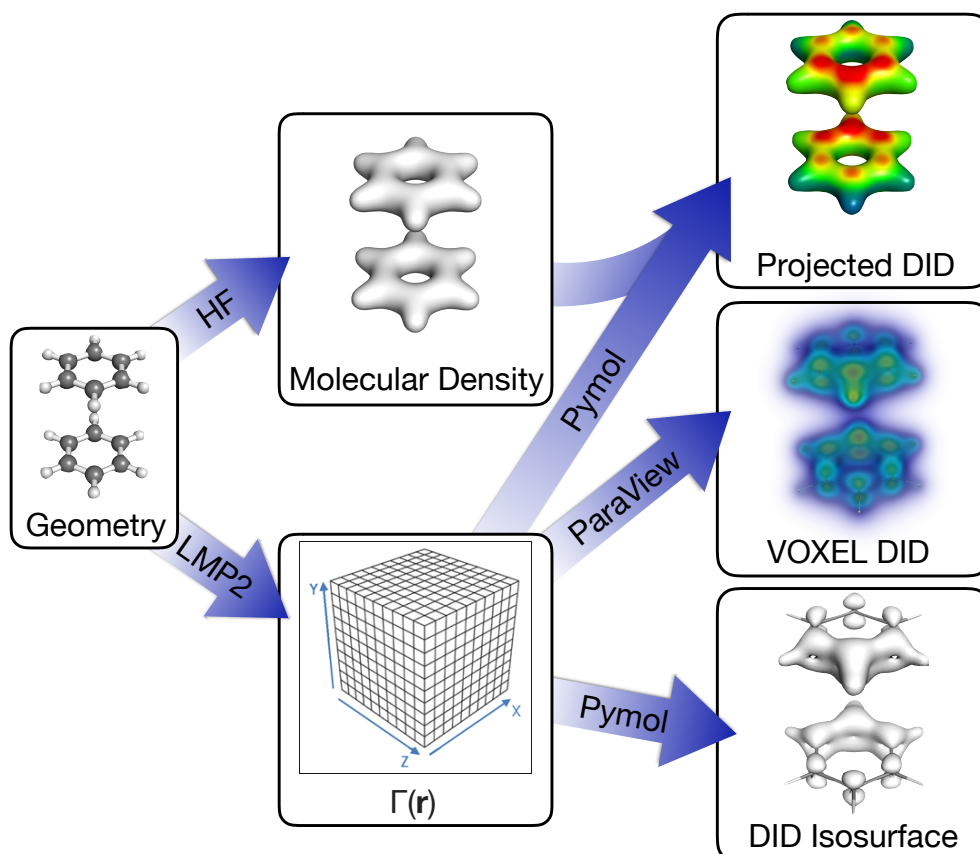
The outlined approach is closely related to F-SAPT<sup>[32]</sup> and A-SAPT<sup>[31]</sup> methods, with the extension to intramolecular interactions. Therefore, a direct comparison between the two is warranted:

- The use of local correlation based analysis only allows for a decomposition of correlation energy terms, while the SAPT ansatz covers all types of interactions (including electrostatics, exchange, and so forth).
- However, the decomposition used in this work is a direct by-product of a wavefunction calculation. It is an interpretation of readily available quantities.
- In the case of A/F-SAPT the virtual space is spanned by monomer canonical virtuals. In our case, PAOs are used, with the excitations being restricted to each orbital domain. For large molecules this can be significant since a transformation of the SAPT terms to the PAO space would show that  $T_{rs}^{ij}$  contributions would be counted, whereby  $r$  and  $s$  are within each respective monomer but not necessarily in the direct vicinity of  $i$  and  $j$ .
- Intramolecular partitioning of SAPT (ISAPT) energies is carried out through the use of embedding Fock-potentials. The latter requires a charge analysis, artificial splitting of the protons of atoms covalently bond and the construction of Coulomb and exchange embedding potentials. This leads to issues such as spurious polarization effects in the linker bonds. In the case of local calculations, the only approximation required is the definition of the local orbital spaces.

### 3.2.2 Visualization of DID

The previously described procedure is implemented in Molpro 2015.1.<sup>[110]</sup> Output is perceived in the form of a *.cube* file. It contains the molecular structure, the dimensions of the grid and the DID values for the grid-points. Any program that can read this file type and render isosurfaces or volumetric pixels (Voxels) is suitable for visualization. We probed 3 different presentation methods for their suitability. A direct representation of the DID as an isosurface, a projection of the DID on the molecular density and as a last a representation using a Voxel engine. The first two depictions were realized with Pymol,<sup>[111]</sup> the Voxel graphics with the Paraview<sup>[112]</sup> software package. A schematic representation from the calculation to the graphical output for a benzene dimer is given in Figure 3.1.

Since one is often interested in coupling information on the interactions with the



**Figure 3.1:** Schematic representation of the different visualizations techniques to represent DIDs on the example of a parallel displaced benzene dimer.

molecular shape, the DID isosurfaces are often insufficient. In order to condense the data into a single representation, the DID values are projected onto an isodensity surface, obtained from the total molecular density. This leads to often times clearer pictures and easier interpretation. Another displayed possibility is the application of a Voxel engine. Thereby, each grid-point is rendered with a transparency and color gradient. The resulting graphics have more depth than the simple application of DID isosurfaces. Furthermore, the plotting of the molecular density often hinders the insight for crowded systems, like those featured in Chapter 5. The illustrations will be discussed further in the following Chapter.





# 4 Applications of Dispersion Interaction Densities

This Chapter is reproduced from the results and discussions and conclusions sections from Reference:

[55] A. Wuttke, R. A. Mata, Visualizing dispersion interactions through the use of local orbital spaces, *Journal of computational chemistry*, **2017**, *38*, 15-23 (DOI: 10.1002/jcc.24508),

with the permission of John Wiley & Sons, Inc. The Voxel representation was integrated and discussed in Sections 4.4, 4.5 and 4.6.

In Chapter 3 fragmentation using local orbital spaces on the basis of NPA and Dispersion Interaction Densities were introduced. To probe both methods different systems are investigated. In the first two Sections the benzene dimer and substituted derivatives are studied. The benzene dimer is a well-known benchmark system for non-covalent interactions and therefore, it is of interest if the method of choice, SCS-LMP2, performs well compared to highly accurate CBS results obtained by CCSD(T) and QCISD(T). Furthermore, the representation of DIDs by isosurfaces and projections on the molecular density as well as the influence of the orbital localization scheme are discussed.

In the case of the substituted benzene derivatives a comparison of local orbital analysis and F-SAPT is made and the sensitivity of both upper DID representation techniques is argued. The methylated anisole derivatives in Section 4.4 provide different binding sites. For the unsubstituted anisole a methanol can either dock at the ether oxygen or the  $\pi$ -ring. If methyl substituents are added to the anisole they are able to affect the methanol binding mainly due to dispersion interactions. It is of interest if the DIDs can display those weak interactions. In the last Section a

coupled diamondoid is examined. It is shown that local orbital analysis also gives the possibility to target intramolecular dispersion interactions.

### 4.1 Computational Details

The local correlation calculations were carried out with a development version of Molpro 2012.1<sup>[113]</sup> using the PAO-based density fitted LMP2 code.<sup>[114]</sup> The basis sets used are referenced in the respective Sections. If not mentioned otherwise the LMOs were obtained according to the Pipek-Mezey<sup>[83]</sup> procedure, and a NPA domain criteria (TNPA = 0.03e)<sup>[94]</sup> was applied in all cases. Density fitting approximations were used throughout for both reference and correlated calculations (the prefix ‘DF’ will be dropped for simplification). The images have been generated with the Pymol<sup>[111]</sup> and Paraview<sup>[112]</sup> program packages. To account for the overestimation of dispersion in the framework of MP2 spin-component scaling is applied to all calculations.<sup>[115]</sup>

### 4.2 Benzene Dimer

The benzene dimer is a common benchmark system for the study of dispersion forces. It serves as a model for  $\pi$ - $\pi$  stacking interactions, such as those found in DNA and aromatic protein residues. A large number of stationary points has been identified through theoretical studies of the potential energy surface (PES).<sup>[116–120]</sup> These are usually grouped under three main categories: the T-shaped (T), the sandwich (S) and the parallel-displaced (PD) structures. There have been differences in identifying the nature of each stationary point (as a minimum or saddle point), given the flatness of the PES. The stability of the dimers is determined by electrostatics (dominated by the quadrupole moment interaction) and dispersion, such that the conclusions may vary depending on the level of theory applied. But there is general agreement in the fact that the lowest lying minima are found for structures of the T or PD type. The objective is not to reassess the PES for the dimer, but instead to visually examine how the dispersion interactions change with the conformation. Five structures were taken from Reference<sup>[120]</sup> (Figure 4.1). These include two sandwich conformations denoted as S-A and S-B. In the S-A structure, the C-H bonds are intercalated. In the case of S-B, the benzene monomers are exactly aligned.

**Table 4.1:** Interaction energies (in  $\text{kJ}\cdot\text{mol}^{-1}$ ) for all five benzene conformers considered in this Section. Dispersion energies (in  $\text{kJ}\cdot\text{mol}^{-1}$ ) are given in parenthesis. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

	SCS-LMP2/A'VQZ	CCSD(T)/CBS	QCISD(T)/CBS	SAPT2/ADZ
$\Delta E_{\text{int}}(\text{PD})$	-11.12 (-19.89)	-11.42	-11.12	-11.00
$\Delta E_{\text{int}}(\text{T-A})$	-9.67 (-10.91)	-11.29		
$\Delta E_{\text{int}}(\text{T-B})$	-8.32 (-11.05)	-10.04	-11.22	-10.92
$\Delta E_{\text{int}}(\text{S-A})$	-7.36 (-16.12)	-7.15	-6.91	-7.11
$\Delta E_{\text{int}}(\text{S-B})$	-7.32 (-16.16)	-7.15		

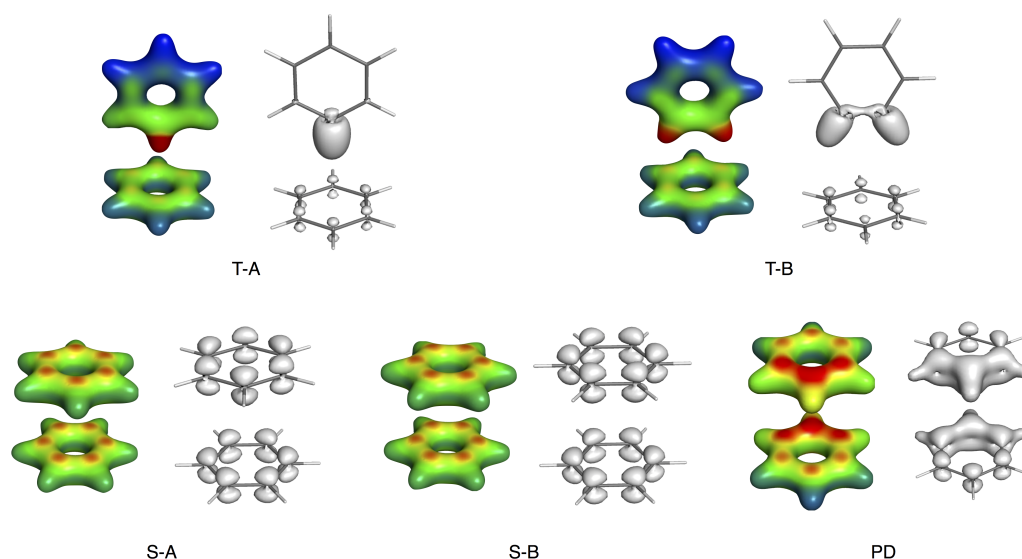
Furthermore, two T-type structures, T-A and T-B are considered. In the  $C_{2v}$  symmetric T-A structure the C-H bond of one monomer is pointing to the center of the other monomer ring. The T-B structure is obtained by a  $30^\circ$  rotation of the upper monomer, orienting a C-C bond perpendicular to the symmetry axis. Finally, there is the PD structure. In this case, the monomers are aligned along the C-H bonds. In Reference<sup>[120]</sup> the latter was identified as the only true minimum in this group of structures. A T-type dimer minimum can be found in a lower symmetry but this corresponds to a marginal stabilization.

The energies of the dimer and monomer were computed at the SCS-LMP2 level of theory, with an aug-cc-pVQZ basis for all atoms, except hydrogen, whereby the diffuse functions were removed.<sup>[121,122]</sup> The respective aug-cc-pVQZ/JKFIT and MP2FIT fitting basis were used.<sup>[123,124]</sup> The basis will be denoted as A'VQZ. Results for the total interaction energy are reported in Table 4.1, comparing to some selected theoretical works. All results have been computed without counterpoise correction. Given the size of the basis set and the reduced basis set superposition error (BSSE) inherent to local correlation methods,<sup>[125]</sup> this should be a small effect. The dispersion energy is also given for all conformers. The SCS-LMP2/A'VQZ results for the electronic interaction energy are in good agreement with previously reported values. The largest deviations are found for the T conformers, with SCS-LMP2 somewhat underestimating the interaction energy. Nevertheless, given the low computational cost of the approach, the agreement is quite reasonable.

The DID and DID projection plots in Figure 4.1 have been obtained as described in Chapter 3. The direct DID plots are displayed by selecting an isovalue of the DID itself. As the pictures reveal, these do not define a contact between the two molecules, they highlight the molecular densities giving the strongest contributions to the intermolecular dispersion energy, which is closer to the physical picture of the

interaction. In the case of DID projection plots, the molecular electronic density is used as a cut surface, and the DID values on the latter surface are represented. A linear color coding scale has been used. There is no specific reference to the magnitude of this scale, since this will be strongly influenced by the choice of isovalue for the molecular density. Such representations are only useful for relative comparisons, between different molecules or portions of a molecule. The ends of the spectrum have been labeled with 'd' for regions where there are large dispersion contributions and 'nd' where no interaction is taking place. The same scale was used for all featured dimers, also in the following Sections. Comparing both types of plots, although the direct DID representations deliver a clearer picture of the orbitals involved, the projected DID values are easier to interpret, as one can also see the molecular electronic density. Regions of strong and weak interactions are then distinguishable, and one can best compare the results among different dimers.

According to Table 4.1, the largest dispersion contributions are found in the PD and S-type conformers. This is also clear in the plots, with visible interaction areas in the rings. In the case of the sandwich conformers, the interaction is evenly distributed across the molecules. In the PD case, one can see that the overlapping regions, as

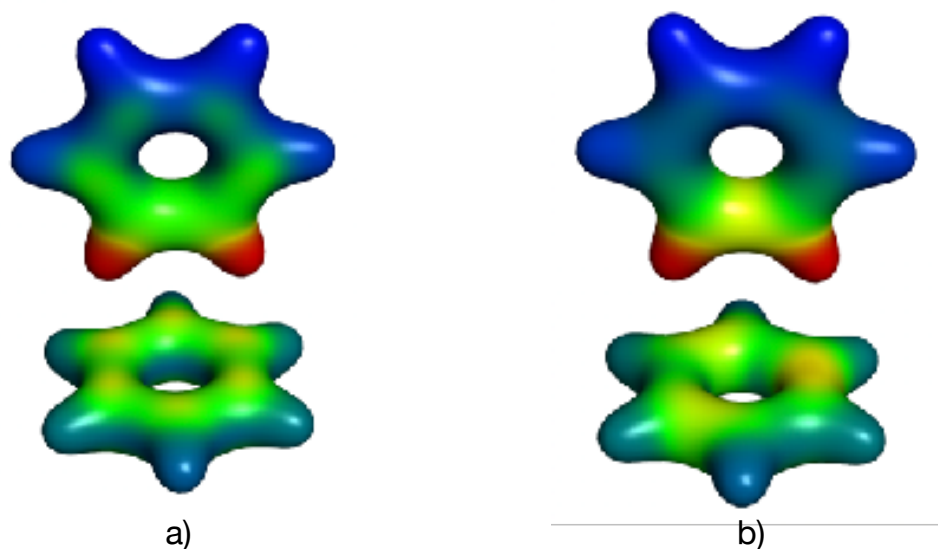


**Figure 4.1:** Projected DID plots (left) and DID plots (right) for five conformers of the benzene dimer. The naming convention is detailed in the text. DID projection plots: The molecular density isosurfaces were generated for a contour value of  $0.1 \text{ e/Bohr}^3$ . The DID color scheme ranges from highest value (red) to zero interaction energy (blue). DID plots: the isosurfaces were generated for a DID value of  $0.12 \text{ kJ}/(\text{mol} \cdot \text{Bohr}^3)$ . Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

expected, lead to larger dispersion forces. In the case of the T-type conformers, the interaction regions are significantly different. There is a strong contribution from the C-H bonds of one monomer which are closer to the ring of the other unit. The donor molecule, however, has little to no contribution from the ring. It is mostly an effect of the  $\sigma$ -orbitals. The acceptor ring itself shows reciprocally lower values than the ones found for other conformers. Overall, the DID plots reflect well the dispersion energy ordering, with  $PD > S-A \approx S-B > T-A \approx T-B$ .

One should note that the resolution of these representations is limited to the locality of the orbitals. Therefore, since all benzene orbitals are more or less symmetric relative to the ring plane (they are distorted due to the presence of the other monomer, but only slightly), the coloring on both sides of the molecule will be almost identical. This is particularly clear in the S and PD images. The faces of the rings which are opposite to the other monomer exhibit the same values as the interacting faces. One would expect that the density on the interacting side would have a larger weight. However, this would only be possible if different sets of orbitals were located on each side. From a chemical point of view, if the interaction is in fact dominated by  $\pi$ -electrons, there is nothing particularly wrong with this picture, since the latter are delocalized over both faces.

Another question would be the impact of the localization method chosen. The use of Pipek-Mezey orbitals allows for a chemical interpretation, since it keeps the separation of  $\pi$  and  $\sigma$  spaces. This particularly important feature when considering dispersion interactions. Nevertheless, this choice will still introduce some level of bias. In Figure 4.2 the results for the S-B dimer are compared using once Boys<sup>[82]</sup> and once Pipek-Mezey orbitals. The two plots are relatively similar. They both identify as top contributors to the dispersion interaction two C-H bonds for the benzene on top and the  $\pi$ -system in the case of the benzene below. However, there are subtle differences due to the way how  $\pi$ -orbitals are localized. It is noticeable on the lower benzene that the result from the Boys procedure strongly highlights three regions, while in the Pipek there is a delocalization of the interaction over the whole ring. These three regions correspond to the well-known banana-like Boys orbitals. The effect is also noticeable in the upper benzene, with the somewhat small highlighted region in the  $\pi$ -system being localized between two carbon atoms in the Boys case, but delocalized when using Pipek orbitals. The results are within expected. The detail in which the interactions are decomposed depends on how strictly localized the orbitals are. However, the overall picture does not change, and



**Figure 4.2:** Projected DID plots with a density contour value of  $0.1 \text{ e/Bohr}^3$ , for a) Pipek-Mezey orbitals, b) Boys orbitals. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

we have also found strong agreement between Pipek-Mezey and other localization schemes, namely NLMOs<sup>[94,126]</sup> and IBOs.<sup>[86]</sup> In such cases, the differences were negligible and a presentation of the plots is not required.

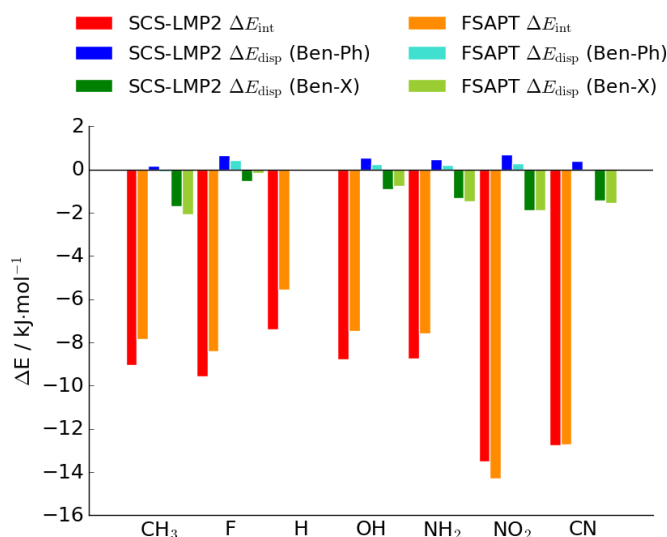
These results set the stage for further studies on benzene derivatives. In the following Sections two main aspects are considered: the impact of substituents in the  $\pi$ - $\pi$  interaction, and the competition between different interaction sites in the example of anisol/methanol dimers.

### 4.3 Substituted Benzene Dimer

As mentioned in Sections 2.5 and 3.2, extensions have been proposed to SAPT in order to distinguish between contributions of different fragments within a molecule. In this Section, the local partitioning scheme is compared to the functional-group partition of SAPT (F-SAPT) of Sherrill and coworkers.<sup>[32]</sup> A closer look is taken into the substituent effect on  $\pi$ -stacking interactions, using the same model systems as Parrish and Sherrill in a recent work.<sup>[58]</sup> By the analysis of different groups, one aims to better understand how different substituents in a benzene ring can affect the  $\pi$ -stacking interaction. Two different models have been suggested in the past. One

relates the changes to nonlocal effects. The ring density is altered due to the inclusion of a substituent, thereby affecting the  $\pi$ -stacking.<sup>[127]</sup> Another possibility is that the change in interaction energy is due to local effects, namely the direct interaction of the substituent and polarization of the  $\sigma$ -bond.<sup>[128]</sup> Both type of effects will impact electrostatic and dispersion interactions. Here, the focus is on the latter, and how the DID plots reflect the changes.

In this comparison the dimers: benzene/benzene, benzene/toluene, benzene/fluorobenzene, benzene/phenol, benzene/aniline, benzene/nitrobenzene and benzene/cyanobenzene are considered. The dimers will be labelled according to the added moiety in the substituted benzene (H, CH<sub>3</sub>, F, OH, NH<sub>2</sub>, NO<sub>2</sub> and CN respectively). It is distinguished between three quantities of interest. The first one will be the total interaction energy ( $\Delta E_{\text{int}}$ ). Secondly, the dispersion interaction of the phenyl moiety from the substituted benzene with the unsubstituted benzene ( $\Delta E_{\text{disp}}(\text{Ben-Ph})$ ). Finally, the dispersion interaction of the substituent and its  $\sigma$ -bond to the phenyl moiety with the unsubstituted benzene ( $\Delta E_{\text{disp}}(\text{Ben-X})$ ). These two last contributions summed up correspond to the total dispersion interaction in the dimer. The results are shown in Figure 4.3, including the F-SAPT/jun-cc-pVDZ values of Parrish and Sherrill.<sup>[58]</sup> In all calculations, and in order to ease the

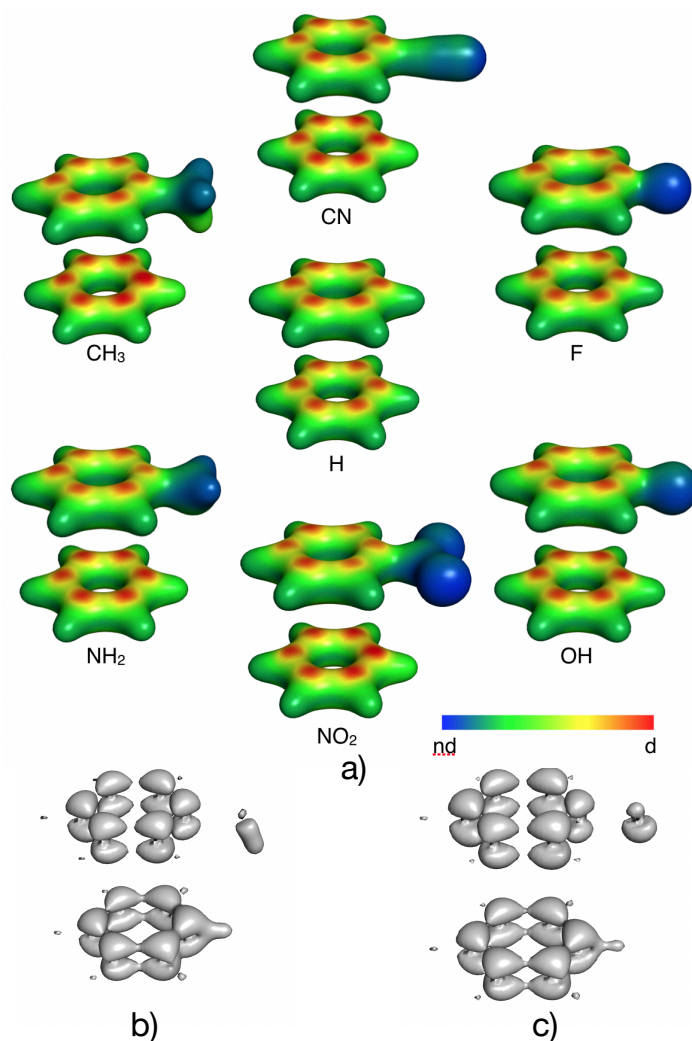


**Figure 4.3:** Interaction energies ( $\Delta E_{\text{int}}$ ), dispersion interactions between the two rings (Ben-Ph) and between the substituent (including the  $\sigma$ -bond) and the unsubstituted benzene (Ben-X) for a selection of dimers (in kJ·mol<sup>-1</sup>). Results are shown for a decomposition of SCS-LMP2/aug'-cc-pVTZ energies (this work) and F-SAPT/jun-cc-pVDZ. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

## 4 Applications of Dispersion Interaction Densities

comparison, the same geometries were used as in the cited study. The SCS-LMP2 calculations were carried out with the aug'-cc-pVTZ basis set (diffuse functions only included for non-hydrogen atoms).

In general, good agreement is observed between SAPT and SCS-LMP2 in the  $\Delta E_{\text{int}}$  values, both predicting the same ordering of stability for the dimers. SCS-LMP2 tends to give stronger binding energies (the only exception being the benzene/nitrobenzene dimer), but the difference is always below  $2 \text{ kJ}\cdot\text{mol}^{-1}$ . This again



**Figure 4.4:** a) Projected DID plots for the mixed benzene-substituted benzene dimers. The density isovalue was set at  $0.1 e/\text{Bohr}^3$ , the color coding for the DID is fixed for all representations and changes linearly from zero (nd) to highest value (d); b) DID plot for the benzene/toluene dimer (CH<sub>3</sub>) and c) DID plot for the benzene/nitrobenzene dimer (NH<sub>2</sub>). The DID isovalues used were  $0.12 \text{ kJ}/(\text{mol} \cdot \text{Bohr}^3)$ . Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.



validates the choice of SCS-LMP2 as electronic structure method. The trends in the dispersion energy, again, are the most relevant quantities and just as before we see a close agreement between both approaches. The local partitioning and F-SAPT reveal the same pattern for all dimers. Most substituent groups tend to slightly decrease the direct interaction between the two rings. The substituents with the strongest dispersion interactions to the unsubstituted benzene are  $\text{CH}_3$  and  $\text{NO}_2$ , closely followed by  $\text{CN}$  and  $\text{NH}_2$ . Looking at the projected DID plots in Figure 4.4, it is hard to distinguish any difference between the various systems. This is within expected, given the small variations in the dispersion energy. Still, one is able to recognize that the increase in  $\Delta E_{\text{disp}}$  for toluene ( $\text{CH}_3$ ) is mostly due to the close proximity of a C-H bond to the benzene. In the other cases, the analysis is somewhat harder, but still it can be seen that in fluorobenzene the direct dispersion interaction with the substituent is almost absent, in agreement with the results of Figure 4.3.

For completeness, also the DID plots in the  $\text{CH}_3$  and  $\text{NH}_2$  cases are shown, where the difference in the interaction is dominated by a  $\sigma$ -bond from the substituent. The plots illustrate this effect rather well. This reveals that for weak local interactions, the direct representation as DID isosurface is a good choice.

## 4.4 Methylated Anisole-Methanol Dimers

In a recent contribution,<sup>[129]</sup> Suhm and coworkers probed molecular dimers of anisole derivatives with methanol through IR-spectroscopy of cold mixed clusters generated in a supersonic jet expansion. In combination with these experiments, they carried out calculations to compare and determine the conformational preference of the most stable dimers. There are two main interaction sites for methanol: the ether oxygen, building a OH-O hydrogen bridge and the  $\pi$ -system, whereby the methanol OH group is directed towards the  $\pi$ -cloud of the anisole derivative. The conformational preference can be inferred from the OH stretching fundamental, with OH-O coordination exhibiting larger shifts than OH- $\pi$ . Methylated variants of anisole have been used in the experiment, with substitutions at the ortho-, para- and meta-positions. This results in a large set of data to which quantum chemical methods can be benchmarked, evaluating how accurate the energy differences between the two bonding motifs are described. Again the focus of the analysis lies on dispersion.

## 4 Applications of Dispersion Interaction Densities

**Table 4.2:** Electronic interaction energies ( $\Delta E_{\text{int}}$ ) and dispersion energies ( $\Delta E_{\text{disp}}$ ), given in  $\text{kJ}\cdot\text{mol}^{-1}$ , for anisole derivatives interacting with methanol. The substituent groups (X) included in anisole are given in the first column. The dispersion energies (interaction to the methanol molecule) are divided between the anisole fragment and the substituents X. The first two values (X='none') correspond to anisole. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

X	bond	$\Delta E_{\text{int}}$	$\Delta E_{\text{disp}}$		
			Anisole	X	Total
none	OH-O	-19.10	-10.86	—	-10.86
	OH- $\pi$	-16.89	-13.30	—	-13.30
2-Me	OH-O	-19.58	-11.73	-2.88	-14.61
	OH- $\pi$	-17.74	-13.37	-1.39	-14.76
3-Me(E)	OH-O	-19.85	-10.98	-0.32	-11.30
	OH- $\pi$	-17.44	-13.40	-0.89	-14.29
3-Me(Z)	OH-O	-19.45	-10.82	-0.05	-10.87
	OH- $\pi$	-17.68	-13.45	-0.96	-14.40
4-Me	OH-O	-19.60	-9.66	-1.25	-10.91
	OH- $\pi$	-17.37	-13.50	-0.78	-14.27
2,3-Me	OH-O	-20.68	-12.05	-3.20	-15.25
	OH- $\pi$	-19.05	-13.49	-2.39	-15.87
2,6-Me	OH-O	-23.37	-10.78	-3.42	-14.20
	OH- $\pi$	-19.27	-14.18	-2.23	-16.41
3,5-Me	OH-O	-20.15	-10.97	-0.37	-11.34
	OH- $\pi$	-18.71	-13.48	-1.62	-15.10
2,3,5-Me	OH-O	-21.07	-12.15	-3.33	-15.48
	OH- $\pi$	-20.07	-13.75	-2.72	-16.47
4-tBut	OH-O	-19.69	-11.00	-0.16	-11.16
	OH- $\pi$	-18.91	-14.14	-3.89	-16.03

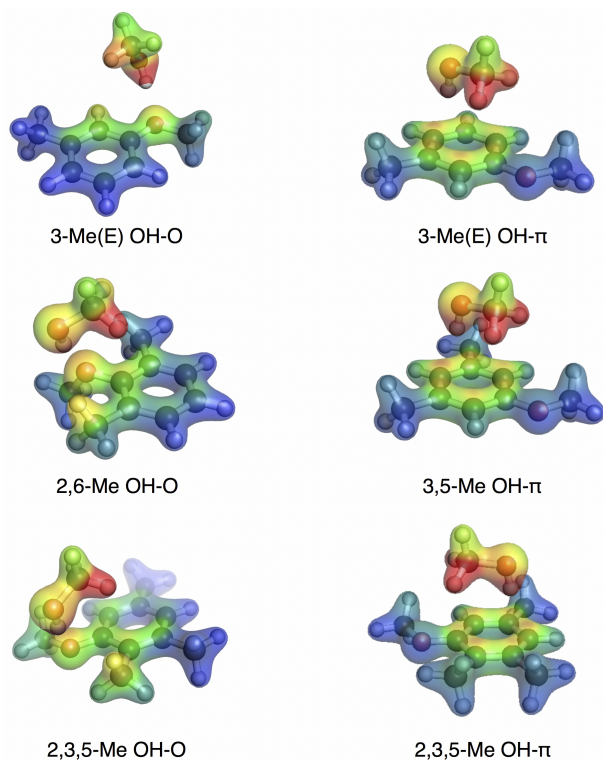
With growing molecular system size, single interaction sites lose significance, as dispersion stabilization can be provided by any added groups. It becomes not only a competition between the ether oxygen and the  $\pi$ -ring, but overall the electron clouds of both monomers. For the application of DIDs it is of huge interest if the DID plots reflect those changes, whether the larger systems show across the board a dominance of the  $\pi$ -electrons, or if the gradual increase in the substituents size is enough to strongly influence the conformation of methanol.

For all systems taken from the aforementioned study the SCS-LMP2/aug'-cc-pVQZ interaction energies were computed. The data is provided in Table 4.2. The full set of structures features 30 different structures for anisole and 8 different derivatives thereof. Also included in the Table is the total dimer dispersion interaction energy,

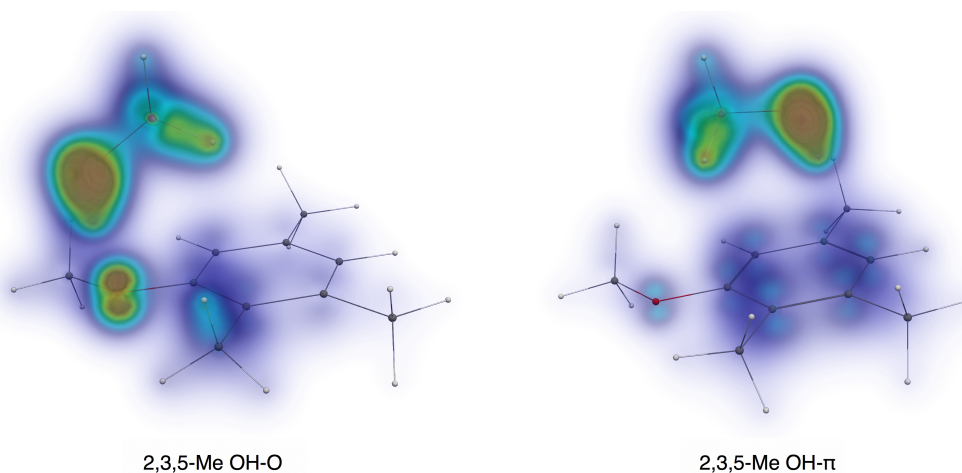
the contribution of anisole (meaning the direct interaction of the anisole fragment with methanol) and the contribution of the substituents (generally labeled as X) to the dimer dispersion. In order to obtain these two separate contributions, use is made of an improved criterium for splitting the energies as described in Section 3.1 by Equation (3.3). The contributions of the covalent bonds between anisole and the substituents were split according to the natural population analysis of the orbital. In all cases the localized orbital was an almost pure C-C bond, so the total orbital charge on each carbon was used to weight the amount of dispersion belonging to the anisole and the substituent.

The decomposition between the main anisole unit and the substituents reveals that the substituents are still too small to dominate the dispersion interaction. The largest substituent contributions are, as expected, observed for the tert-butyl system, but these are still only one quarter of the total dispersion interaction. Similar values are also observed for di- and tri-methylated anisoles. One particular interesting case is the 2,6-dimethylanisole OH-O conformer. Although only two methyl groups have been included, their proximity to the ether oxygen leads to a  $3.42 \text{ kJ}\cdot\text{mol}^{-1}$  dispersion stabilization through the substituents.

The projected DID plots in Figure 4.5 reflect the aforementioned trends. In this Figure, the profiles are combined with a ball-and-stick representation of the molecules. Six plots are included, for which direct comparison is of interest. The DID values for the OH-O and OH- $\pi$  conformers of the methanol/3-Me-anisole show a clear picture of the dispersion interaction. In the OH-O case, the 3-Me-anisole interacts through the electron cloud of the ether moiety, in the OH- $\pi$  case the ring system is highlighted. Adding a further methyl group to the ring (3,5-Me) has little effect on the dispersion interactions for the OH- $\pi$  conformer. This can also be seen in Table 4.2. The projected DID plot of the 2,6-Me-anisole OH-O conformer is represented in the same Figure. As previously mentioned, and confirmed in the plot, the larger  $\Delta E_{\text{disp}}(X)$  value is due to a methyl close to the interacting ether oxygen. Finally, the two conformers for the trimethyl substituted system are compared. The OH- $\pi$  conformer shows little differences relative to the previous OH- $\pi$  conformers. Table 4.2 shows that the  $\Delta E_{\text{disp}}$  contribution of anisole is  $-13.4$ ,  $-13.5$  and  $-13.8 \text{ kJ}\cdot\text{mol}^{-1}$ , for 3-Me, 3,5-Me and 2,3,5-Me, respectively. The DID profiles seem consistent with this result. However, the dispersion interaction in total raises slightly along the series. The project DID plot does not capture these few  $\text{kJ}\cdot\text{mol}^{-1}$  difference since it is a small cumulative effect over a large region of space. As shown in Fig-



**Figure 4.5:** DID projected plots of methanol and anisole derivative dimers for a DID value of  $0.12 \text{ kJ}/(\text{mol} \cdot \text{Bohr}^3)$ . Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.



**Figure 4.6:** Voxel DID plots of methanol and 2,3,5-methylated anisole dimers.

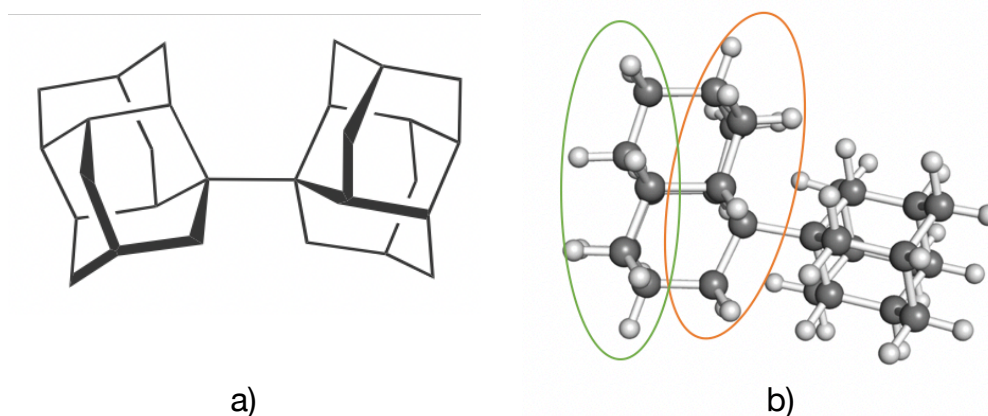
ure 4.6 based on the example of 2,3,5-Me, Voxel DIDs are able to recover also these marginal differences. In case of the O-dimer a strong interaction between methanol and the ether oxygen is observed, while the interaction to the  $\pi$ -ring is displayed as very small and the two outer methyl groups do not contribute to the stabilization. For the  $\pi$ -dimer the ring shows stronger interactions and the contributions of the

$\pi$  orbitals become clearly visible. All substituted methyl groups are involved in the dispersive binding.

## 4.5 Coupled Diamondoids

As a last application system a coupled diamantane-diamantane molecule is investigated. The latter was synthesized and characterized by Schreiner *et al.*<sup>[69]</sup> The Lewis structure of the latter is given in Figure 4.7 a). Two diamantane moieties are connected through a particularly long single C-C bond, with dispersion interactions playing a determining role in the thermodynamic stabilization of this product.

One difficulty in characterizing such a system is that the two interacting fragments are actually covalently bound. One is then forced to look into model systems or the bond dissociation enthalpy, but these do not tap directly to the properties of the molecule. By the use of local orbital spaces, the connecting C-C bond can be removed from the analysis in order to calculate the DID directly, including only orbitals strictly localized on each side. The latter was carried out once again at the SCS-LMP2 level, with the aug-cc-pVTZ basis set for all atoms, except hydrogen whereby the diffuse functions were removed.<sup>[122]</sup> The plot is shown in Figure 4.8 a). One can clearly observe that the major interaction terms arise from the first



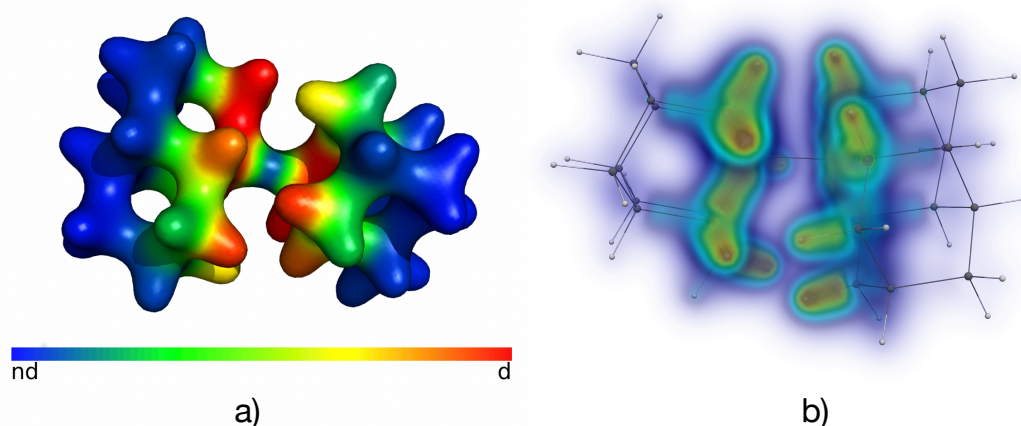
**Figure 4.7:** a) Lewis structure of the coupled diamantane-diamantane molecule, b) definition of two layers within a single diamondane unit. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

## 4 Applications of Dispersion Interaction Densities

layer, the atoms closer to the elongated C-C bond. This supports the observations of Schreiner and coworkers.

The total dispersion interaction between the fragments amounts to  $-73.1 \text{ kJ}\cdot\text{mol}^{-1}$ . This value can be compared to the difference between the B3LYP and B3LYP-D results published in the original work,<sup>[69]</sup> of  $-112.1 \text{ kJ}\cdot\text{mol}^{-1}$ . This would indicate an overestimation of the dispersion energy by the D correction. However, the latter should be taken as a qualitative measure, rather than a quantification of the effect. The D correction is a rather early approach, which has been improved over the years. The newer D3 formulation should be more reliable, but even then there is the issue of separating close- and long-range dispersion contributions.<sup>[130]</sup> The former are already to some extent present in conventional DFT functionals.

As noted, the project DID plot indicates that the major contribution comes from the first contact shells. One can also quantify these by further splitting the molecule as shown for one of the diamondoid units in Figure 4.7 b) and applying the fragmentation scheme as outlined in Section 3.1. One can then identify atoms in a first layer, close to the C-C bond, and a layer lying further away (layer 2). The dispersion interaction between the first layers on each side (1-1) amounts to  $-57.0 \text{ kJ}\cdot\text{mol}^{-1}$ , which is expected given the information contained in the DID plot. The outer layers (2-2 interaction) only contribute  $-1.4 \text{ kJ}\cdot\text{mol}^{-1}$  and contacts between the first layer and the second layer on the other side (1-2 interactions) sums up to  $-14.7 \text{ kJ}\cdot\text{mol}^{-1}$ .



**Figure 4.8:** a) DID projection plot for a DID value of  $0.12 \text{ kJ}/(\text{mol} \cdot \text{Bohr}^3)$  and b) Voxel DID of the molecule, considering the two diamondane units as the interacting fragments. Reproduced from Reference [55] with the permission of John Wiley & Sons, Inc.

While comparing the projected DID and the Voxel DID in Figures 4.8 a) and b), it is observed that the Voxel representation is again more sensitive to small dispersion contributions. While in the projected DID the second layer displays no dispersive interactions, the Voxel representation captures these contributions through a blueish tone.

## 4.6 Conclusions

In order to probe the usefulness of Dispersion Interaction Densities as a visualization tool, different systems were investigated. The quantity itself, being proportional to a local electronic density and its contribution to the London forces acting on neighboring atoms, can be applied with minimal extra computational work. It highlights regions of space where dispersion is strongly stabilizing interfragment contact. Given that the partitioning is carried out solely at the orbital level, the way results can be visualized is rather flexible. The fragments can be defined readily as intra- or intermolecular, and one can also focus on specific moieties of the total molecule. No embedding, models or extra partitioning schemes are required.

In the applications featured in this Chapter, use was made of spin-component scaled local MP2, which proved for most systems to be in close agreement with SAPT results. Furthermore, even for the benzene dimer, SCS-LMP2 performed rather well compared to the computationally much more demanding Coupled-Cluster and Configuration Interaction approaches. Also the dispersion fractions of the total interaction energies were found to be consistent to F-SAPT, when a comparison was available. These results are encouraging, since the two procedures are based on an interpretation of the same operator (two single excitations occurring on different interacting fragments). They should, if the interaction is correctly described, ideally converge.

There are, however, limitations to the approach. The degree of detail depends on the locality of the orbitals and the final representation also depends on the particular method chosen for representing the occupied space. The preferred option so far has been the use of Pipek-Mezey orbitals, since the  $\sigma$ - $\pi$  separation is kept, allowing for a chemical intuitive interpretation. Although it was found that the overall information

is not too sensitive to this choice, it does influence the distribution, as the whole procedure depends on the shape of the orbitals.

Regarding the representation technique one has the agony of choice. The right selection strongly depends on the system and all three options have advantages and disadvantages. For the most purposes the Voxel DID is preferred. It is most sensitive to small contributions and gives a detailed insight to the orbitals involved in the interaction. In Chapter 5, when supramolecular host-guest systems will be discussed this representation is advantageous, targeting weak contributions arising through the whole pocket. Furthermore, Voxels allow to treat all systems with the same display settings. The projected DIDs as well as the DID isosurfaces suffer here. Considering one weakly interacting and one strongly interacting system, in case of DID isosurfaces this would end in to small or to large contours, respectively. For the projected DID in one case there will be to much highlights and in the other to little.



# 5 Supramolecular Coordination Cages

This Chapter is partly reproduced from References:

[59] S. Löffler, J. Lübben, A. Wuttke, R. A. Mata, M. John, B. Dittrich, B. H. Clever, Internal dynamics and guest binding of a sterically overcrowded host, *Chemical Science*, **2016**, *7*, 4676-4684 (DOI: 10.1039/C6SC00985A),

[61] S. Löffler, A. Wuttke, B. Zhang, J. J. Holstein, R. A. Mata, H. G. Clever, Influence of size, shape, heteroatom content and dispersive contributions on guest binding in a coordination cage, *Chemical Communications*, **2017**, *53*, 11933-11936 (DOI: 10.1039/C7CC04855F),

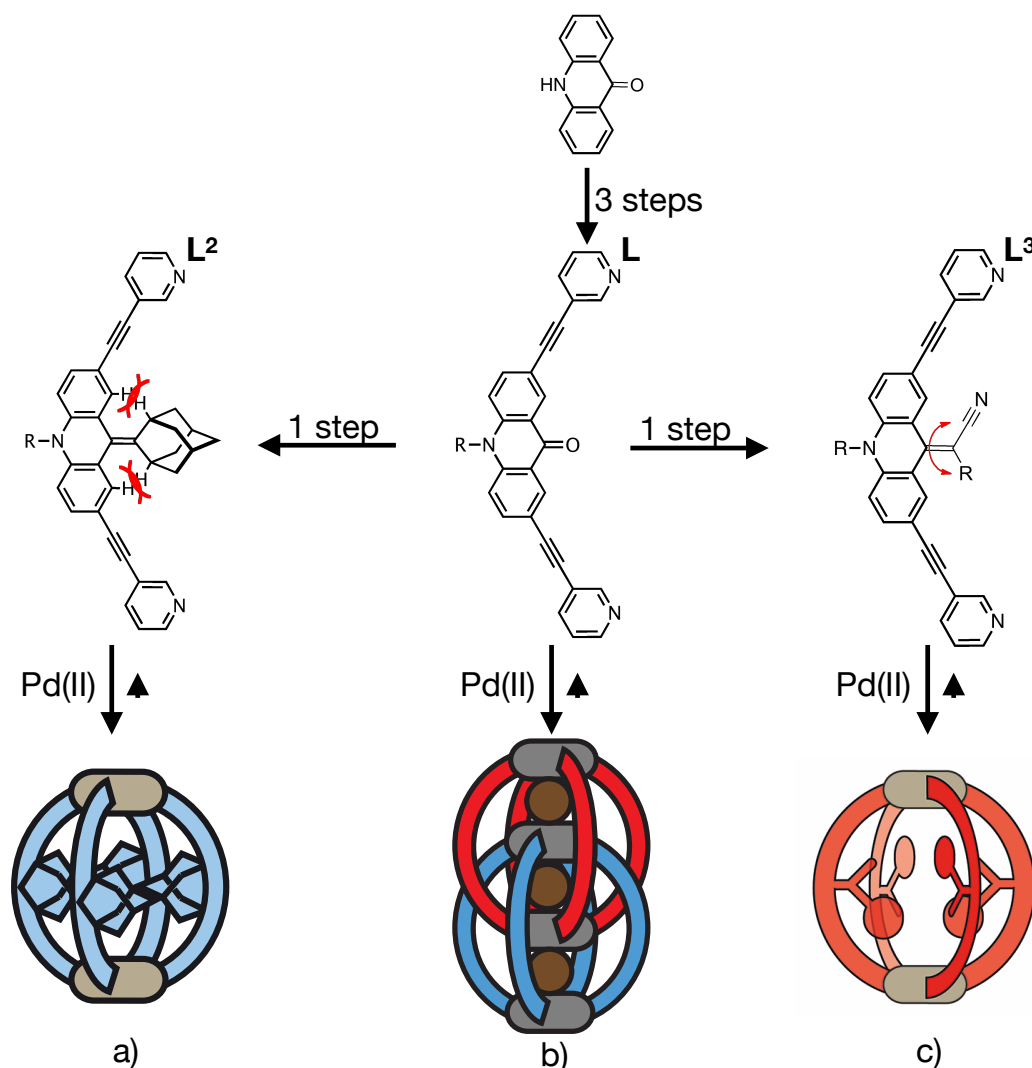
with the permission of the Royal Society of Chemistry. The description of the experimental procedures were provided by Clever and coworkers. All computational results were calculated on my own.

The binding of small molecules inside the nanoscopic cavities of natural or artificial hosts such as enzymes or supramolecular cavitands is far from being fully understood. For example, the roles of shape match, conformational restrictions, steric repulsion vs. dispersive attraction as well as desolvation processes are still under debate.<sup>[131–133]</sup> Synthetic supramolecular chemistry is able to supply a plethora of container systems of variable size, shape and chemical makeup as model systems helping to understand non-covalent binding in nature.<sup>[134–137]</sup> The metal-mediated self-assembly of coordination cages<sup>[138–140]</sup> under modular variation of ligand scaffolds and functionalities has yielded sophisticated bottom-up nano systems such as switchable receptors,<sup>[141–143]</sup> selective catalysts,<sup>[144–147]</sup> and light-driven charge separators.<sup>[148]</sup> Guest binding in these systems has been studied intensively with ionic guests being of opposite charge than the cage, where Coulombic interactions dominate the driving force for encapsulation.<sup>[149,150]</sup> Less explored is the binding of neutral guest species inside ionic cages.<sup>[151,152]</sup> Here, solvophobic effects together

## 5 Supramolecular Coordination Cages

with enthalpically beneficial host-guest contacts seem to play a major role but further systematic experimental and theoretical efforts are required to shed light on the situation in selected hosts.

Through the substitution of two 3-ethynylpyridines to an acridone the Clever group was able to synthesize the banana shaped ligand **L** shown in Figure 5.1 b). By a self-assembly process in the presence of  $[\text{Pd}(\text{CH}_3\text{CN})_4]-(\text{BF}_4)_2$  an interpenetrated double cage with three pockets filled by  $\text{BF}_4^-$  ions is obtained. This system can be activated through the exchange of the outer pocket ions by  $\text{Cl}^-$  or  $\text{Br}^-$  to bind



**Figure 5.1:** Schematic representation of the pathways to synthesize supramolecular cages starting from acridone. The addition two 3-ethynylpyridines to the latter leads to the bis-monodentate pyridyl ligand **L**. This ligand can be further functionalized by the substitution of adamantyl (**L**<sup>2</sup>) or 2-cyanoacetic esters (**L**<sup>3</sup>) in the carbonyl position. In the presence of  $[\text{Pd}(\text{CH}_3\text{CN})_4]-(\text{BF}_4)_2$  the cages are formed in a self-assembly process.

neutral guest molecules in the middle pocket. The theoretical derivation of free binding energies for such systems is a challenging task, since several effects have to be accounted for. Since the guests are neutral it seems obvious that dispersion plays a crucial role in the stabilization within the pocket. Local correlation calculations were carried out in order to prove this hypothesis.

A further modification of **L** by the substitution of adamantyl groups in the carbonyl position, lead to a sterically overcrowded coordination cage as shown in Figure 5.1 a). Different ionic guest molecules are able to bind in its pocket and it is of interest if dispersion has a determining role, since one would suspect that electrostatics are the dominant driving force. Furthermore, the system revealed interesting dynamics which are discussed in detail in Reference [59].

By the substitution of 2-cyanoacetic esters to the carbonyl position of **L** the ligands **L<sup>3</sup>** shown in Figure 5.1c) were obtained. NMR experiments revealed that the free ligands rotate around the double bond, connecting the backbone and the esters. To describe this behavior DFT calculations were carried out to simulate these pathways. Since this thesis mainly aims the description of non-covalent interactions, the interested reader is referred to Reference [60].

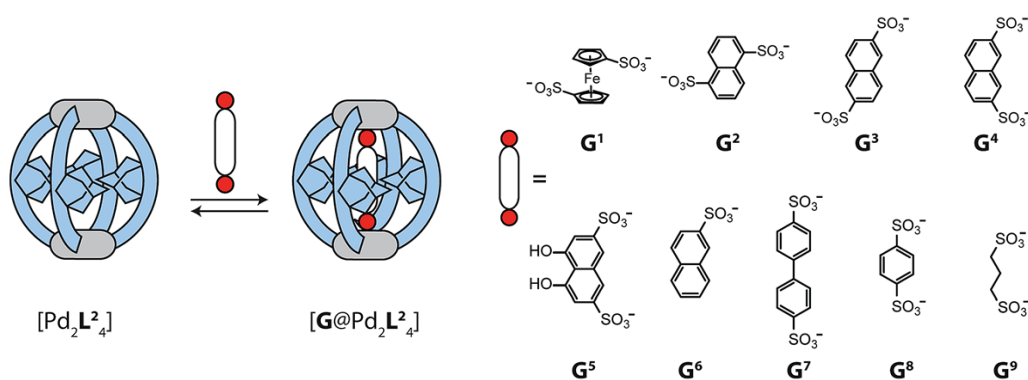
## 5.1 A Sterically Overcrowded Cage

Comparing the ligands **L** and **L<sup>2</sup>**, as shown in Figure 5.1, the steric bulk in the centre of **L<sup>2</sup>** is dramatically increased by incorporating the doubly-bound adamantylidene group. In a cooperative study with the Clever group the capability of forming  $[\text{Pd}_2\text{L}_4^2]$  self-assembled cages in which the adamantyl groups serve as dispersion energy donors was investigated. Cationic palladium centres are arranged at opposite ends across the globular cavity. Rod-shaped bis-anionic guests carrying a negative charge at either end are encapsulated in such  $[\text{Pd}_2\text{L}_4^2]$  cages in a way that the guests major axis is collinear with the Pd-Pd axis, thereby bringing the anionic centres as close as possible to the metal cations.<sup>[153]</sup> In order to probe the interactions within the sterically crowded equatorial area of the cage  $[\text{Pd}_2\text{L}_4^2]$ , the uptake of various bis-anionic guests (as shown in Figure 5.2) with respect to their relative lateral bulk is studied.

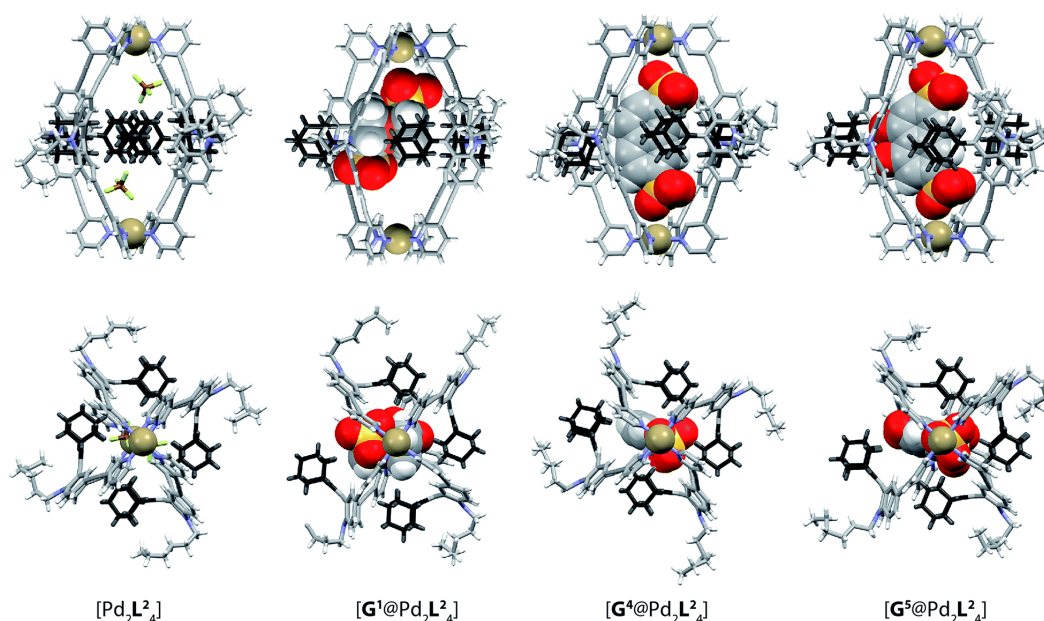
## 5.1.1 Summary of Experimental Findings

The mono- and bis-anionic guests  $\mathbf{G}^1$ - $\mathbf{G}^9$  were titrated into a solution of the cage  $[\text{Pd}_2\mathbf{L}_4^2]$  in acetonitrile. It was initially expected that the tremendous steric bulk of the four equatorial adamantyl groups would only allow for guests with a thin backbone (i.e. an alkyl chain) to bind inside the cage. Surprisingly, even quite bulky guests are bound to form host-guest complexes  $[\mathbf{G}@\text{Pd}_2\mathbf{L}_4^2]$ . This was verified by NMR spectroscopy, high-resolution ESI mass spectrometry and single crystal X-ray analysis. Due to the substantial size of the guests, the guest molecules are in close contact with the surrounding adamantyl residues in the equatorial region of the cavity as can be observed in the X-ray structures (Figure 5.3).

In order to obtain a measure of the binding strength one equivalent of the guest molecules has been added into a solution of the cage  $[\text{Pd}_2\mathbf{L}_4^2]$  in acetonitrile. For the guest molecules  $\mathbf{G}^1$ - $\mathbf{G}^5$  the signals of the empty  $[\text{Pd}_2\mathbf{L}_4^2]$  cage disappeared and only signals of the host-guest systems  $[\mathbf{G}^{1-5}@\text{Pd}_2\mathbf{L}_4^2]$  remained in the  $^1\text{H}$ -NMR spectra. Hence, guests  $\mathbf{G}^1$ - $\mathbf{G}^5$  are strongly bound inside the cage and the guest exchange with the solvent exterior is slow on the NMR time scale. In contrast, addition of guests  $\mathbf{G}^6$ - $\mathbf{G}^9$  to the cage resulted in a gradual shifting of the cages  $^1\text{H}$ -NMR signals, indicating a fast exchange of these guests. Since more than one equivalent of guests  $\mathbf{G}^6$ - $\mathbf{G}^9$  was required to saturate the cages, these guests bind more weakly to the cage.



**Figure 5.2:** Structures of various guest molecules encapsulated by coordination cage  $[\text{Pd}_2\mathbf{L}_4^2]$ . Reproduced from Reference [59] with the permission of the Royal Society of Chemistry.



**Figure 5.3:** X-ray crystal structures of  $[\text{Pd}_2\text{L}_4]$ ,  $[\text{G}^1@\text{Pd}_2\text{L}_4]$ ,  $[\text{G}^4@\text{Pd}_2\text{L}_4]$  and  $[\text{G}^5@\text{Pd}_2\text{L}_4]$ . Color scheme: C, grey; N, blue; O, red; S, yellow; F, green; B, brown; Pd, tan; Fe, orange. For clarity, solvent molecules and free (non encapsulated) anions are omitted. Reproduced from Reference [59] with the permission of the Royal Society of Chemistry.

These results seem to be a consequence of the different size and shape of the guest structures. Guest molecules  $\text{G}^1$ - $\text{G}^5$  are bis-anionic and have a good size match for encapsulation,<sup>[154]</sup> while the other guests are either too small ( $\text{G}^8$  and  $\text{G}^9$ ), too large ( $\text{G}^7$ ) or lack a second sulfonate group ( $\text{G}^6$ ) to be encapsulated with high affinity inside the coordination cage.

### 5.1.2 Theoretical Investigations

In order to obtain further insight into the binding of the guest molecules, in particular the role of dispersion forces, electronic structure calculations were carried out for  $[\text{Pd}_2\text{L}_4]$ ,  $[\text{G}^4@\text{Pd}_2\text{L}_4]$  and  $[\text{G}^5@\text{Pd}_2\text{L}_4]$ . Due to the large size of these systems, the calculations involving the binding pocket were limited to the four adamantane moieties (with the covalent double bond to the ligand backbone cut and capped with two hydrogen atoms). The adamantane fragments and guest compounds were optimized at the B3LYP-D3/def2-TZVP level of theory<sup>[130,155,156]</sup> (with Becke-Johnson type damping)<sup>[157,158]</sup> using the ORCA software.<sup>[159]</sup> Afterwards the optimized geometries were overlapped with the crystal structure positions. These should corre-

**Table 5.1:** Interaction and dispersion energies (in  $\text{kJ}\cdot\text{mol}^{-1}$ ) between different guest molecules and the saturated adamantane residues of the  $[\text{Pd}_2\mathbf{L}_4^2]$  cage at the SCS-LMP2/A'VTZ level of theory.

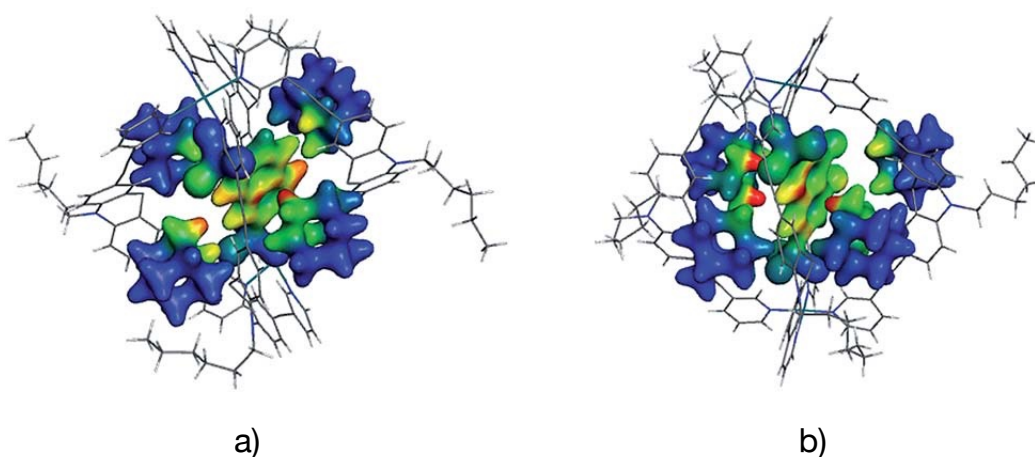
$[\text{X}@\text{Pd}_2\mathbf{L}_4^2]$	$\Delta E_{\text{int}}$	$\Delta E_{\text{disp}}$
$\mathbf{G}^4$	-52.8	-86.3
$\mathbf{G}^5$	-67.2	-64.0

spond to a suitable average conformation of the ligands and guest molecules. Also, given the small space available in the pocket, dynamic fluctuations in the structure should not affect the results severely. On the basis of the structures obtained, electronic structure calculations were carried out at the density fitted spin-component-scaled local second-order Møller-Plesset perturbation theory DF-SCS-LMP2<sup>[114,115]</sup> using the aug'-cc-pVTZ (A'VTZ) basis set (diffuse functions on all atoms, except hydrogens)<sup>[121,122]</sup> implemented in the Molpro 2015.1<sup>[110]</sup> program package (for simplification the prefix 'DF' will be dropped in the remaining text). All interaction energies are solely electronic. The total energies are given according to the supermolecular approach, with fixed host and guest structures. The dispersion energies were obtained through the decomposition of the SCS-LMP2 pair energies.<sup>[47]</sup>

Calculations in the system  $[\text{Pd}_2\mathbf{L}_4^2]$  were primarily intended to estimate the stabilization of the cage brought forth by dispersion interactions between the adamantane moieties in the pocket. This system was chosen since it contains the smallest guest ( $\text{BF}_4^-$ ), allowing for the adamantane moieties to come closest. For comparison the  $[\mathbf{G}^4@\text{Pd}_2\mathbf{L}_4^2]$  cage was also considered. Here, the X-ray analysis revealed that the adamant-moieties are shifted outwardly. This should depend on steric repulsion between the latter and the guest molecule taken up in the pocket. The SCS-LMP2/aug'-cc-pVTZ calculations gave a dispersion contribution of  $-14.3 \text{ kJ}\cdot\text{mol}^{-1}$  for the  $[\text{Pd}_2\mathbf{L}_4^2]$  system. The dispersion contribution in the  $[\mathbf{G}^4@\text{Pd}_2\mathbf{L}_4^2]$  system was calculated to be  $-9.3 \text{ kJ}\cdot\text{mol}^{-1}$ , somewhat smaller given the larger distances between the four ligands that result from the guest-induced expansion of the cage structure. This is not a particularly large effect and although it will counterbalance the steric repulsion of the bulky adamantyl moieties, it should not be a major factor in the cage formation itself.

In the calculations of the host-guest complexes, only the interactions of the guests with the adamantanes were considered, disregarding the interaction with the re-

maining parts of the ligands and the Pd ions. The latter would be particularly difficult to estimate given that the electrostatics will be strongly influenced by the solvent model used. The remaining parts of the ligands should have no strong interaction with the guests considered. Interestingly, dispersion forces seem to play a significant role in the binding of the guest molecules inside the cages interior. The overall interaction energies between the adamantyl residues and the guests (excluding the cation-anion interactions) for the systems  $[\mathbf{G}^4@Pd_2\mathbf{L}_4^2]$  and  $[\mathbf{G}^5@Pd_2\mathbf{L}_4^2]$  were computed as  $-52.8 \text{ kJ}\cdot\text{mol}^{-1}$  and  $-67.2 \text{ kJ}\cdot\text{mol}^{-1}$ . The dispersion energy contributions obtained are  $-86.3 \text{ kJ}\cdot\text{mol}^{-1}$  and  $-64.0 \text{ kJ}\cdot\text{mol}^{-1}$  for  $[\mathbf{G}^4@Pd_2\mathbf{L}_4^2]$  and  $[\mathbf{G}^5@Pd_2\mathbf{L}_4^2]$ , respectively (see Table 5.1). These correspond to a strong contribution in binding, even exceeding the total interaction in the case of  $[\mathbf{G}^4@Pd_2\mathbf{L}_4^2]$ . In  $[\mathbf{G}^5@Pd_2\mathbf{L}_4^2]$ , the adamantyl units are further apart, reducing the relative weight of the dispersion forces. Consequently, it is anticipated that the main contributions for the adamantyls interaction with the guests to be Pauli repulsion and dispersion, the two of opposite sign. These results clearly show that dispersion forces can easily add up to large values in such a supramolecular construct. To better visualize this effect, the Dispersion Interaction Densities (DIDs) calculated between the guest and the four adamantane units, are shown in Figure 5.4 (the plots were prepared with Pymol<sup>[111]</sup>). The images show that the  $\pi$ -systems of the guests interact strongly with the surface of the pocket. In particular, some hot spots are identifiable where the adamantyl groups are in close contact with the encapsulated guests.



**Figure 5.4:** Dispersion interaction densities (DIDs) between the guests and the surrounding adamantyl groups in a)  $[\mathbf{G}^4@Pd_2\mathbf{L}_4^2]$  and b)  $[\mathbf{G}^5@Pd_2\mathbf{L}_4^2]$  (red: high DID; blue: low DID). Reproduced from Reference [59] with the permission of the Royal Society of Chemistry.

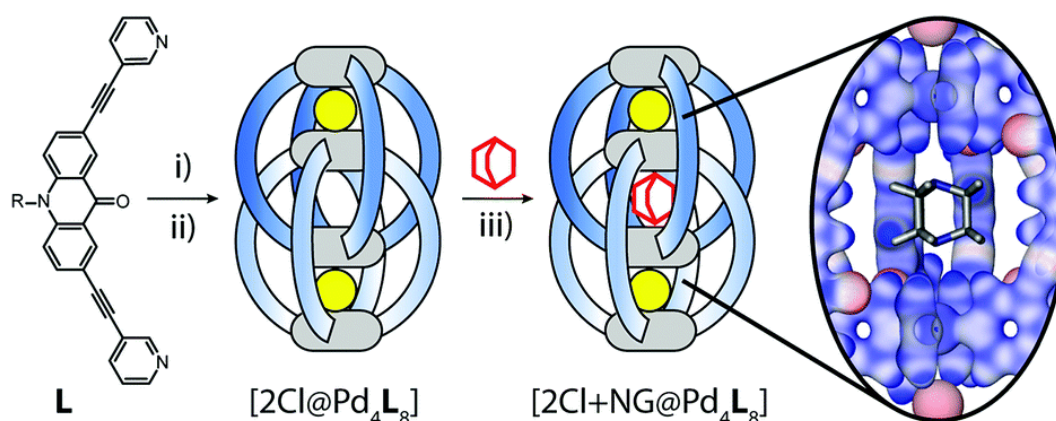
## 5.2 An Interpenetrated Double-Cage

In recent publications, the Clever group and cooperation partners reported on an interpenetrated double cage featuring three consecutive binding pockets, filled with tetrafluoroborate anions.<sup>[160]</sup> These host molecules can be activated by the addition of chloride anions to bind neutral guests such as benzene, cyclohexane and norbornadiene.<sup>[161]</sup> A schematic representation is shown in Figure 5.5. While the effect of halide binding has been well explained as an allosteric contraction of the whole architecture along the Pd<sub>4</sub>-axis,<sup>[162,163]</sup> thereby shrinking the size of the outer pockets and expanding the inner pocket, the observed binding of neutral guests remained poorly understood in terms of driving force, structural preferences and scope.

In cooperation with the Clever group a systematic study comprising more than 50 neutral guest molecules has been carried out. It reveals the influence of the guests size, shape and chemical nature on the binding affinity.

### 5.2.1 Summary of Experimental Findings

Binding free enthalpies between  $-6.1 \text{ kJ}\cdot\text{mol}^{-1}$  (furan) and  $-27.1 \text{ kJ}\cdot\text{mol}^{-1}$  (DABCO (1,4-diazabicyclo[2.2.2]octane)) were determined in acetonitrile by NMR

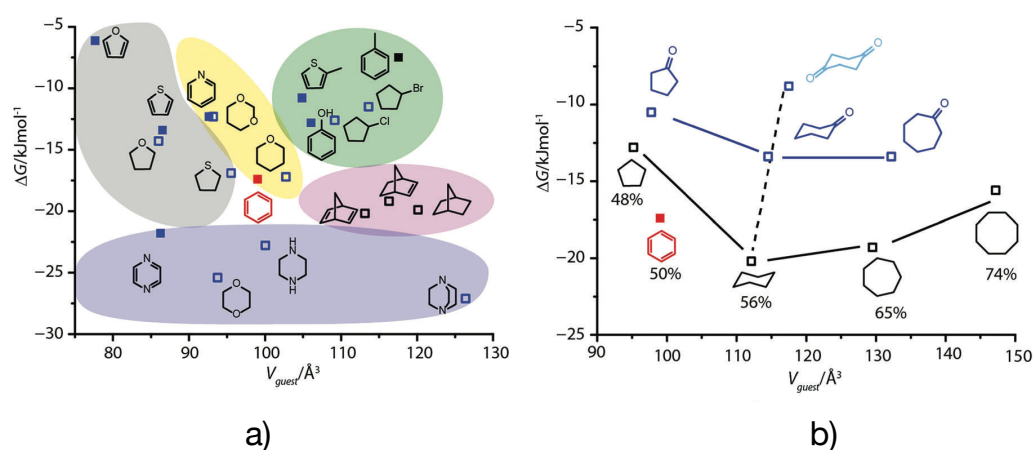


**Figure 5.5:** Ligand structure, cage assembly, chloride activation and uptake of neutral guests. (i) [Pd(CH<sub>3</sub>CN)<sub>4</sub>](BF<sub>4</sub>)<sub>2</sub> in CH<sub>3</sub>CN; (ii) 2 eq. Cl<sup>-</sup>; (iii) neutral guest. The inset shows the guest DABCO binding inside the central cavity (electrostatic potential mapping). Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.



signal integration of the free and filled host species. Figure 5.6 relates the obtained free enthalpies of binding to the guest volume and its packing coefficient (PC) in the cavity, respectively. The PC was computed as the quotient of the CPK volume of EDF2/6-31G\* optimized molecular models using Spartan<sup>[164]</sup> with the volume of the empty pocket calculated with the Voidoo software.<sup>[165]</sup> Thereby, the crystallographically determined structures containing DABCO and benzene as guest served as the reference. Both structures were symmetry expanded and the guest molecules as well as the chloride counter ions were removed. The resulting inner cavities were calculated using a primary grid and plot grid spacing of 0.1 Å and 40 cycles of volume refinement with the default water size probe radius of 1.4 Å. For the benzene containing cage a volume of 201.66 Å<sup>3</sup> was calculated, while for the DABCO structure a volume of 200.84 Å<sup>3</sup> was obtained. Thus, the averaged volume of the central cavity was set to 201.25 Å<sup>3</sup> and used to calculate the packing coefficient.

While oversized and non-cyclic guests were found to refrain from encapsulation, smaller 5-ring guests bind weakly (PC between 39 and 48%) as do more voluminous guests carrying substituents such as a methyl group or oxygen atom. Among the 6-ring guests, a competition experiment revealed that benzene binds faster than cyclohexane while the latter binds stronger. Interestingly, bridged 6-rings such as norbornadiene and its relatives bind rather strongly, showing that a 3-dimensional extension of the guest structure is not detrimental to encapsulation as long as the



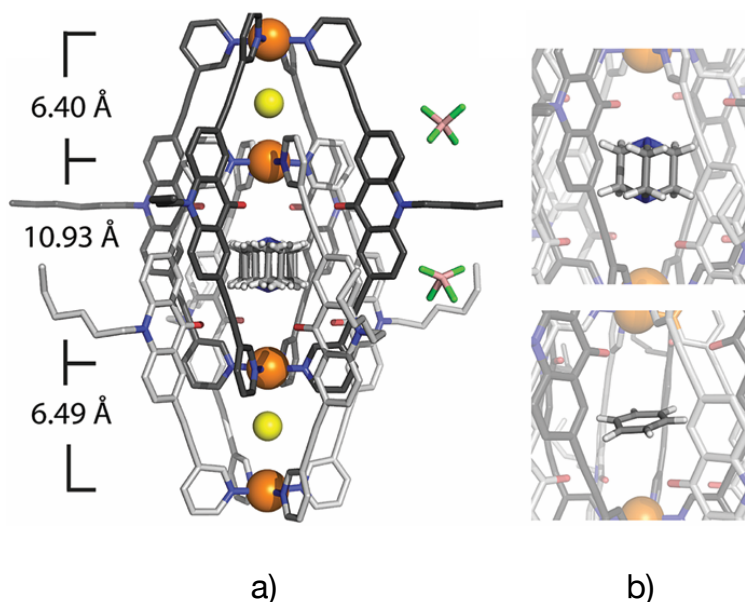
**Figure 5.6:** Binding free enthalpies over guest volume for (a) various guest families with different ring sizes, 3D structure and hetero atom content and (b) selected series of cyclic alkanes and alkanones (packing coefficients are shown for the alkanes; solvent:acetonitrile). Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

## 5 Supramolecular Coordination Cages

overall size does not increase beyond a PC of 74% (adamantane, for example, is not encapsulated; tentative PC = 81%).

A further trend was observed when comparing the affinity/volume relationships for a series of cycloalkanes (C<sub>5</sub>-C<sub>8</sub>), cyclic ketones (C<sub>5</sub>-C<sub>7</sub>) and 1,4-cyclohexanedione (Figure 5.6 b)). Here, the addition of substituents leads to a decrease in binding affinity, demonstrating that lateral expansion of the oblate guest dimension is disfavoured with respect to bridging above the ring plane. Furthermore, the ring size effect is nicely reflected in these series. Small 5-rings bind weaker than their 6-ring siblings but further ring expansion leads to a decrease in affinity. Nevertheless, voluminous cyclooctane (PC = 74%) still binds with considerable affinity, which we denote to its structural flexibility, allowing the guest to adopt a favourable conformation inside the cage.

Among the strongest binders, guests carrying two heteroatoms in opposite positions such as 1,4-pyrazine, 1,4-piperazine and 1,4-dioxane were found. Comparison of the latter guest to its 1,3-isomer is particularly intriguing since it demonstrates that the relative positioning of the heteroatoms is of paramount importance for strong



**Figure 5.7:** Single crystal X-ray structure of  $[2\text{Cl}+\text{DABCO}@\text{Pd}_4\text{L}_8]^{6+}$  (solvents and most counter anions outside the cage structure omitted): a) full side view with disordered DABCO guest; b) central pocket with one selected guest orientation. Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

binding, most likely due to the perfect orientation of the hetero atoms lone pairs towards the two cationic Pd(II) centres lining the cavity. The latter explanation was further supported by a single crystal X-ray structure of the cage carrying its strongest binding guest DABCO in its central pocket (Figure 5.7 a)). While the fourfold crystallographic symmetry of the structure along the Pd<sub>4</sub>-axis complicated the modelling of the axially disordered D<sub>3h</sub>-symmetric guest, the DABCO molecule could be unambiguously refined to stand upright inside the cavity with its N<sub>2</sub>-axis indeed coinciding with the cages Pd<sub>4</sub>-axis (Figure 5.7 b), top). It is interesting to compare this binding mode with the orientation of the previously reported benzene-containing complex<sup>[59]</sup> in which the benzene molecular plane is oriented perpendicular to the Pd<sub>4</sub>-axis, thereby exposing its  $\pi$ -surfaces towards the Pd(II) complexes (Figure 5.7 b), bottom).

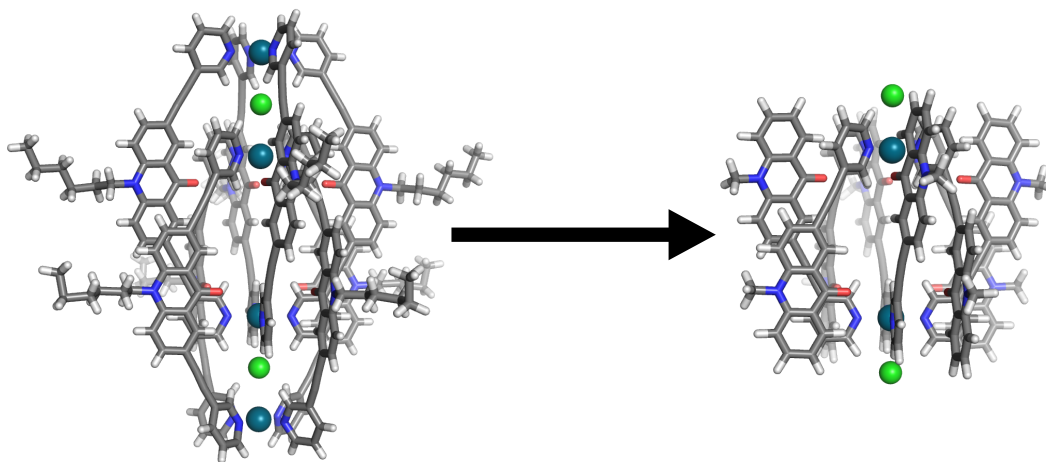
### 5.2.2 Theoretical Investigations

The host system contains over 400 atoms and therefore, accurate calculations at the correlated level using at least a triple zeta basis set have too high demands with regard to the required computation power. To overcome this issue the host structure was taken from X-ray crystallography results at a distance of 10.16 Å between the inner palladium atoms and a model of the double cage was built. For the latter the top and bottom palladium-pyridine planes were cut off and saturated with hydrogens as shown in Figure 5.8. The chlorides in the outer pockets were kept to neutralize part of the charge. Only the hydrogens were optimized with DF-BP86<sup>[23,166]</sup> using the def2-SVP<sup>[156]</sup> orbital basis and its corresponding J-basis.<sup>[167]</sup> For the Pd atoms the ECP-28MWB pseudopotential<sup>[168]</sup> was used. These calculations were carried out with the Turbomole package.<sup>[169]</sup>

To obtain for a selected subset of guests the conformation in the pockets interior they were manually inserted into the model system. Different conformations were used as starting structures for optimizations at the DF-BP86/def2-SVP level of theory. Thereby, the cage geometry was kept frozen. This approximation builds on the assumption that the chloride interaction in the outer pockets is much stronger than the binding in the inner pocket and thus determines the host geometry. The geometry optimizations reflect the experimental observed trends rather well. All guest molecules containing two heteroatoms at opposite positions are aligned along the Pd-Pd axis, while cycloalkanes and aromatic hydrocarbons are aligned perpen-

dicular to this axis. Considering the D3 correction as an estimate for the dispersion interactions in all cases the perpendicular arrangement is favored. However, in the case of doubly substituted heterocycles electrostatics force the Pd-Pd alignment.

Out of this conformational sampling, the most stable structures were selected. For the obtained minimum structures single point DF-SCS-LMP2<sup>[114,115]</sup> calculations (for simplification the prefix ‘DF’ will be dropped in the remaining text) were performed using the Dunning cc-pVTZ<sup>[121]</sup> orbital basis set in combination with the cc-pVTZ-PP<sup>[170]</sup> basis and ECP28MDF<sup>[170]</sup> pseudopotential for Pd (this basis will be referred to as VTZ). The density fitting basis used were the corresponding defaults for cc-pVTZ,<sup>[124,171]</sup> except for calculations including Pd where the def2-TZVPP JK-FIT and MP2FIT<sup>[172,173]</sup> basis sets were used. The orbitals were localized by the Pipek-Mezey scheme.<sup>[83]</sup> The corresponding orbital domains were determined with the Boughton-Pulay criterion<sup>[84]</sup> at a threshold of 0.985. To speed up the calculations use was made of a multipole-approximation<sup>[174]</sup> for very-distant pairs ( $r=10$  Bohr) and local fitting.<sup>[114]</sup> The parameters were chosen by test calculations of the host pocket at increased distance parameters between 8 and 12 Bohr. According to the locality of the orbitals, fragments were built (guest plus host) to determine the dispersion contributions<sup>[47]</sup> at the SCS-LMP2/VTZ level of theory. All calculations were carried out with a development version of Molpro 2015.1.<sup>[110]</sup>

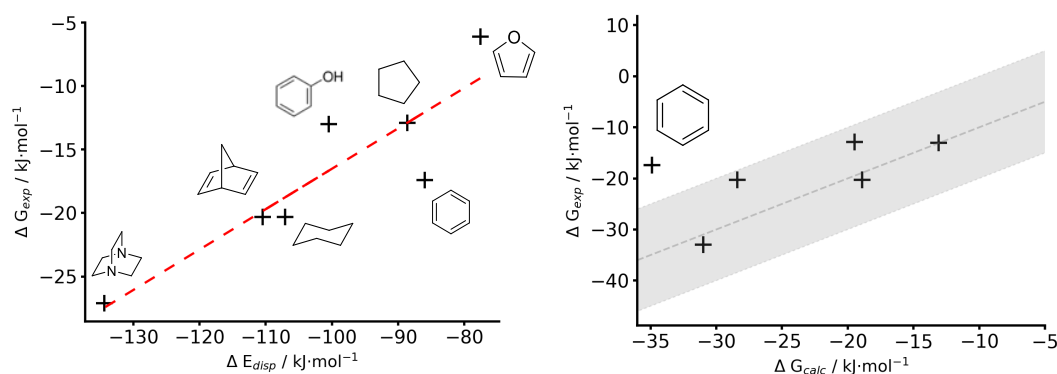


**Figure 5.8:** left: X-ray double cage structure, right: model used for the cage system in the calculations. The ligands were truncated (C–C bonds replaced by C–H) and the bottom and lower planes were removed. The chlorides in the outer pockets were kept in order to neutralize part of the charge. Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

**Table 5.2:** Experimental and computed (SCS-LMP2/VTZ) free binding energies as well as dispersion contributions in  $\text{kJ}\cdot\text{mol}^{-1}$ . Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

Guest	$\Delta G^{296}$	$\Delta G_{\text{theo}}$	$\Delta E_{\text{disp}}$
Furan	-6.1	1.4	-77.6
Cyclopentane	-12.8	-19.5	-88.6
Benzene	-17.4	-34.9	-86.0
Cyclohexane	-20.2	-18.9	-107.1
Norbornadiene	-20.2	-28.4	-110.5
Phenol	-12.8	-13.1	-100.5
DABCO	-27.1	-31.0	-134.4

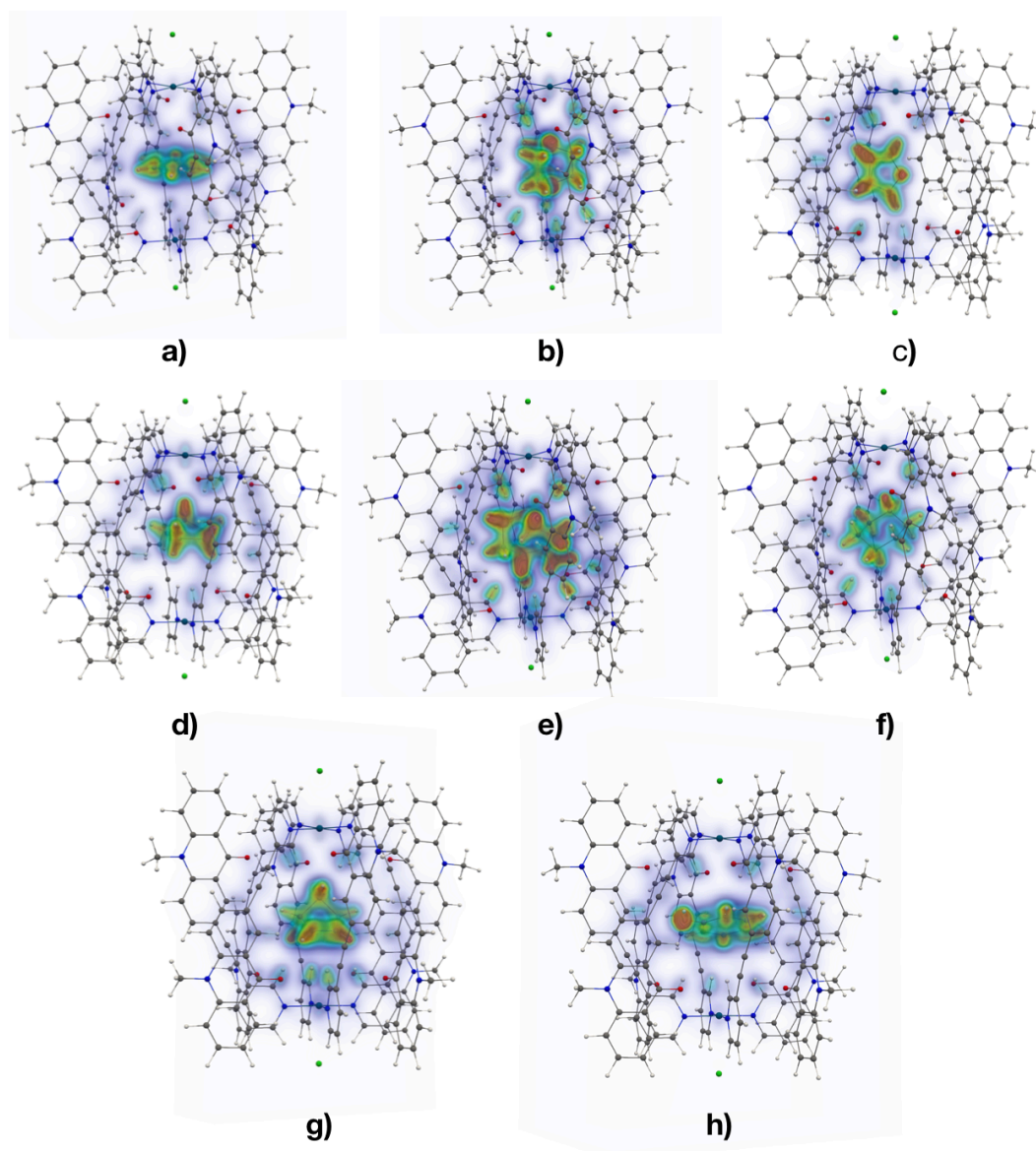
The obtained dispersion energies are provided in Table 5.2 ( $\Delta E_{\text{disp}}$ ). As expected, these contributions increase with system size, with the 5-ring guests in the range of 70 up to 90  $\text{kJ}\cdot\text{mol}^{-1}$ , and the 6-rings going beyond 100  $\text{kJ}\cdot\text{mol}^{-1}$ . The DID profiles depicted in Figure 5.10 show that close C-H contacts to the interior of the pocket dominate the dispersion profile (the plots were prepared with Paraview<sup>[112]</sup>). This is observed also in other systems, somewhat reminiscent of coupled diamondoids.<sup>[55,69]</sup> Individually, these are all small contributions, but they significantly add up. It also explains the difference in  $\Delta E_{\text{disp}}$  between benzene and cyclohexane, since the latter has a larger number of C-H contacts. Interestingly, the dispersion itself correlates with the experimental binding affinity for the subset of guests which were computed (Figure 5.9 left side). The slope of the least-squares fit is 0.41, reflecting the attenuation from solvent effects (i.e. stabilizing dispersion contacts with solvent in the unbound state).



**Figure 5.9:** left side: plot of the experimental free binding energies against computed dispersion energies (SCS-LMP2/VTZ), right side: plot of the experimental free binding energies against computed free binding energies (see approach in the text). Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

## 5 Supramolecular Coordination Cages

In addition, we devised a composite approach for the calculation of the free enthalpy of binding. Solvent effects (acetonitrile) were included by D-COSMO-RS<sup>[175]</sup> implemented in Turbomole. The calculations were done at the DF-BP86/def2-TZVP<sup>[156,167,176]</sup> level. The obtained  $\Delta E_{\text{solv}}$  was computed by a supermolecular approach. For all solvent and thermodynamic corrections the environment was set to ambient pressure (1 bar) and temperature (298.15 K). Thermochemistry corrections were computed with the RRHO-approach.<sup>[67]</sup> The frequencies were computed at the DF-BP86/def2-SVP level for guests@host and a truncated Hessian containing



**Figure 5.10:** Voxel DID plots for different guests@host, a) benzene, b) DABCO, c) furan, d) cyclopentane, e) cyclooctane, f) cyclohexane, g) norbornadiene and h) phenol. Reproduced from Reference [61] with the permission of the Royal Society of Chemistry.

only the guest molecules second derivatives was weighted and diagonalized. Those values are compared to the free guest molecule, building  $\Delta E_{\text{ZPVE}}$ ,  $\Delta H_{(0 \rightarrow T)}$  and  $\Delta S$ . Furthermore, a correction to the solvation energies was added, accounting for the missing translational degrees of freedom in solution, based on the Ardura *et al.* equation for bimolecular reactions<sup>[177]</sup> ( $\Delta G_{\text{trans}}$ ). Finally the electronic energies  $\Delta E_{\text{ele}}$  were calculated at the SCS-LMP2/cc-pVTZ level of theory. The total free energy of binding  $\Delta G$  was then calculated as:

$$\Delta G = \Delta E_{\text{ele}} + \Delta E_{\text{solv}} + \Delta E_{\text{ZPVE}} + \Delta H_{(0 \rightarrow T)} - T\Delta S + \Delta G_{\text{trans}} \quad (5.1)$$

In general, the computed values agree well with the experimental estimates for guest molecules with a small spatial extent (see Figure 5.9 right side), the only outlier being benzene. The mean absolute deviation of the computed binding affinities is  $6.5 \text{ kJ}\cdot\text{mol}^{-1}$  ( $4.6 \text{ kJ}\cdot\text{mol}^{-1}$  excluding benzene), which is expectable given the size and complexity of the system.<sup>[178]</sup> In particular, in the case of large guest molecules like cyclooctane, the calculation suggest that a fixed pocket size for all uptaken molecules is insufficient for the correct description of the host-guest interaction. Looking more closely at the DID profile of cyclooctane (Figure 5.10 e)), the spotlight is found for the interaction between the pyridine planes and the guest molecule. In return, however, this also means that the repulsion becomes significant. Therefore, we expect that in such cases the model systems pocket region is too small and inferentially the dispersion interactions are superimposed by repulsive effects.

### 5.2.3 Outlook

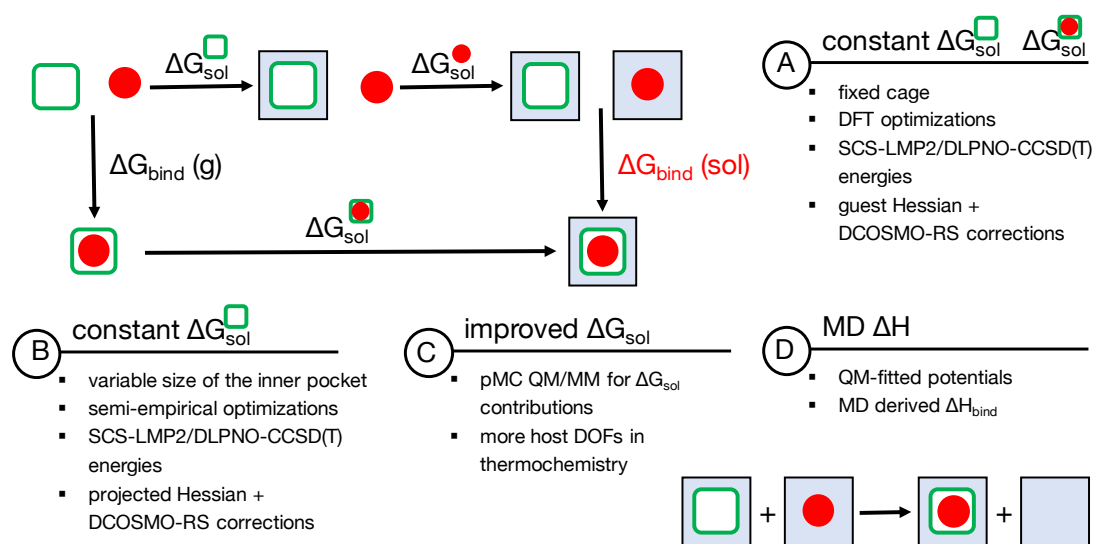
As reported in the previous Subsection, the theoretical description of the system is very difficult. The interpenetrated cages can shift against each other, changing the potential within the middle as well as the potential of the two outer pockets. Furthermore, by these movements the interactions between the banana ligands are amplified or weakened and it has an effect on the solvation. The latter is established by an increased or decreased contact surface. To account for these difficulties three schemes are suggested as shown in Figure 5.11 (scheme A was applied in the previous Subsection).



## 5 Supramolecular Coordination Cages

In scheme B the Pd-Pd distance is shifted subsequently by 0.1 Å with respect to a starting value of 10.16 Å up to 11.56 Å. The host systems are then optimized with HF-3c.<sup>[179]</sup> Afterwards, the guest molecules are projected into the pocket and optimized with fixed pocket geometries. The resulting structures are re-evaluated with PBE/def2-TZVP<sup>[180]</sup> single point calculations to obtain the optimum pocket size for the respective guest molecule. This scheme has already been applied to a test set of guest molecules. The obtained distances are with 10.16 Å and 10.76 Å for benzene and DABCO, respectively, in good agreement to the X-ray measurements. The upcoming local correlation calculations benefit from the developments over the recent years. The initial calculations for these systems used the DF-HF and DF-LMP2 code implemented in Molpro 2015.1.<sup>[110]</sup> Thereby, each single point calculation took at least 10 days on 16 processors. With the latest LDF-HF<sup>[181]</sup> and PNO-LMP2<sup>[42,43]</sup> approaches implemented in Molpro 2018.2<sup>[182]</sup> this is dramatically reduced to 10 hours, illustrating the computational efficiency of these methods. The thermodynamic corrections will be determined in a supermolecular approach using HF-3c. The obtained results will then be fitted empirically to the experimental derived free binding energies obtaining a pocket dependent correction to account for the outer pockets potential and solvation effects.

In Scheme C, the use of perturbative Monte Carlo (pMC)<sup>[183]</sup> is proposed to obtain an



**Figure 5.11:** Schematic representation of possible approaches to computationally determine thermodynamic quantities. Scheme A has been applied in the previous Subsection. The color/shape coding is as follows: red circles denote the guest molecules, green rectangles the host system and blue boxes the solvent.



accurate description of the solvent environment. In combination with an additional potential for the outer pockets (as already computed in Reference [163]), and scheme B one obtains a theoretical description without empirical fittings.

Scheme D suggests a molecular dynamics ansatz as presented by Fenley *et al.*<sup>[184]</sup> In their report an approach to directly determine binding enthalpies for the cucurbit[7]uril (CB7) host system was applied. Thereby,  $\Delta H$  is computed as the difference between the mean energies of four simulation setups, shown in Figure 5.11 scheme D. The existing problem is to obtain an accurate force field for the cage structures. Therefore, we implemented an extension to the quantum mechanically derived force field (QMDF),<sup>[185]</sup> developed by Grimme, into ChemShell.<sup>[186,187]</sup> Unfortunately these calculations could not be finished in time.

## 5.3 Conclusions

In this Chapter two supramolecular host-guest systems were investigated. The first Section dealt with the bis-monodentate pyridyl ligand  $\mathbf{L}^2$  that carries a bulky adamantyl group protruding sideways from its concave face. Through self-assembly a sterically overcrowded cage is formed which is able to encapsulate a range of guest molecules in its interior. Single crystal X-ray structures of the free cage and three host-guest complexes were obtained showing that the larger guests squeeze into the cavity by pushing the adamantyl residues further aside. This in turn leads to a slight compression of the cage along its Pd-Pd-axis. Non-covalent contacts between the guests and the surrounding host were investigated by local orbital analysis using SCS-LMP2. A substantial contribution of attractive dispersion interactions conveyed by the adamantyl groups was identified. The computed DID images show that the  $\pi$ -systems of the guests interact strongly with the surface of the pocket. In particular, some hot spots are identifiable where the adamantyl groups are in close contact with the encapsulated guests. Since both dynamic scaffolds and densely packed molecular surroundings play important roles in the chemical processes taking place inside biological cavities (such as enzyme pockets), the type of artificial self-assemblies described herein will assist in understanding the interplay between molecular crowding and reactivity. In addition, it is anticipated that such sterically overcrowded host structures in which the guests are snapping into a narrow inner cavity contribute new ideas to the field of supramolecular catalysis.<sup>[188]</sup>

In the second Section, an interpenetrated double cage was discussed. It is formed from the bis-monodentate pyridyl ligand **L** in self-assembly and features three consecutive binding pockets, filled with tetrafluoroborate anions. These host systems can be activated by the addition of chloride anions to bind neutral guests. The here presented study curtails the scope of neutral guest molecules that can be encapsulated by the interpenetrated host system. In terms of size and shape, 6-rings and norbornene-type compounds are favoured. One substituent with the size of a methyl group is tolerated (leading to a drop in affinity) but two substituents are not. The opposite arrangement of hetero atoms in bridged or unbridged 6-ring guests is favourable, indicating that orientation of lone pairs towards both cationic Pd(II) centres leads to a stabilizing two-point attachment. This assumption was supported by the X-ray structure of the host-guest complex containing DABCO. Electronic structure calculations reveal that binding is strongly favoured by dispersion interactions. Moreover, a correlation between the experimental free binding and computed dispersion energies was obtained. The DID profiles showed that close C-H contacts to the interior of the pocket dominate the dispersion profile. By a rather simple approach, with regard to the complexity of the system, free binding energies could be determined. In general, the computed values agree well with the experimental estimates, the only outlier being benzene. In the case of large guest molecules like cyclooctane, it became obvious that a more sophisticated approach is required to account for the cage relaxation by incorporating such molecules.

## 6 Metallophilic interactions

This Chapter is reproduced from Reference:

[54] A. Wuttke, M. Feldt, R. A. Mata, All That Binds Is Not Gold—The Relative Weight of Auophilic Interactions in Complex Formation, *The Journal of Physical Chemistry A*, **2018**, *122*, 6918-6925 (DOI: 10.1021/acs.jpca.8b06546),

with the permission of the American Chemical Society. In this Reference parts of my Masters Thesis on "Theoretical Investigations of Auophilic Interactions in Molecular Dimers and Trimers" are included. All calculations except the basis set extrapolation for the Cl-Au-Me<sub>2</sub>bimy system were carried out by myself.

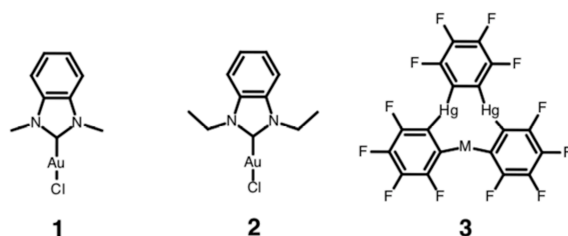
Albeit the many advances observed in quantum chemical methods over the recent years, the ab initio prediction of molecular crystal structures still remains a grand challenge. Part of the problem is inherent to the systems under study. The number of energetically low-lying polymorphs increases rapidly with the size and chemical complexity of the molecular monomers. This leads to a vast number of structures which need to be identified and compared for their thermodynamic and/or possibly kinetic stability. The other issue is, of course, the multitude of interactions which have to be accurately described in order to carry out such comparisons. The competition between the latter will in the end define the form of the crystal, in particular how the monomers are oriented relative to their neighbors.

One approach to the problem is the brute-force application of high-level electronic structure methods to a selected number of candidates. This is certainly possible today for crystals composed of small sized organic molecules, with one notable example being that of benzene.<sup>[189]</sup> With the continuing development of density functional approaches and low-order scaling correlated methods, more complex systems can be targeted today,<sup>[190]</sup> as long as the number of polymorphs is small and the relevant minima are already characterised.

However, in many cases one is not only interested in knowing what the global minimum structure is, but also in the reason behind it. One would like to understand the interplay between the different intermolecular forces. This information, for example, can be used to design compounds where certain polymorphs are avoided.

If one is interested in systems with more than one interaction site, the relative weight of the different contributions can be much more difficult to obtain. Upon localizing the occupied space orbitals (and possibly the virtual space as well), the different interactions in the electron correlation energy can be attributed to different interacting regions, even if the fragments are covalently bound.<sup>[47–50,52,53]</sup> Such analysis depends either on a suitable separation of the orbital space,<sup>[191]</sup> or in identifying the region where the orbitals localize. In order to identify the direct Au-Au interaction energy in a series of model Au(I) complexes, Andreijic and Mata made use of local orbital analysis and the fragmentation scheme discussed in Section 3.1.<sup>[51]</sup> The results of this work revealed two main trends. The strength of the latter interaction is independent of the nature of the ligands, and can be safely derived from a simple distance criterion. Secondly, larger complexes reveal strong ligand-ligand interactions, which can easily outweigh the aurophilic interaction. This goes against the prevalent logic that the structure of complexes with short Au-Au contacts are primarily influenced by the so called aurophilic interaction.<sup>[192,193]</sup>

The investigation of Au(I) complexes was continued, discerning the relative weight of such interactions in building molecular complexes. With this purpose in mind, the crystal structure of two closely related complexes are studied. These are the [Au(Me<sub>2</sub>bimy)Cl] and [Au(Et<sub>2</sub>bimy)Cl] complexes (**1** and **2**, respectively, see Figure 6.1)(bimy = benzimidazol-2-ylidene). The latter were characterized by Wang *et al.*<sup>[62]</sup> Furthermore, Au(I)-Hg(II) complexes aroused interest as examples of mixed



**Figure 6.1:** Lewis structures of **1** [Au(Me<sub>2</sub>bimy)Cl] **2** [Au(Et<sub>2</sub>bimy)Cl] and **3** MHg<sub>2</sub>(o-C<sub>6</sub>F<sub>4</sub>)<sub>3</sub> where M can be either Au or Hg. Reproduced from Reference [54] with the permission of the American Chemical Society.

metallophilic interactions. It has been stated that the forces between those metals are much stronger than the aurophilic ones.<sup>[63,64]</sup> To investigate the effect a trinuclear complex  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  containing one short  $\text{Au(I)} \cdots \text{Hg(II)}$  contact (**3** see Figure 6.1,  $M = \text{Au, Hg}$ ) is considered.<sup>[64]</sup>

The central question is how the competing forces contribute to the final crystal structures. Use is made of a fragment orbital analysis of the electron correlation energy, as outlined in Section 3.1. This is particularly well suited for such systems, since a significant percentage of the interactions is due to the latter. One should take into account that in all systems observed in this Chapter  $\pi$ - $\pi$  interactions are present. These are not purely electron correlation effects, since they include sizable quadrupole (and higher multipole) electrostatic contributions. However, as later demonstrated, an analysis of the electron correlation energy suffices to understand the relative impact of metallophilic interactions. Furthermore, a closer look is taken into dispersion interactions in the aforementioned systems. Plotting Dispersion Interaction Densities (DID) allows one to gain a greater insight into which fragments of the molecules contribute to the complex stabilization by London forces.

## 6.1 Computational Details

Calculations carried out in this Chapter are based on local correlation methods, as first suggested by Pulay and later developed by Werner and coworkers.<sup>[79,194,195]</sup> As verified in a previous study,<sup>[51]</sup> SCS-LMP2<sup>[115]</sup> results are in very close agreement with Coupled-Cluster calculations (LCCSD(T0)) for the study of aurophilicity. Furthermore, it has already been shown in Chapter 4 and other studies,<sup>[117,196]</sup> that the spin-component scaled method is also adequate for the description of  $\pi$ - $\pi$  interactions. The fragmentation of the intermolecular electron correlation energies is performed as in an earlier report on aurophilic interactions<sup>[51]</sup> and outlined in Section 3.1.

Density fitting approximations were used throughout. The prefix ‘DF-’ will be dropped. Through the use of SCS-LMP2, one also obtains a straightforward correction of MP2 energies at the pair energy terms which are used in the decomposition. In all local correlation calculations the occupied orbitals were Pipek-Mezey

localized.<sup>[83]</sup> The orbital domains were determined according to the NPA criteria<sup>[94]</sup> with  $T_{NPA} = 0.03e$ .

The basis sets used were the Dunning aug-cc-pVTZ orbital basis set<sup>[121,122]</sup> in combination with the aug-cc-pVTZ-PP basis and the ECP60MDF pseudopotentials<sup>[197,198]</sup> for Au and Hg (this basis will be referred to as AVTZ). For larger complexes the hydrogens diffuse basis functions were removed (this basis will be referred to as A'VTZ). The density fitting basis used were the corresponding defaults (for the aug-cc-pVTZ and aug-cc-pVTZ-PP basis)<sup>[123,124]</sup> except for SCF calculations where the def2-QZVPP/JKFIT<sup>[168,199]</sup> basis set for Au and Hg was used. All calculations were carried out with a development version of Molpro 2015.2.<sup>[110]</sup>

In all calculations the monomers were optimized with the BP86<sup>[23,166]</sup> functional using the def2-TZVP basis set<sup>[156]</sup> for all atoms except Au whereby the ECP60MDF core potential was used in combination with the dhf-TZVP basis functions<sup>[200]</sup> and Hg with a ECP60MWB<sup>[168]</sup> pseudopotential. The different arrangements included in this study (dimers and trimers) were built from frozen monomers. To ensure a smooth potential profile of the interaction curves, the orbital domains were computed at the r(Au-Au) distance of 100 Å and kept fixed in all calculation steps.

The starting structures for **1** and **2** were taken from the published X-Ray structures by Wang *et al.*<sup>[62]</sup> For complex **3**, X-ray structures provided in Reference<sup>[64]</sup> served as starting point. Dispersion Interaction Densities (DID) are computed as proposed in Chapter 3. The Voxel DIDs are plotted using the ParaView Software.<sup>[112]</sup> Energy decomposition analysis (EDA) of the reference wave function is performed with a modified Kitaura-Morokuma scheme as proposed in Reference <sup>[100]</sup>.

## 6.2 Results and Discussion

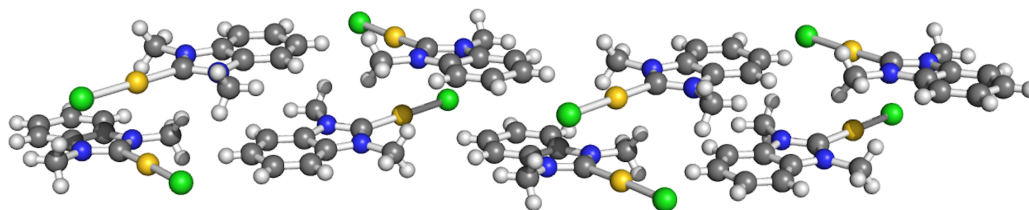
### 6.2.1 Cl-Au-Me<sub>2</sub>bimy

A cut of the crystal structure of **1** is shown in Figure 6.2. The chain can be understood as a collection of dimer units translated in space, with the Au centers closely stacked. The distance between them is measured at about 3.16 Å. The aromatic rings are also favorably stacked, building contacts to the neighboring dimer. The

distance between the center of the rings is 3.45 Å. Therefore, the minimal system to contain all of the relevant interactions is a trimer, where one of the complexes interacts aurophilically with another monomer, and through  $\pi$ - $\pi$  interactions with the other neighbor.

The first interesting observation is the monomers orientation forming the short Au...Au contact. In the theoretical investigations of the aurophilic effect, it is common that monomers are oriented with a 90° dihedral angle (C1-Au1-Au2-C2) along the normal axis<sup>[48,50,192,201,202]</sup> (Figure 6.3 a)). This arrangement leads to a null dipole-dipole interaction term, facilitating the interpretation of the results. Electrostatics are removed, and the stabilization is exclusively due to electron correlation effects. In the crystal structure the dihedral angle differs slightly from 90°. In one set of calculations, a dimer with a dihedral angle of exactly 90° is considered, noted as model dimer (**1m**). As a comparison, use is made of a dimer with a dihedral angle C1-Au1-Au2-C2 of 94.7° with Au1-Au2-C2 forming an angle of 84.4°, in agreement with the crystal structure. This will be referred to as the crystal dimer (**1a**) structure. This set up will reveal how much the results are actually influenced by such assumptions.

The energy profiles were calculated with fixed monomer geometries at different  $r(\text{Au-Au})$  distances at the HF and SCS-LMP2 level using the A'VTZ basis set and they can be seen on Figure 6.3 b). The SCS-LMP2/A'VTZ level of theory shows for the crystal dimer an increased well depth of  $-17 \text{ kJ}\cdot\text{mol}^{-1}$  compared to the model system. This effect cannot be explained on the basis of electron correlation effects, as the latter are of similar magnitude in both cases. The change in the angles has little effect on such contributions. This is within expected, since these should be mostly anisotropic. However, the crystal dimer shows an attraction of about  $-10 \text{ kJ}\cdot\text{mol}^{-1}$  at the  $r(\text{Au-Au})$  distance of 4.2 Å in the Hartree-Fock energy profiles. An EDA was performed to verify which kind of interactions stabilize the crystal complex.



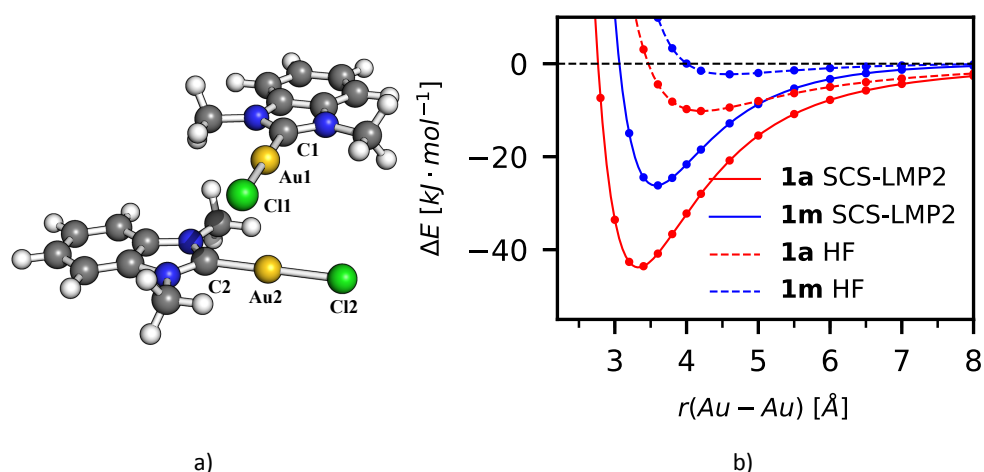
**Figure 6.2:** Chain structure of  $[\text{Au}(\text{Me}_2\text{bimy})\text{Cl}]_n$ . Reproduced from Reference [54] with the permission of the American Chemical Society.

**Table 6.1:** EDA for **1a** and **1m** at  $r(\text{Au-Au})$  3.4 Å (in  $\text{kJ}\cdot\text{mol}^{-1}$ ). Reproduced from Reference [54] with the permission of the American Chemical Society.

	$\Delta E_{\text{elst}}$	$\Delta E_{\text{exch}}$	$\Delta E_{\text{rep}}$	$\Delta E_{\text{pol}}$	$\Delta E_{\text{HF}}$
<b>1a</b>	-39.0	-103.6	166.6	-21.1	2.9
<b>1m</b>	-28.6	-115.0	189.1	-23.0	22.5
<b>1a-1m</b>	-10.4	11.4	-22.5	2.1	-19.6

Thereby, **1a** and **1m** were computed at the  $r(\text{Au-Au})$  distance of 3.4 Å. The results are shown in Table 6.1. The analysis shows that the crystal arrangement is preferred due not only to improved electrostatics (dipole-dipole interactions and higher order terms), but also by removing some of the repulsion between the monomers wavefunctions. These are most likely linked to the contact between the methyl and the ring, which is avoided by tilting the monomer.

To obtain a deeper insight into the intermolecular electron correlation energies the monomers were split in three fragments, namely Au, Cl and L(the carbene ligand). Comparing the results for the crystal (**1a**) and the model dimer (**1m**), some slight deviations are observed (Figure 6.4 a)-b). The electron correlation interaction energy  $\Delta E(\text{L-L})$  is lowered while  $\Delta E(\text{L-Cl})$  is increased in the crystal conformation. However, the total electron correlation interaction energy has still a similar value. In the obtained equilibrium distance the Au-Au electron correlation interaction does not change from the model system to the crystal structure and has a similar con-



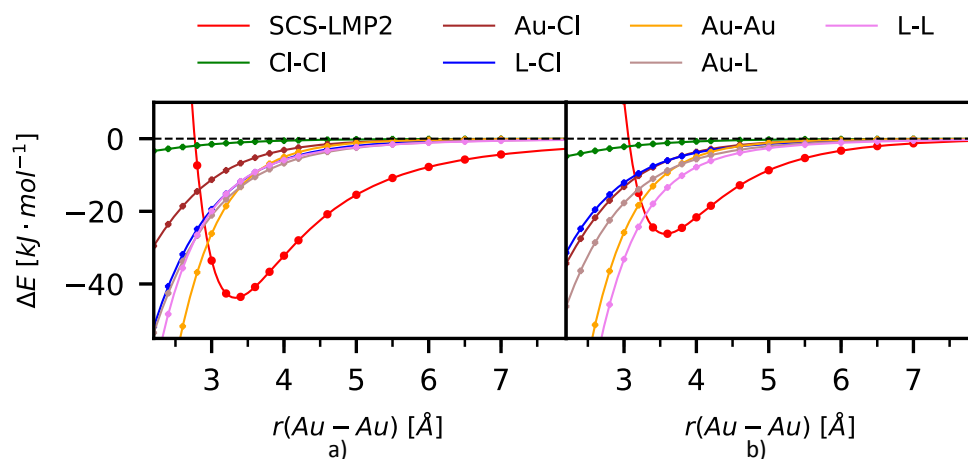
**Figure 6.3:** a) Orientation of Au(Me<sub>2</sub>bimy)Cl in dimers, b) Potential energy curve of the Au(Me<sub>2</sub>bimy)Cl model (**1m**) and crystal (**1a**) dimers at HF/A'VTZ and SCS-LMP2/A'VTZ theory levels. Reproduced from Reference [54] with the permission of the American Chemical Society.



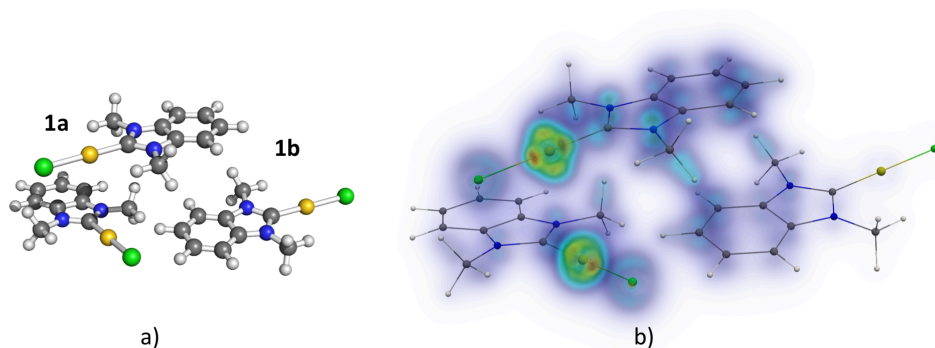
tribution like L-L, L-Cl and Au-L. Taking this into account and the stabilization at the HF level, one is lead to the conclusion that the Au-Au interaction is not the dominant stabilizing force in the crystal dimer.

To confirm the role of aurophilicity in the complex formation the trimer shown in Figure 6.5 a) has also been considered. As previously mentioned, the trimer can be rationalized as the combination of two dimers, one with a conformation showing Au-Au interaction, same as before (**1a**), and the other one showing  $\pi$ - $\pi$  stacking, in the following referred to as **1b** (see Figure 6.5 a)).

By observing the SCS-LMP2/A'VTZ results for the trimer, an equilibrium distance of 3.4 Å is obtained. This value differs from the experimental result by 0.2 Å. In



**Figure 6.4:** Decomposition of the intermolecular SCS-LMP2/A'VTZ electron correlation energies together with the full result for a) **1a** dimer and b) **1m** dimer. Reproduced from Reference [54] with the permission of the American Chemical Society.



**Figure 6.5:** a) representation of the initial Au(Me<sub>2</sub>bimy)Cl trimer, b) Voxel DID of the Au(Me<sub>2</sub>bimy)Cl trimer at SCS-LMP2/A'VTZ. Reproduced from Reference [54] with the permission of the American Chemical Society.

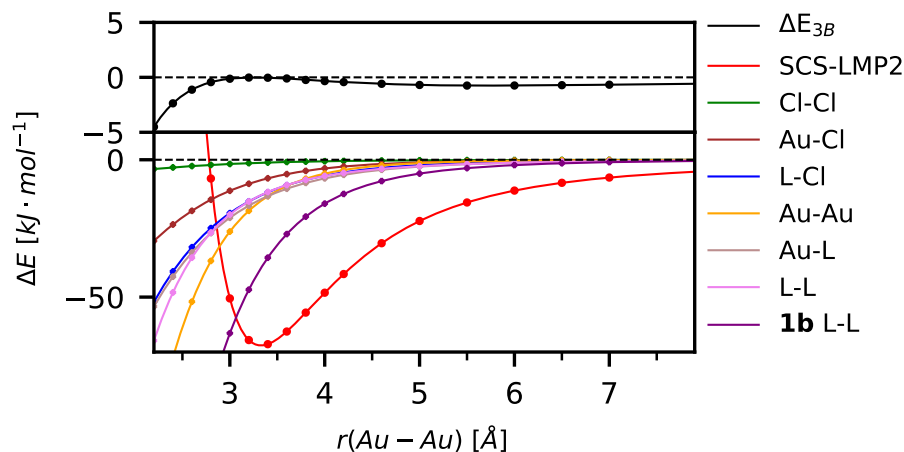
former studies it had already been shown a strong basis set dependence of the equilibrium distance for unsupported aurophilic dimer systems.<sup>[203]</sup> In order to converge the results of the interaction curves, a basis set extrapolation was performed. This has been carried out with the A'VTZ and A'VDZ basis sets. The resulting SCS-LMP2/CBS[2:3] curves give a minimum at 3.20 Å, in excellent agreement with the X-ray structure. Furthermore, it should be considered that the energy difference between the A'VTZ minimum at 3.4 Å and the CBS[2:3] minimum at 3.2 Å is only ~1 kJ·mol<sup>-1</sup>.

In Figure 6.5 b) a DID plot of a trimer is depicted. Two main regions of interaction can be identified. One corresponds to the direct Au-Au interaction, the other to the interaction between the two ring systems. The alternating disposition of the monomers allows contacts with neighboring molecules, both through an aurophilic contact and  $\pi$ -stacking. Even the methyl groups contribute to the total stabilization. These results once again strengthen the idea that the aurophilic interaction is not the leading force in this crystal. It can be also seen that such plots offer the possibility of separating individual contributions, helping to identify the specific role of different moieties.

In order to obtain a clearer picture of the interaction energy for the trimer the systems additivity was evaluated. If the interactions involved are additive, the total trimer interaction energy should be coincident with the interaction energy of both dimers. In such a case, the three body term

$$\Delta E_{3B} = E_{0,\text{trimer}} - \sum_x^{\text{dimers}} \Delta E_x - \sum_y^{\text{monomers}} E_{0,y}, \quad (6.1)$$

should be close to zero ( $E_0$  are single point energies and  $\Delta E_x$  denote dimer interaction energies). SCS-LMP2/A'VTZ energies were used for computing said 3-body energies. The results are given in a onset of Figure 6.6. As it can be observed, no significant three body interactions are obtained in the region of the equilibrium distance. Therefore, one should be able to safely describe the aggregation forces in the crystal by pairwise dimer interaction energies. This is done in Figure 6.6, where the main electron correlation contributions, which have been extracted from dimer **1a** and **1b** calculations, are plotted. Since in the dimer **1b** only the L-L interaction is significant, this is the only term plotted for ease of visualization.



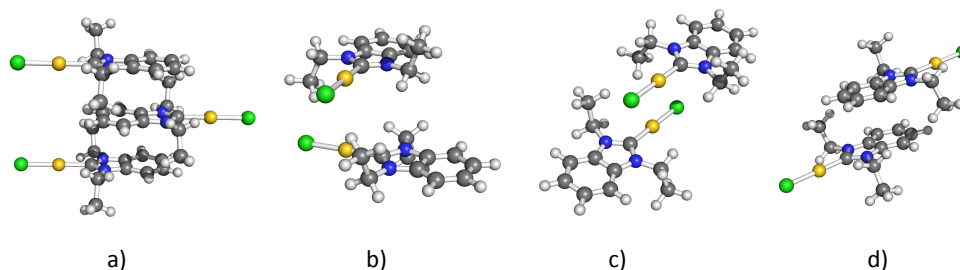
**Figure 6.6:** SCS-LMP2/A'VTZ electron correlation energy decomposition of the Au(Me<sub>2</sub>bimy)Cl **1a** dimer and **1b** dimer (only L-L contribution) potential curves and three-body energies as computed from Eq (1). Reproduced from Reference [54] with the permission of the American Chemical Society.

Considering these results for the whole trimer, in particular the large L-L contribution from dimer **1b**, the Au-Au interaction is dwarfed even further. Given the stoichiometry of the crystal and its arrangement, the number of **1a** and **1b** dimers is the same, so the relative weights can be directly deduced from Figure 6.6. The electron correlation interaction energy for Au-Au only contributes ~14% to the total correlation energy. Thus, even though the molecular crystal might hint at a strong role of aurophilicity, it has in fact a minor impact in its stability. The main players are the London type forces between the aromatic rings, electrostatic interactions and weak electron correlation components as detailed in dimer **1a**.

### 6.2.2 Cl-Au-Et<sub>2</sub>bimy

Closely related to **1** is the complex [Cl-Au-Et<sub>2</sub>bimy] (**2**), which only differs by the substitution of two methyl groups by two ethyl moieties. However, the crystal structure of the two complexes differ strongly, with the latter showing no favored aurophilic contacts.

In Figure 6.7 a), part of the crystal structure for complex **2** is shown. It can be seen that two ethyl groups on the bimy moiety point to opposite sides of the bimy plane. Such orientation of the ethyl groups allow a columnar stacking of the bimy planes. Also, it can be noticed that only ring  $\pi$ - $\pi$  interactions are present, but no



**Figure 6.7:** a) Crystal structure of the  $[\text{Au}(\text{Et}_2\text{bimy})\text{Cl}]_n$  b) complex  $90^\circ$  perpendicular arrangement (**2o**), c)  $180^\circ$  perpendicular arrangement (**2p**) and d) arrangement as in the crystal (**2c**). Reproduced from Reference [54] with the permission of the American Chemical Society.

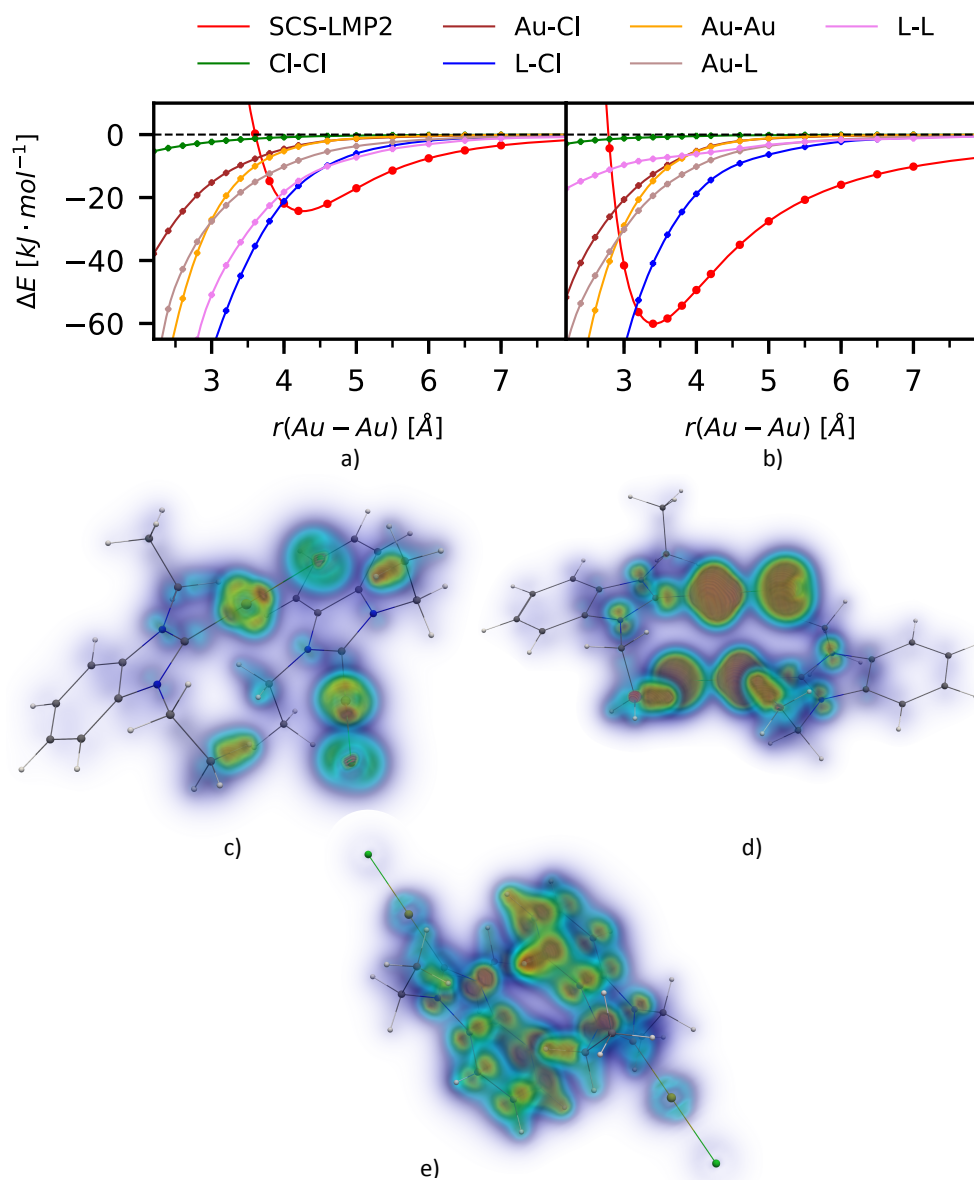
Au-Au interactions. The distance between two ring planes is  $3.53 \text{ \AA}$ . In this case each monomer interacts with upper and lower bimy rings only through the bimy ring in a head-to-tail fashion. Since only the  $\pi$ - $\pi$  interactions are present, and this is pairwise repeated, a model dimer will suffice for the theoretical analysis.

Given the absence of close  $\text{Au} \cdots \text{Au}$  contacts, the aim was to investigate two different (hypothetical) orientations where aurophilic interactions could exist, in an attempt to understand this occurrence. The use of similar structures as those used in the previous Sections also allows for a more direct comparison of the complexes, and the influence of the ethyl groups. At first, there are two systems with a face-to-face orientation and dihedral angles of  $90^\circ$  (ortho dimer, **2o**) and  $180^\circ$  (para dimer, **2p**). Lastly, the dimer present in the crystal is used (this will be denoted as **2c**) (Figure 6.7 b-d)). Minimal distances between monomer planes and their energies at the minimum are compared.

In the ortho dimer **2o** the Au-Au distance measures  $4.2 \text{ \AA}$  and the well depth amounts to about  $-24 \text{ kJ}\cdot\text{mol}^{-1}$  (Figure 6.8 a)). Comparing these results with those obtained for the model dimer with methyl groups (**1m**), one can see that the equilibrium distance is increased by  $0.6 \text{ \AA}$  and the well depth differs by  $\sim 2 \text{ kJ}\cdot\text{mol}^{-1}$ . By studying the molecular structure it can be seen that the ethyl groups point towards the chlorides of the neighboring monomer. In Figure 6.8 b) one can observe that this results in large dispersion contributions even at apparently large intermonomer distances. However, this also leads to stronger Pauli repulsion due to orbital overlapping. By taking a closer look into the electron correlation contributions of different interactions, it is observed that the Cl-L interaction is the strongest one. The Au-L

interaction seems to be strengthened by the larger ethyl groups, comparing to the **1m** dimer studied in the previous Section. However, there is a high steric repulsion by the Et<sub>2</sub>bimy ligands and chloride, which leads to a large Au-Au distance. As a result, the Au-Au term now contributes less than 10% to the total electron correlation interaction energy at the equilibrium distance.

The para dimer **2p** is already at the HF level largely stabilized due to electrostatic interactions. Taking the potential curve at the SCS-LMP2/AVTZ level, the well



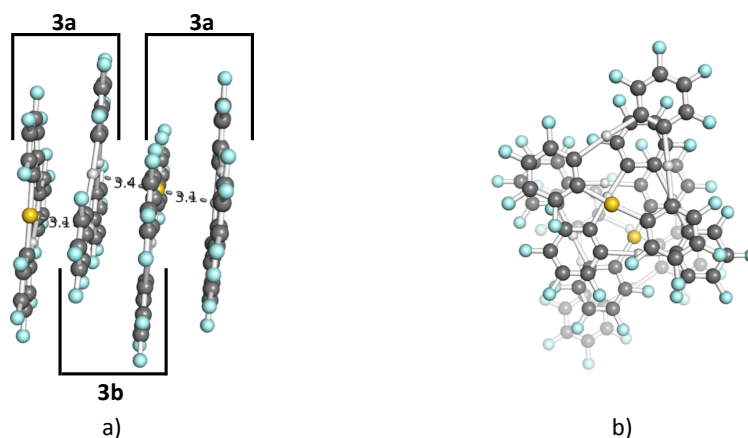
**Figure 6.8:** Decomposition of the intermolecular SCS-LMP2/AVTZ electron correlation energies together with the full result for a) **2o** and b) **2p** dimers. Voxel DIDs at SCS-LMP2/AVTZ for c) **2o** d) **2p** and e) **2c** dimers. Reproduced from Reference [54] with the permission of the American Chemical Society.

depth stands at about  $-60 \text{ kJ}\cdot\text{mol}^{-1}$  with a  $r(\text{Au-Au})$  of  $3.4 \text{ \AA}$ . A strong interaction between the  $\text{Et}_2\text{bimy}$  ligand and chloride is observed. The Pauli repulsion between them is lowered, since the ethyl groups are laterally offset. Moreover, the para conformation results in a low L-L interaction, but the Au-Cl term is maximized by this conformation. The electron correlation energies curves for  $\Delta E(\text{Au-L})$  and  $\Delta E(\text{Au-Au})$  are not significantly affected by changing the conformation from ortho to para. However, due to the shorter Au-Au distance,  $\Delta E(\text{Au-Au})$  contributes more to the stabilization in **2p**. Moreover, the L-L contribution is much lower compared to **2o**. Summed up the Au-Au relative contribution increases from the ortho to the para conformation about  $\sim 15\%$ . In Figures 6.8 c)-d) DID plots are shown.

In one last set of calculations, a closer look is taken at dimer **2c**. Compared to dimer **1b** the bimyl ligands in this conformation have a larger overlap (Figure 6.7 d)). Due to the dispersion character of the  $\pi$ - $\pi$  sandwich stacking, the resulting electron correlation energy should be increased. The DID Plot in Figure 6.8 e) shows that the ethyl groups only have a small influence on the interaction energy, due to their outwardly orientation. At the SCS-LMP2/AVTZ level a well depth of about  $-40 \text{ kJ}\cdot\text{mol}^{-1}$  at the ring distance of  $3.6 \text{ \AA}$  is observed. The calculated equilibrium distance is very close to the experimental result ( $0.07 \text{ \AA}$  deviation). As expected the well depth from the former investigated dimer **1b** (Section 6.2.1) is outbid by  $\sim 17 \text{ kJ}\cdot\text{mol}^{-1}$ .

### 6.2.3 $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$

In this Section the trinuclear  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  complex is studied. A part of the crystal structure is shown in Figure 6.9 a)-b). The compounds are sandwich-staked and show two slightly different conformations. The complexation of the crystal can be described by a trimer including both arrangements. There is one added complication when discussing this system. There is no clear distinction between gold and mercury in X-ray diffraction methods, due to the fact that they are isoelectronic.<sup>[64]</sup> However, based on an earlier publication,<sup>[63]</sup> which showed strong Au(I)-Hg(II) interactions in mononuclear Au(I)-Hg(II) complexes, it has been argued that the close contacts are caused by strong Au(I) $\cdots$ Hg(II) contacts. This hypothesis will be verified by interchanging those atoms. Further, this allows to effectively compare mercuriphilic and mixed Au(I)-Hg(II) interactions.



**Figure 6.9:** a) side view of the trinuclear  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  complex and b) front view. Reproduced from Reference [54] with the permission of the American Chemical Society.

First of all, it was tested whether there are any significant three body interactions, just as in the previous systems. For that, Equation 6.1 has been applied to the crystal arrangement. To save computational time the basis set was lowered for this procedure, so that the carbons and fluorines were treated with a cc-pVTZ basis, the basis sets and pseudopotentials for the heavy metals were kept fixed.  $\Delta E_{3B}$  has a value of  $4.4 \text{ kJ}\cdot\text{mol}^{-1}$  compared to an interaction energy of  $-299.3 \text{ kJ}\cdot\text{mol}^{-1}$ . Therefore, three body interactions are deemed to be only a minor effect and the discussion is based only on the dimers.

The dimer with the shortest  $\text{Au}\cdots\text{Hg}$  contact was already theoretically studied.<sup>[64]</sup> Their results have shown a purely repulsive HF interaction energy. However, at the MP2 level of theory a large stabilization energy of  $-199.2 \text{ kJ}\cdot\text{mol}^{-1}$  at a distance  $r(\text{Au-Hg})$  of  $3.19 \text{ \AA}$  was obtained. The authors came to the conclusion that the main attraction is provided by the close  $\text{Hg(II)}\cdots\text{Au(I)}$  contact.<sup>[64]</sup>

Our interest lies not in reassessing the potential energy surface. Therefore, the  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  dimer was only computed at the experimental Au-Hg distance of  $3.1 \text{ \AA}$ . This dimer will be referred to as **3a**. **3b** was computed in the second crystal arrangement at the Au-Hg distance of  $3.4 \text{ \AA}$  (see Figure 6.9 a)). SCS-LMP2/AVTZ gives an interaction energy of  $-162.1 \text{ kJ}\cdot\text{mol}^{-1}$  for **3a** and  $-158.1 \text{ kJ}\cdot\text{mol}^{-1}$  for dimer **3b**. Without spin-component scaling, the interaction energies are with  $-220.8 \text{ kJ}\cdot\text{mol}^{-1}$  and  $-216.3 \text{ kJ}\cdot\text{mol}^{-1}$  largely overestimated by

**Table 6.2:** Decomposition of intermolecular SCS-LMP2/AVTZ electron correlation energies (in  $\text{kJ}\cdot\text{mol}^{-1}$ ) and their relative weights for  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3 - \text{Hg}_3(\text{o-C}_6\text{F}_4)_3$ . Reproduced from Reference [54] with the permission of the American Chemical Society.

		<b>3a</b>		<b>3b</b>	
		$\Delta E_{\text{corr}}$	%	$\Delta E_{\text{corr}}$	%
Au	$\text{Hg}_C$	-16.4	7.0	-12.0	5.3
Au	$\text{Hg}_{D+E}$	-3.3	1.4	-5.9	2.6
Au	$\text{R}_B$	-19.4	8.3	-16.9	7.4
$\text{Hg}_{A+B}$	$\text{Hg}_C$	-2.3	1.0	-0.9	0.4
$\text{Hg}_{A+B}$	$\text{Hg}_{D+E}$	-5.9	2.5	-10.6	4.7
$\text{Hg}_{A+B}$	$\text{R}_B$	-28.9	12.4	-31.3	13.8
$\text{R}_A$	$\text{Hg}_C$	-12.4	5.3	-11.9	5.2
$\text{R}_A$	$\text{Hg}_C$	-29.5	12.6	-29.9	13.2
$\text{R}_A$	$\text{R}_B$	-115.4	49.5	-107.5	47.4

MP2 as expected from similar studies.<sup>[115,202,204]</sup> Dimer **3a** is favored by  $4 \text{ kJ}\cdot\text{mol}^{-1}$  which is well within the accuracy limit of the method used.

To obtain the relative weight of the Au(I)-Hg(II) interactions the monomers were fragmented. For  $\text{AuHg}_2(\text{o-C}_6\text{F}_4)_3$  three fragments were built: Au,  $\text{Hg}_{A+B}$  containing both mercury atoms and  $\text{R}_A$  containing all other atoms. The  $\text{Hg}_3(\text{o-C}_6\text{F}_4)_3$  fragments were built in a similar fashion.  $\text{Hg}_C$  denotes the mercury with the close contact to the gold atom,  $\text{Hg}_{D+E}$  stands for the other two mercury atoms and in  $\text{R}_B$  all other atoms are collected. The results for **3a** and **3b** are shown in Table 6.2.

First dimer **3a** is considered. The Au- $\text{Hg}_C$  interaction is about 7% of the intermolecular electron correlation, and when counting all pure metal-metal interactions (Au- $\text{Hg}_C$ , Au- $\text{Hg}_{D+E}$ ,  $\text{Hg}_{A+B}$ - $\text{Hg}_C$ ,  $\text{Hg}_{A+B}$ - $\text{Hg}_{D+E}$ ) together the overall weight is only 12%. Focusing on the dispersion contribution, the DID representation in Figure 6.10 clearly illustrates what is happening. The fluorinated benzene ligands show strong interactions involving the benzenes  $\pi$ -systems and some of the fluorines. 39% of the intermolecular interactions at the correlated level arise from metal-ligand contacts. However, the pure ligand-ligand interaction adds up to account for nearly half the intermolecular electron correlation energy. The discussion above neglects stabilization effects present already at the uncorrelated (Hartree-Fock) level, such that the role of the direct metal-metal interactions is even smaller than what is postulated here.

Comparing the two different crystal conformations the Au(I)-Hg(II) interaction is

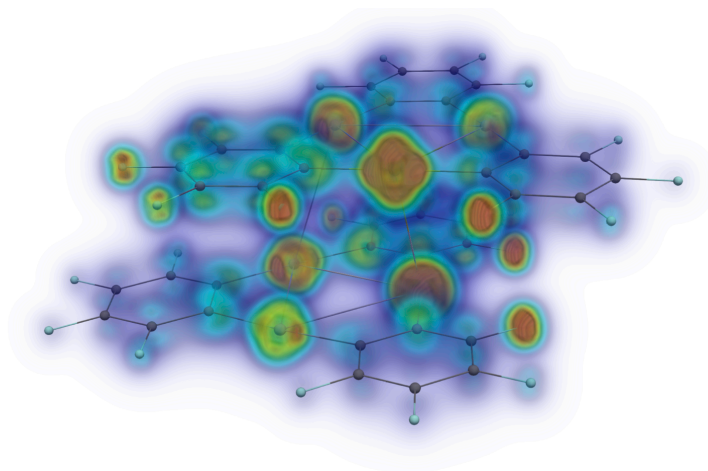


just slightly increased by bringing them closer together from **3b** to **3a**. Some of the values are slightly shifted but the overall intermolecular pair energies only differ by 6 kJ·mol<sup>-1</sup> favoring **3a**. The two dimers are effectively very close in energy.

Lastly, the interactions of a pure mercury complex are considered. For that, the gold atom of the middle plane was exchanged by mercury. As expected the HF energies become more repulsive for both dimer structures by about 70 kJ·mol<sup>-1</sup>. This is caused by the larger nuclear repulsion arising from mercury. The intermolecular pair energies compared to the complexes containing gold only differ by ~6 kJ·mol<sup>-1</sup> for both dimers favoring the Au complexes.

Analyzing the Hg(II)···Hg(II) close contacts, the interaction strength decrease compared to Au(I)···Hg(II) contacts at the same distance. Moreover, the total metal-metal and metal-ligands interactions decrease while the remaining ligands interactions increase. This effect can be explained by the smaller van der Waals radius of Hg. The larger core results in a higher orbital contraction which leads to a smaller overlap with the other monomer orbitals compared to gold.

To sum up, the hypothesis that this type of complex is determined by large Au(I)-Hg(II) interactions can not be confirmed. Furthermore, stating that the complex formation is driven by the trinuclear metal center misses the dominant forces at work in the crystal, the ligand interactions. The only observation in line with the previous studies is that Au(I)-Hg(II) interactions are stronger than pure mercurophilic interactions.



**Figure 6.10:** Voxel DID at SCS-LMP2/AVTZ for dimer **3a**. Reproduced from Reference [54] with the permission of the American Chemical Society.

## 6.3 Conclusions

An interaction analysis for three different compounds containing the heavy  $d^{10}$  metals gold(I) and mercury(II) has been presented. For that, extensive use of local orbital approaches combined with SCS-LMP2 was made. The electron correlation energy was decomposed in fragments and therefore, allowed to identify the relative weight of ligands and metal centers stabilizing the complexes.

The calculations carried out undoubtedly show that even if favorable short-distance Au...Au contacts are observed in the crystal, this serves as no indication of the driving force. It is necessary to consider the latter relative weight. For the ClAuMe<sub>2</sub>bimy and ClAuEt<sub>2</sub>bimy crystals this appears to be clearly outweighed by the ligand-ligand interactions, which are mainly of  $\pi$ - $\pi$  character. Even in the case of the methyl substituted carbene ligand, where the aurophilic bond is observed, the structure is determined by an optimal arrangement of the rings, a face-to-face orientation.

For the AuHg<sub>2</sub>(o-C<sub>6</sub>F<sub>4</sub>)<sub>3</sub> - Hg<sub>3</sub>(o-C<sub>6</sub>F<sub>4</sub>)<sub>3</sub> crystal no structure-inducing strong Au(I)...Hg(II) contacts were found. The ligand-ligand interaction adds for the overall system up to nearly the half of the intermolecular electron correlation energy. Even though there are 6 heavy  $d^{10}$  metals interacting per dimer, the pure metal-metal interactions only play a small role. Furthermore, evidence was given that Au(I) and Hg(II) centers will interact stronger than in the case of Hg(II)/Hg(II).

The existence of attractive interactions between closed-shell cationic centers is, without a doubt, fascinating and enticing. It is also clear that such interactions contribute to smaller intermolecular distances in crystals. They partly counteract the electrostatic repulsion between the metal centers. Considering also the ratio  $\Delta E/\text{atom}$ , there are good reasons to highlight the strength of the aurophilic interaction. However, mistaking such an interaction for the main driving force in molecular aggregation should be avoided. The calculations carried out for the gold complexes show that simple changes in the ligands can lead to crystals where even no aurophilic contacts are observed. In some cases, metals may just be found in close contact because the ligand interactions dictate it, respecting the monomer symmetry and the templating effect of the metal. The  $d^{10}$  cations can contribute in stabilizing a molecular crystal, but the ligand composition seems a main factor. This should always be considered even if gold shines a bit brighter.

# 7 Energy Decomposition Analysis for PNO-LMP2

The quantification of intermolecular interactions has been an important field in computational chemistry for several years. The derivation of physical quantities like electrostatics or dispersion can serve as a guideline for the tailored design of new drugs or molecular catalysts.<sup>[205–207]</sup> Over the past decades, the seminal work of Kitaura and Morokuma<sup>[95,96]</sup> has inspired many new analytical methods for the determination of such physical quantities.<sup>[97–100]</sup> Thereby, local orbital spaces have been used successfully in various methods, with one example given by the extensions to SAPT-based approaches presented by the Sherrill group.<sup>[31–33]</sup> The basis to analyze the interaction energies obtained by local correlation methods goes back to the work of Schütz *et al.*<sup>[47]</sup> They presented a scheme which decomposes the double excitations from MP2 or CCSD calculations into different contributions by the use of local molecular orbitals and projected atomic orbitals as occupied and virtual spaces, respectively. The latest generation of local correlation methods uses pair natural orbitals (PNOs) for the latter. The computational cost of such methods scales linearly in regard to the molecular system size.<sup>[39,42,44]</sup> These features allow the treatment of large molecular clusters, containing hundreds of atoms, at quantum mechanical accuracy. Also for these methods it is of interest to have tools that allow a deep insight into the physical forces that establish molecular binding.

The group of Neese presented an approach to determine intermolecular physical forces such as dispersion, charge transfer or electrostatics on the basis of DLPNO-CCSD(T).<sup>[1]</sup> Their local energy decomposition (LED) scheme consists of an analysis of interactions at the Hartree-Fock and correlated levels. In this chapter the focus lies on the decomposition of the dynamic electron correlation part. In order to decompose the interaction energy at the correlated level, a similar scheme as presented by Schütz *et al.* is applied. Instead of ionic contributions in which all charge trans-

fer contributions are summarized, charge transfer terms between different fragments are introduced. Furthermore, dynamic double polarization excitations are not omitted, because the spatial location of PNOs is not known at first glance. Effectively, there is no possibility to remove these excitations by an *ad hoc* criterion. To decompose the correlated interaction energy, the spatial distribution of the PNOs has to be determined once the CC equations have been solved. For this purpose, the obtained PNOs are localized by the Pipek-Mezey scheme and afterwards assigned to fragments by Mulliken population analysis. The amplitudes and integrals are then transformed to the localized PNO basis to finally derive the decomposed energy contributions.

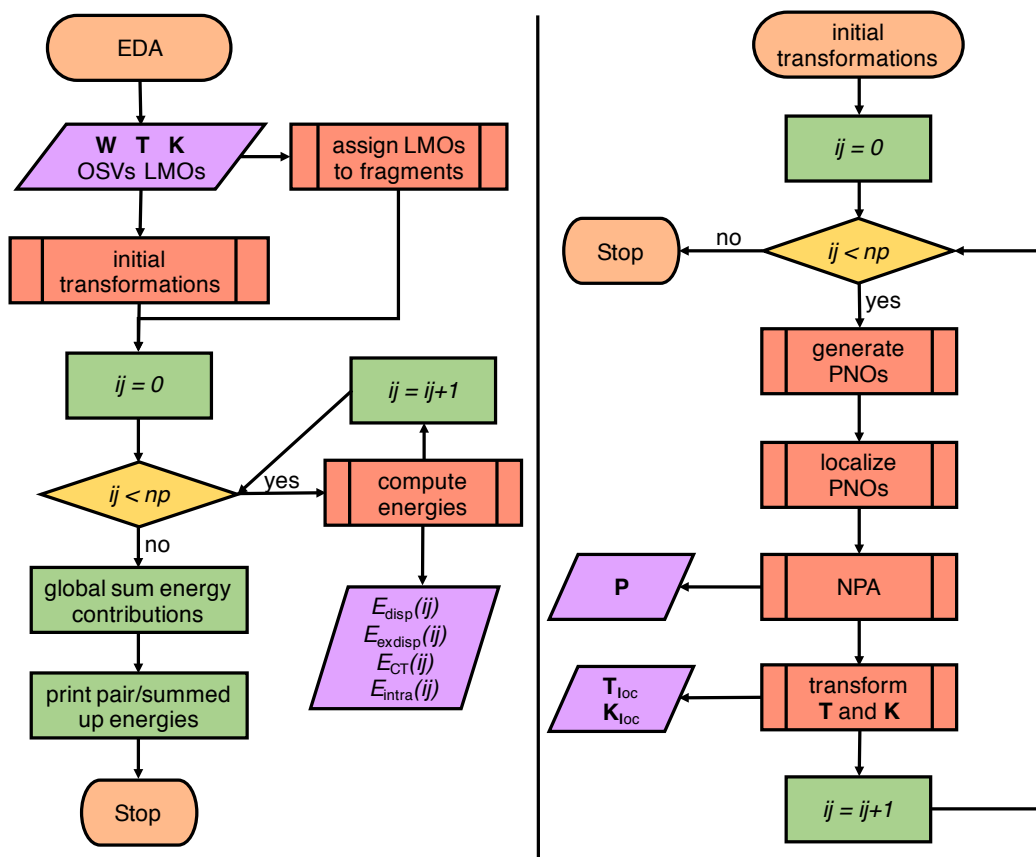
For dispersion interactions it was shown that the approach converges well with the basis set size. However, the charge transfer energies and intramolecular interactions are unstable if large augmented basis sets are applied. To investigate this behavior, a similar scheme has been implemented for the PNO-LMP2 code<sup>[42,43]</sup> in a development version of Molpro 2018.2.<sup>[182]</sup> The influence of different localization methods and population analyses on the energy contributions was examined on that basis.

### 7.1 Implementation

The following energy partitioning scheme is based on a preceding PNO-LMP2 calculation. In contrast to DLPNO-CCSD(T) implemented in ORCA,<sup>[159]</sup> the PNOs are obtained by OSVs in an intermediate step. This brings two advantages.<sup>[42]</sup> First, the OSV to PNO transformation matrices are much smaller than the corresponding PAO to PNO transformation matrices. Therefore, less memory is required. Secondly, the PNOs of a pair domain can be determined by an energy threshold based on an initial LMP2 evaluation in the OSV basis. In earlier implementations the natural occupation numbers of the PNOs served as a guideline to determine the pair domain. However, if large basis sets are applied the PNO occupation numbers become small and the domain definition becomes non-trivial. Loose thresholds of the occupations numbers do not capture enough PNOs to properly recover the interaction energy, while tight thresholds unnecessarily increase the size of pair domains. The subsequent energy decomposition analysis was implemented in a development version of Molpro 2018.2.<sup>[182]</sup> A flowchart for this scheme is depicted in Figure 7.1.

The application of the decomposition scheme requires an assignment of the occupied and virtual orbitals to interacting fragments. The occupied space is localized by intrinsic bond orbitals.<sup>[86]</sup> Thereby, intrinsic atomic orbital charges are computed, which can then be used to determine the spatial placement. To exclude atoms which only have small contributions to an orbital cut off of 0.1  $e$  is used. Thus, for each orbital, the atoms over which it is spanned are preserved. With this information, the orbitals can then be assigned to fragments. Occupied orbitals shared between fragments will be removed from the interaction analysis.

To assign PNOs to fragments, a sophisticated approach is necessary. The PNOs of a pair domain are often strongly delocalized, making it difficult to unambiguously assign them to fragments. By localizing the PNO space for each pair domain, with a standard procedure like Pipek-Mezey<sup>[83]</sup> or Boys,<sup>[82]</sup> this problem can already



**Figure 7.1:** Flowchart of the energy decomposition analysis implementation for PNO-LMP2. On the left side the main program computing the energy contributions  $E_{disp}$ ,  $E_{exdisp}$ ,  $E_{CT}$  and  $E_{intra}$  is shown. The right side contains the sub-routine which localizes the PNOs, determines their locality  $\mathbf{P}$  and transforms the amplitudes  $\mathbf{T}$  and integrals  $\mathbf{K}$  in the local PNO basis ( $\mathbf{T}_{loc}$  and  $\mathbf{K}_{loc}$ ).

be solved for most virtual orbitals. To account for the PNOs which still remain shared between fragments a sophisticated approach by the use of natural population analysis<sup>[93,94]</sup> is introduced as discussed later in this section. On the right side of Figure 7.1 the necessary steps for the spatial determination of the PNOs are summarized under ‘initial transformations’ in a flow chart.

The local PNO-MP2 calculations using the Hylleraas functional only require the amplitudes and integrals. In the initial implementation the latter are derived by transforming the OSV amplitudes and integrals in the PNO basis without creating the PNO coefficient matrices. Therefore, in order to localize the PNOs they must first be built from OSVs

$$\mathbf{C}_{\text{PNO}}^{ij} = \mathbf{C}_{\text{OSV}}^{ij} \mathbf{W}^{ij} , \quad (7.1)$$

with  $\mathbf{C}_{\text{PNO}}^{ij}$  and  $\mathbf{C}_{\text{OSV}}^{ij}$  being the PNO and OSV coefficient matrices and  $\mathbf{W}^{ij}$  as transformation matrix from the OSV to the PNO basis for an orbital pair  $ij$ . Afterwards, the PNOs are localized in the same fashion as occupied orbitals. The choice of the localization method has a huge influence on the robustness of the scheme as will be discussed later in this chapter. Once the localized PNOs are obtained as  $\mathbf{C}_{\text{PNO-loc}}^{ij}$ , the transformation matrix from the initial to the localized PNOs is computed as

$$\mathbf{U}^{ij} = \mathbf{C}_{\text{PNO}}^{ij} \mathbf{S} \mathbf{C}_{\text{PNO-loc}}^{ij} . \quad (7.2)$$

Next the integrals and amplitudes are transformed to the localized PNO basis by unitary transformations

$$\mathbf{T}_{\text{loc}}^{ij} = \mathbf{U}^{ijT} \mathbf{T}^{ij} \mathbf{U}^{ij} , \quad (7.3)$$

$$\mathbf{K}_{\text{loc}}^{ij} = \mathbf{U}^{ijT} \mathbf{K}^{ij} \mathbf{U}^{ij} . \quad (7.4)$$

As a last step of the ‘initial transformations’ natural populations analysis is used to determine the spatial distribution of the PNOs. Thereby, the PNO population matrix  $\mathbf{P}$  is build, which has the dimension  $m \times n_{\text{PNO}}$  with  $m$  being the number of fragments and  $n_{\text{PNO}}$  the total number of PNOs. The matrix stores the percentual

weighting of the PNOs on the fragments, which is derived in a similar fashion as described for LMOs in section 3.1. For simplification the case with two fragments  $X$  and  $Y$  will be discussed.

The energy contributions are then derived by following the decomposition scheme. LMOs  $i$  are excited to PNOs  $a$ , the same applies to  $j$  and  $b$ . For the dispersion and exchange dispersion interactions the orbitals  $i$  and  $j$  are located on different fragments  $X$  and  $Y$ . The contributions are then obtained by

$$\Delta E_{\text{disp}}(ij) = \sum_{ab}^{\text{PNOs}} \tilde{T}_{ab}^{ij} K_{ab}^{ij} P_a^X P_b^Y, \text{ with } i \in X, j \in Y, \quad (7.5)$$

$$\Delta E_{\text{exdisp}}(ij) = \sum_{ab}^{\text{PNOs}} \tilde{T}_{ab}^{ij} K_{ab}^{ij} P_a^Y P_b^X, \text{ with } i \in X, j \in Y, \quad (7.6)$$

with  $P_a^X$  being the percentual weight of PNO  $a$  on fragment  $X$  and  $\tilde{T}_{ab}^{ij}$  the contravariants of the amplitudes in the localized PNO basis. In terms of charge transfer contributions one has to distinguish between intramolecular and intermolecular orbital pairs. In case of an intramolecular pair the charge transfer energy is given by

$$\Delta E_{\text{CT}_{X \rightarrow Y}}(ij) = \sum_{ab}^{\text{PNOs}} \tilde{T}_{ab}^{ij} K_{ab}^{ij} (P_a^Y P_b^Y + P_a^X P_b^Y + P_a^Y P_b^X), \text{ with } i, j \in X, \quad (7.7)$$

For an intermolecular pair the charge transfer contribution is computed as

$$\Delta E_{\text{CT}_{X \rightarrow Y}}(ij) = \sum_{ab}^{\text{PNOs}} \tilde{T}_{ab}^{ij} K_{ab}^{ij} P_a^Y P_b^Y, \text{ with } i \in X, j \in Y. \quad (7.8)$$

The intramolecular energy term is derived by excitations from LMOs to PNOs lo-

cated in the same fragment

$$\Delta E_{\text{intra}}(ij) = \sum_{ab}^{\text{PNOs}} \tilde{T}_{ab}^{ij} K_{ab}^{ij} P_a^X P_b^X, \text{ with } i, j \in X. \quad (7.9)$$

Both main procedures ‘EDA’ and ‘initial transformations’, shown in Figure 7.1, are programmed in parallel using global arrays. To each processor a number of orbital pairs is assigned, which it processes one after the other. In an earlier implementation all transformation matrices were computed first and stored in memory before transforming the integrals and amplitudes. However, this was disadvantageous due to the large memory storage needed. In the here presented implementation the transformation matrices are built to directly transform the amplitudes and integrals. Therefore, each processors only needs a memory storage  $n_{\text{PNOmax}} \times n_{\text{PNOmax}}$ , with  $n_{\text{PNOmax}}$  being the number of PNOs for the largest pair domain. After finishing the energy evaluation for all pairs the results of the processes are merged and the output is generated. Finally, the energy contributions for each pair and the overall interaction energies are printed. Additionally, Dispersion Interaction Densities and charge transfer densities, which will be introduced later in this chapter, can be computed.

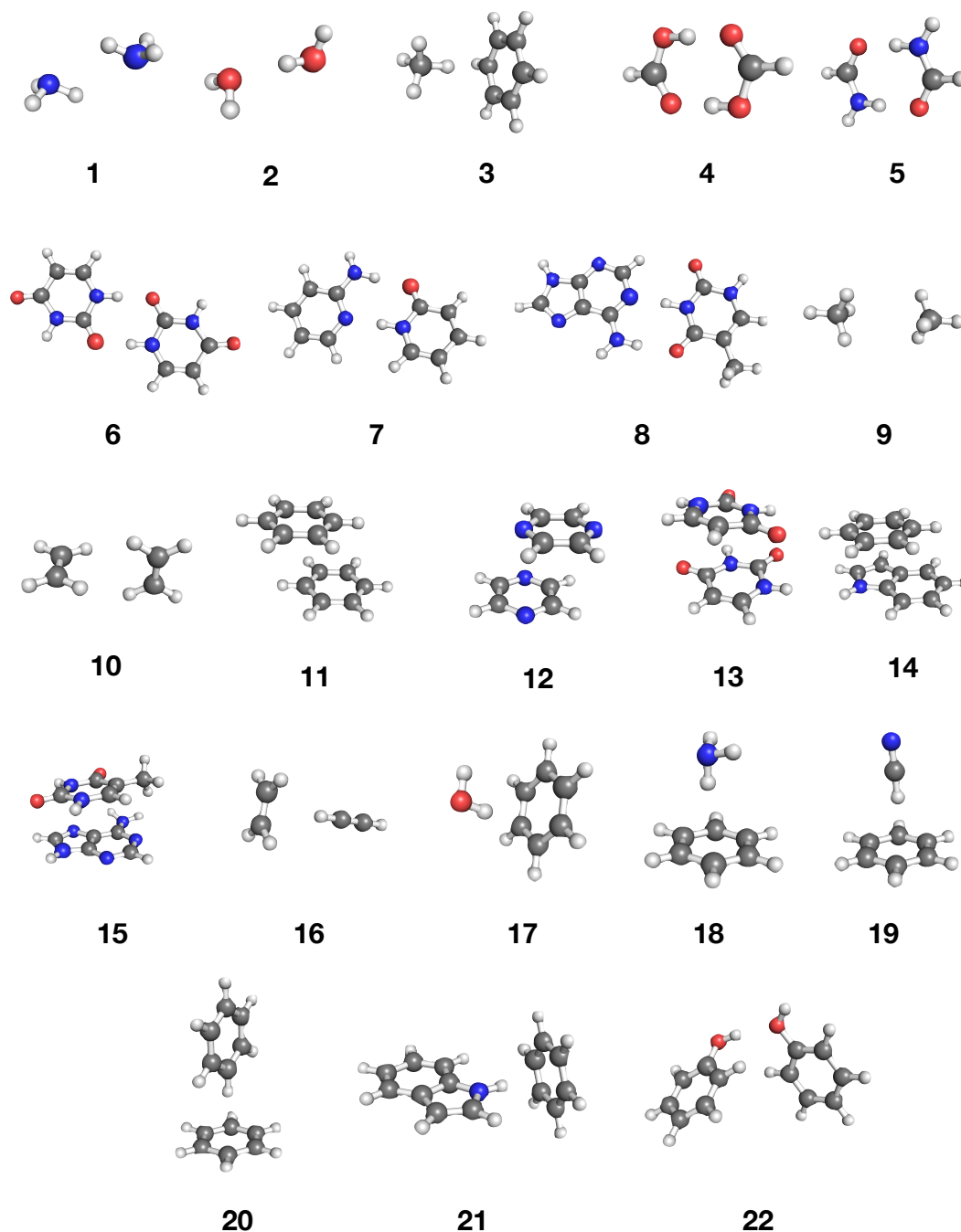
## 7.2 Results for Intermolecular Interactions

The implemented EDA scheme was examined for its stability by calculations on the S22 benchmark set for non-covalent interactions.<sup>[65]</sup> The latter provides the structures shown in Figure 7.2, ranging from the water dimer (**1**) up to adenine thymine complexes (**8** and **15**) (retained names are given in Table 7.1). In this section, the focus is on the study of the intermolecular interactions, namely dispersion and charge transfer.

The conformational arrangement of some S22 structures already gives some idea of the binding interactions involved. One example are the benzene dimers, where it is well-known that the parallel displaced arrangement **11** is dominated by dispersion, while in the T-shaped conformation **20** electrostatics become more important. Therefore, the analysis should be able to reflect these trends. However, the consistency of the method has to be ensured first. It has to converge with the basis set



size and the dispersion energies should behave with an  $r^{-6}$  dependence. For this purpose, all S22 structures were computed with Dunning's correlation consistent basis sets from double up to quintuple zeta, with and without diffuse functions.<sup>[121,122]</sup> The calculations were carried out with a development version of Molpro 2018.2. Thereby, the initial Hartree-Fock calculations were computed with the local den-



**Figure 7.2:** Structures of the S22 benchmark set. The retained names are provided in Table 7.1.

**Table 7.1:** Retained names of the S22 structures provided in Figure 7.2.

structure	retained name
<b>1</b>	ammonia dimer
<b>2</b>	water dimer
<b>3</b>	benzene methane complex
<b>4</b>	formic acid dimer
<b>5</b>	formamide dimer
<b>6</b>	uracil dimer h-bonded
<b>7</b>	2-pyridoxine 2-aminopyridine complex
<b>8</b>	adenine thymine Watson Crick complex
<b>9</b>	methane dimer
<b>10</b>	ethene dimer
<b>11</b>	benzene dimer parallel displaced
<b>12</b>	pyrazine dimer
<b>13</b>	uracil dimer stacked
<b>14</b>	indole benzene complex stacked
<b>15</b>	adenine thymine complex
<b>16</b>	ethene ethyne complex
<b>17</b>	benzene water complex
<b>18</b>	benzene ammonia complex
<b>19</b>	benzene cyanide complex
<b>20</b>	benzene dimer T-shaped
<b>21</b>	indole benzene complex T-shaped
<b>22</b>	phenol dimer

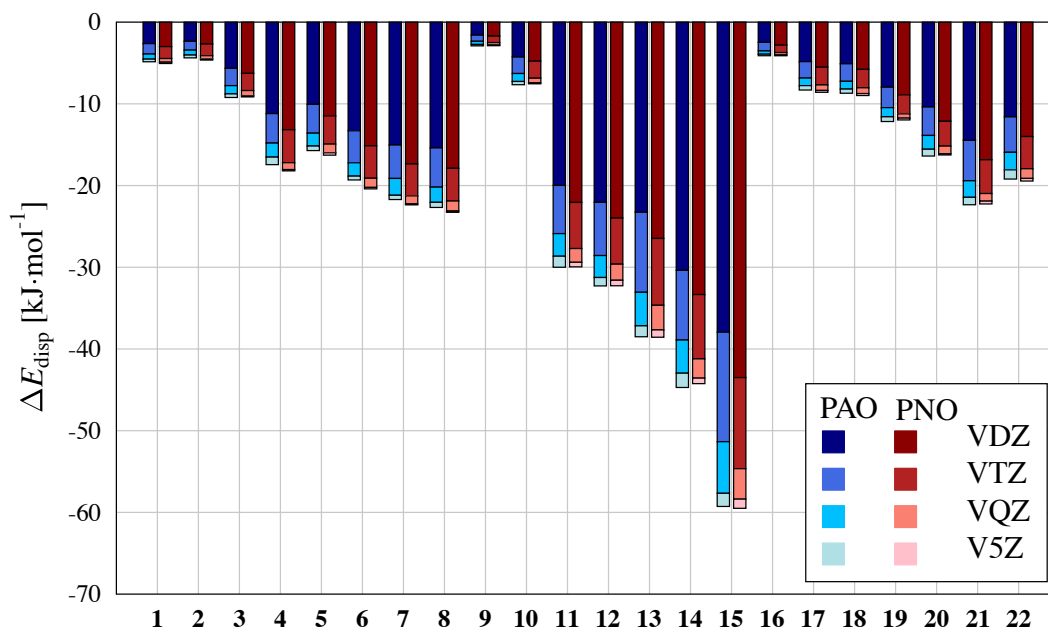
sity fitting approach.<sup>[181]</sup> The PNO-LMP2 calculations were performed with the PAO-OSV-PNO-LMP2 code implemented by the Werner group.<sup>[42,43]</sup> Thereby, the occupied space was localized by intrinsic bond orbitals.<sup>[86]</sup> The PNOs were generated by the default settings. PAO-LMP2 calculations were performed with density fitting<sup>[114]</sup> and Pipek-Mezey<sup>[83]</sup> localization. The corresponding orbital domains were determined by the default Boughton-Pulay criterion.<sup>[84]</sup> The corresponding defaults were used for the fitting basis sets.<sup>[123,156]</sup>

The PNO-EDA schemes discussed in this chapter differ mainly in the applied localization schemes and population analyses to determine the spatial distribution of the PNOs. In several publications it has been shown that the localization method used for the valence orbitals does not significantly affect the interaction energies and therefore, its discussion can be neglected.<sup>[84]</sup> The proposed set up localizes the PNOs by the Boys scheme. Afterwards, the PNOs are assigned to fragments using natural population analysis. Thereby, PNOs shared between fragments are considered with their percentual portion to the fragments as described in section 7.1. We start with a closer look to the dispersion contributions obtained by this scheme. In Figure

7.3 the results of PAO- and PNO-EDA are compared for non-augmented cc-pVXZ (VXZ) basis sets.

Both methods converge with only very small deviations to the same values for all structures. Moreover, it is observed that the PNO approach converges faster regarding the basis set size. In the case of PAOs, the virtual space is constructed by an *ad hoc* criterion and restricted to the near vicinity of the corresponding LMOs. Therefore, the PAO domains need to become large enough to properly receive the interaction energy. In the case of PNOs, the domains are constructed from an extended PAO space, containing the PAOs of neighboring atoms.<sup>[42]</sup> This highlights that by the strong restriction of the virtual space in case of PAOs the dispersion contributions are underestimated for small basis sets.

By adding augmented functions (AVXZ) the dispersion contributions are already close to convergence with a double zeta basis set, as shown in Figure 7.4. These results are quite encouraging, if one is interested in the influence of dispersion forces for large molecular systems. Increasing the basis set size the dispersion contributions stagnate quickly. For structures with a missing AV5Z bar, a problem arises with the linear dependency of the basis.

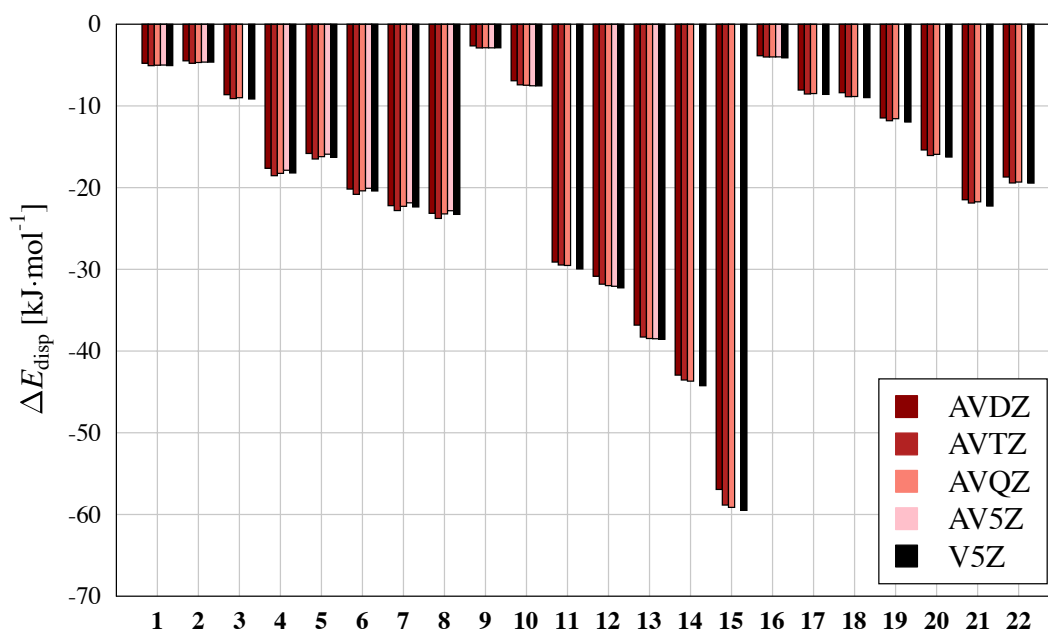


**Figure 7.3:** Comparison of dispersion energies for the S22 benchmark set for cc-pVXZ basis sets. PAO results are shown in blue tones, whereas PNO results are in red tones.

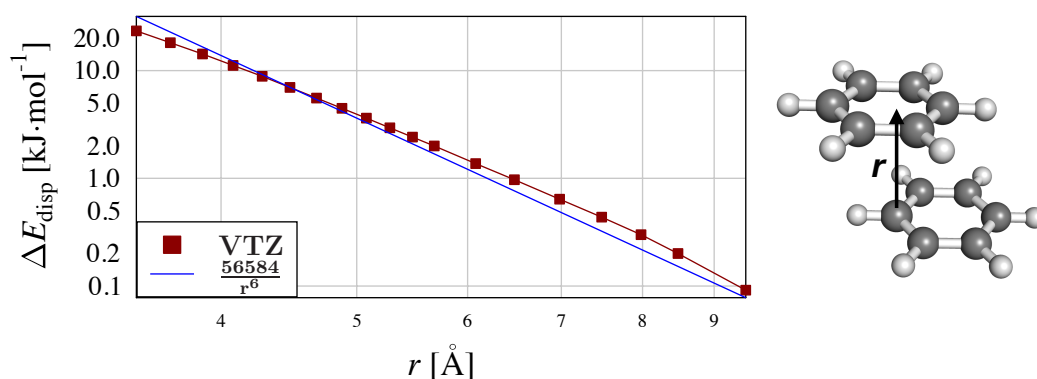
## 7 Energy Decomposition Analysis for PNO-LMP2

To study the distance behavior of the dispersion interactions, the parallel displaced benzene dimer was computed at different distances using a VTZ basis. The results are plotted with a double logarithmic scale (Figure 7.5). As it can be seen, the dispersion contributions are in close agreement to a  $r^{-6}$  decay.

Having verified the proper description of dispersion interactions, the focus moved to the charge transfer contributions. However, a comparison to PAO-LMP2 is mean-



**Figure 7.4:** Comparison of dispersion energies for the S22 benchmark set with aug-cc-pVXZ basis sets and the cc-pV5Z basis set in red tones and black, respectively.



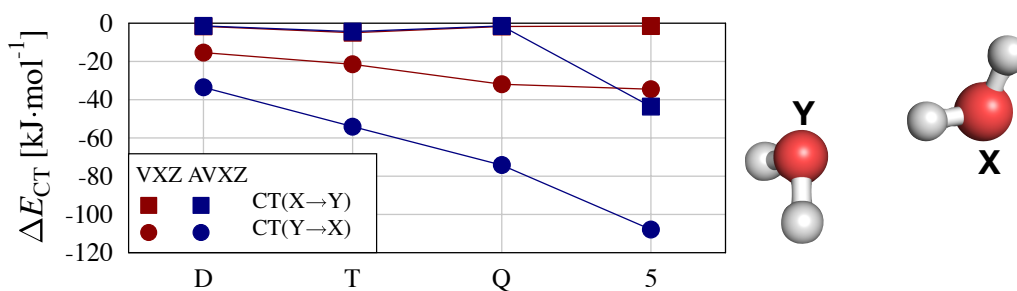
**Figure 7.5:** Dispersion contributions for a parallel displaced benzene dimer in dependency of the distance. Both axis are plotted as a logarithmic scale. For the dispersion contributions the logarithm of the absolute values was taken.

ingless since several excitations terms are omitted by definition. The LED analysis of the Neese group applies a different partitioning scheme to decompose the interaction energies. In that case, the calculated charge transfer and intramolecular contributions diverge, especially when large augmented basis sets are used for highly polarized complexes. One example is given by the water dimer depicted in Figure 7.6.

The investigations revealed that this behavior is mainly due to the poor description of the electron population by applying the Mulliken analysis<sup>[91]</sup> for large basis sets. Furthermore, the Pipek-Mezey localization turned out to be disadvantageous, since the latter produces  $\pi$ -like PNOs, which are in several cases strongly delocalized over different fragments. This feature hinders a clear determination of the PNOs spatial location. Moreover, as expected the strongest contributing excitations, are into close lying PNOs. Therefore, a spatial localization criterion like Boys should be favored, minimizing the number of PNOs shared over fragments. To further account for the PNOs which are still shared between fragments after their localization, the percentual weighting according to the NPA charges is introduced.

The obtained charge transfer energies for the S22 structures are provided for cc-pVXZ and aug-cc-pVXZ basis sets, in Figures 7.7 and 7.8, respectively. To show the overall convergence behavior, the two contributions  $E_{CT_{X \rightarrow Y}}$  and  $E_{CT_{Y \rightarrow X}}$  are summed up.

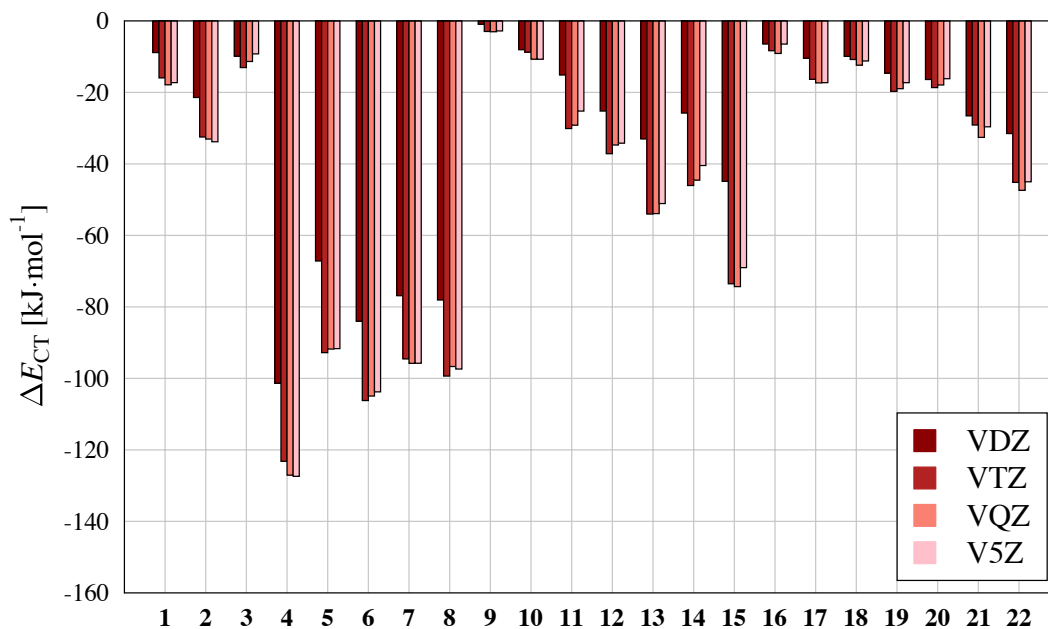
As observed in Figure 7.7 the method performs rather well for non-augmented basis sets. Interestingly, for complexes which are more electrostatically driven (like the systems 4-8) the scheme seems to be almost converged with a quadruple zeta basis set, while for dispersion dominated systems (like 10-15) a significant decrease



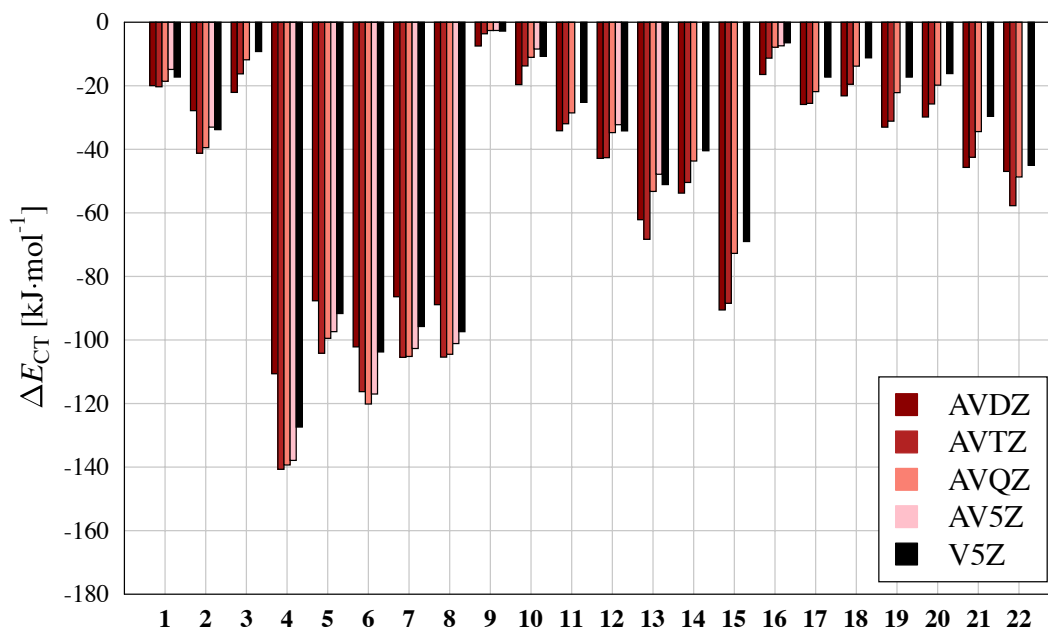
**Figure 7.6:** Charge transfer contributions for a water dimer computed with LED implemented in ORCA 4.0.0.2.<sup>[159]</sup>

## 7 Energy Decomposition Analysis for PNO-LMP2

from VQZ to V5Z is observed in several cases. By adding augmented functions the charge transfer component fluctuates, as shown in Figure 7.8. This should mainly rely on the larger overlap of the basis functions and the resulting higher BSSE for



**Figure 7.7:** Comparison of charge transfer energies for the S22 benchmark set with cc-pVXZ basis sets.  $\Delta E_{CT}$  is derived as the sum of  $\Delta E_{CT_{X \rightarrow Y}}$  and  $\Delta E_{CT_{Y \rightarrow X}}$ .



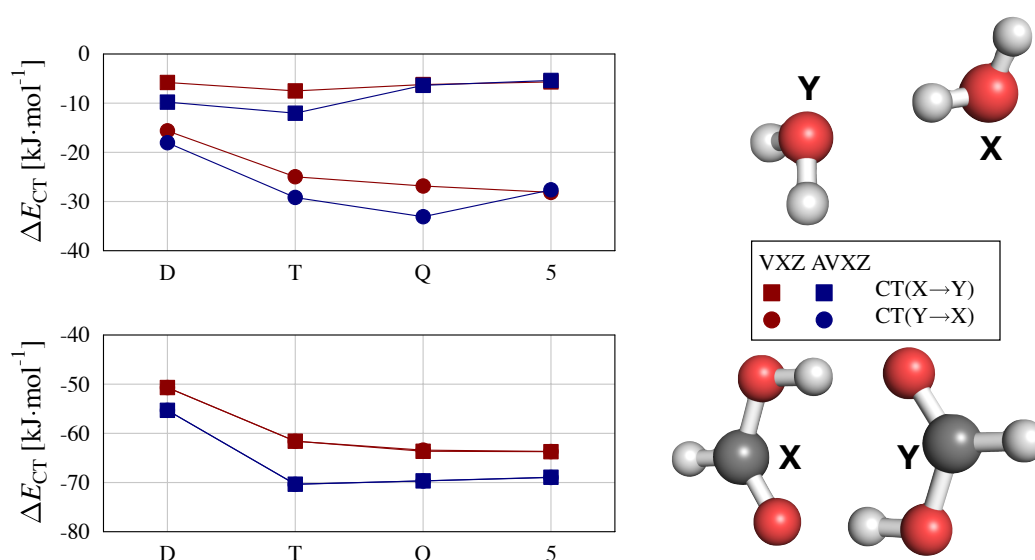
**Figure 7.8:** Comparison of charge transfer energies for the S22 benchmark set with aug-cc-pVXZ and cc-pV5Z basis sets, in red tones and black, respectively.  $\Delta E_{CT}$  is derived as the sum of  $\Delta E_{CT_{X \rightarrow Y}}$  and  $\Delta E_{CT_{Y \rightarrow X}}$ .

small basis sets. Compared to the results at V5Z, the contributions for the strong electrostatically bound systems **4-9** deviates at most by  $15 \text{ kJ}\cdot\text{mol}^{-1}$  for the formic acid dimer (**4**). By splitting up the charge transfer in its two components,  $\Delta E_{\text{CT}_{X\rightarrow Y}}$  and  $\Delta E_{\text{CT}_{Y\rightarrow X}}$ , the maximum deviation is given by  $7 \text{ kJ}\cdot\text{mol}^{-1}$  at an absolute value of  $70 \text{ kJ}\cdot\text{mol}^{-1}$ .

The examination of the individual charge transfer components provided the expected results for all structures. Two examples, water and formic acid dimers, are highlighted in Figure 7.9. For the water dimer it is expected that one monomer polarizes the other more than vice versa, while for the symmetric formic acid dimer both contributions should display the same values. These trends are well reflected. The contributions at least converge to similar values for basis sets with and without diffuse functions.

To obtain further insight into the charge transfer components a similar visualization technique as for dispersion interactions was implemented in a development version of Molpro 2018.2. Thereby, only the cross excitations are considered to build the density matrix, as shown in Figure 7.10.

In order to build the charge transfer density matrix  $\mathbf{D}^{X\rightarrow Y}$  intra- (Equations 7.10 and 7.12) and intermolecular (Equations 7.11 and 7.13) pairs have to be distinguished.



**Figure 7.9:** Analysis of the individual charge transfer components for the water and formic acid dimers.

$$D_{\mu\nu}^X = \sum_{ij}^{\text{LMOs PNOs}} \sum_{ab} \tilde{T}_{ab}^{ij} K_{ab}^{ij} [P_a^Y P_b^Y (C_{\mu i} C_{\nu i} + C_{\mu j} C_{\nu j}) + P_a^Y P_b^X (C_{\mu i} C_{\nu i}) + P_a^X P_b^Y (C_{\mu j} C_{\nu j})], \text{ with } i \in X, j \in Y \quad (7.10)$$

$$D_{\mu\nu}^X = \sum_{ij}^{\text{LMOs PNOs}} \sum_{ab} \tilde{T}_{ab}^{ij} K_{ab}^{ij} [P_a^Y P_b^Y (C_{\mu i} C_{\nu i})], \text{ with } i, j \in X \quad (7.11)$$

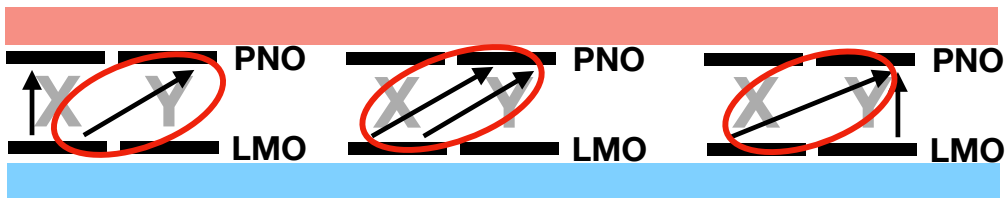
$$D_{\mu\nu}^Y = \sum_{ij}^{\text{LMOs PNOs}} \sum_{ab} \tilde{T}_{ab}^{ij} K_{ab}^{ij} [P_a^Y P_b^Y (C_{\mu a} C_{\nu a} + C_{\mu b} C_{\nu b}) + P_a^Y P_b^X (C_{\mu a} C_{\nu a}) + P_a^X P_b^Y (C_{\mu b} C_{\nu b})], \text{ with } i \in X, j \in Y \quad (7.12)$$

$$D_{\mu\nu}^Y = \sum_{ij}^{\text{LMOs PNOs}} \sum_{ab} \tilde{T}_{ab}^{ij} K_{ab}^{ij} [P_a^Y P_b^Y (C_{\mu a} C_{\nu a})], \text{ with } i, j \in X \quad (7.13)$$

$$\mathbf{D}^{X \rightarrow Y} = \mathbf{D}^X - \mathbf{D}^Y \quad (7.14)$$

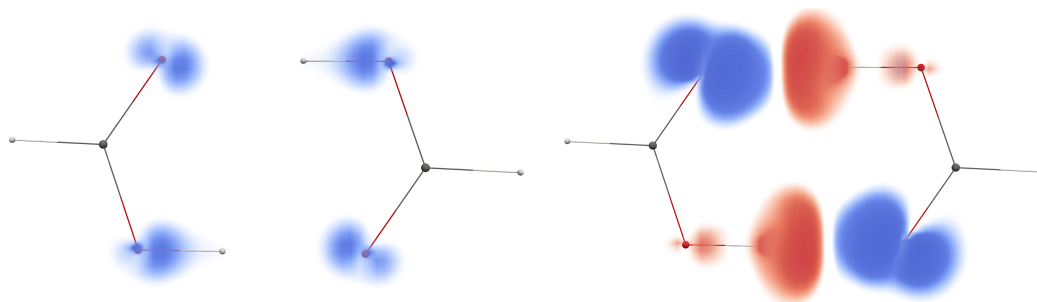
The closed shell densities of the LMOs and PNOs corresponding to cross excitations, are scaled with their energy contributions. The subtraction defines a positive (LMO) range and a negative (PNO) range. Using a visualization tool, the densities can be displayed in a similar fashion as DIDs (the plots were prepared with Paraview<sup>[112]</sup>). An example is given for the formic acid dimer in Figure 7.11.

In the charge transfer plot (right side of Figure 7.11) both densities,  $\mathbf{D}^{X \rightarrow Y}$  and  $\mathbf{D}^{Y \rightarrow X}$ , are plotted simultaneously. The differences to dispersion interactions already become apparent. While the dispersion interaction between certain elements can be well described by an  $r^{-6}$  decay, the charge transfer component highlights strongly polarized regions of the molecules. In combination with the robustness of the examined scheme one has the possibility to investigate the nature of these contributions.



**Figure 7.10:** Excitation scheme to build the charge transfer density. Only the red rimmed excitations are considered.



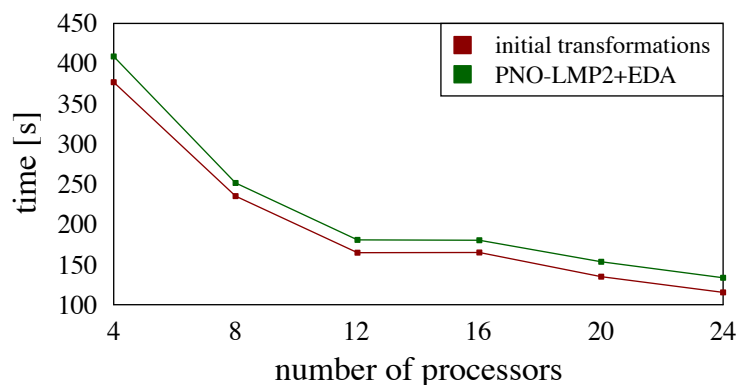


**Figure 7.11:** Comparison of the dispersion (left side) and charge transfer (right side) interaction densities for the formic acid dimer.

## 7.3 Time Scaling

As shown in the previous Section, the EDA was significantly stabilized by a strong spatial localization criterion in combination with a reliable population analysis, especially for the charge transfer energies. However, this also brings its disadvantages, which is mainly the time scaling. The bottleneck of the whole scheme is the localization of the PNOs. The Boys algorithm scales with  $N^3$  regarding the basis set size, while the Pipek-Mezey scheme scales with  $N^2$ . However, the Pipek-Mezey scheme produces  $\pi$ -like PNOs which are at least partly delocalized over fragments and thus, hinder the clear assignment of PNOs to fragments. Moreover, the natural population analysis is more time consuming in comparison to the Mulliken analysis, because several transformations are needed. Nonetheless, the application of the computationally cheaper Mulliken or Löwdin analysis should be avoided, since they do not converge with increasing the basis set size. An example for the time scaling of the localization processes is given for the benzene methane complex, provided in Figure 7.12.

As can be seen in the plot, the initial transformations procedure claims nearly 90% of the total computation time (PNO-LMP2+EDA). Furthermore, the computing time is not halved by doubling the number of processors. These scaling issues are mainly due to the different complexity of the PNO spaces. To each processor a certain number of orbital pairs is assigned, whose virtual spaces have to be localized. It can be observed that the localization procedure for some orbital pairs converges very slow and therefore, takes more computing time. While some processors have already finished their pairs, they have to wait for the remaining processes until the EDA can proceed. For larger molecular systems than discussed here the number of orbital pairs increases and it is expected that the speedup will be higher by taking more



**Figure 7.12:** Time scaling for the localization process (red) and the PNO-LMP2 calculation plus additional EDA (green) for the benzene methane complex with a cc-pVQZ basis set.

processors. A worker farmer model could possibly fix the problem since thereby all processors are used until all elements are calculated.

## 7.4 Conclusions

An energy decomposition analysis for PNO-LMP2 has been implemented into a development version of Molpro 2018.2 and probed on the S22 benchmark set for non-covalent interactions. The dispersion contributions converge rather well, regarding the basis set size. By adding augmented functions, already the double zeta basis set provides a good estimate of the dispersion interactions. These results are encouraging with respect to the analysis of large molecular clusters. Furthermore, it was shown that the dispersion decays with  $r^{-6}$  and that the results are in good agreement with the initial PAO-LMP2 implementation.

The charge transfer energies show a stable pattern if non augmented-basis sets are applied. It was observed that, especially for dispersion dominated systems, the energies become too large by applying small basis sets with diffuse functions. This should be linked to the larger BSSE compared to the non-augmented basis sets. In particular for the charge transfer energies it could be clearly shown that the PNOs have to be localized by a strong spatial criterion in order to obtain a stable scheme. Thereby, the Boys localization is favored over the PM-scheme, since the latter produces  $\pi$ -like PNOs which are often shared between fragments. Furthermore, the

population analysis has to be robust with regard to the basis set size. Therefore, unstable schemes like Mulliken or Löwdin analysis should be omitted.

A technique related to the Dispersion Interaction Density has been implemented for the charge transfer contributions. The resulting plots can be directly compared to the dispersion interactions by using the same plotting settings. The results for the S22 benchmark set suggest that the charge transfer contributions mainly compensate the overestimation of the polarization by HF.

The improvement of the stability is accompanied by a larger computational effort. The Boys localization as well as the natural population analysis are computationally more demanding than the PM localization and Mulliken analysis used in the ORCA LED. However, it was clearly shown that this is necessary to obtain stable results. Comparing the required computing times for PAO-LMP2 and PNO-LMP2 plus the additional EDA the PAO analysis was faster in all cases. Tests for large systems with a large number of processors still have to show whether the scheme brings a time advantage over PAO methods.



## 8 Summary

The study of non-covalent interactions has been an important task in computational chemistry since its beginnings. However, in order to accurately describe the delicate balance of forces, high-level correlation methods are required. For years these methods were limited to systems containing a few atoms/electrons. The development of local orbital approaches allows nowadays the application of such methods to large scale molecular systems.

For experimentalists, understanding the physical nature behind molecular aggregation is of great value, as it allows for the rational design of materials and/or catalysts. Some tools of computational chemistry are popular for these purposes, like the electrostatic potential map or the noncovalent interaction plots put forward by Yang and co-workers.<sup>[109]</sup> What these methods have in common is that they are easy in application and allow for a pictorial output facilitating an intuitive chemical interpretation. However, for electrostatic potential maps a grid has to be computed, which can become extremely time consuming for large molecules. Moreover, both methods do not allow for the quantification of intermolecular interactions. In order to obtain quantitative measures more sophisticated approaches are needed. One possibility is the powerful family of SAPT-approaches.<sup>[29,30,101]</sup> By using local orbital spaces the Sherrill group presented different extensions providing atomic and fragment-wise partitioning of the systems under investigation.<sup>[31,32]</sup> In addition, extensions were presented which facilitate a graphical representation of SAPT contributions.

In this thesis, a method to visualize dispersion interactions on the basis of local correlation methods has been presented, the Dispersion Interaction Density (DID).<sup>[55]</sup> It can be deduced from the closed-shell densities of orbitals and their contribution to the dispersion energy. Based on the DID, three dimensional grids can be generated. The latter can be visualized as contour plots, color projections on the electronic

density or Voxel graphics by suitable software. Local correlation methods only allow for a decomposition of correlation energy terms, while the SAPT ansatz covers all types of interactions (including electrostatics, exchange, and so forth). However, the decomposition used in this thesis is a direct by-product of a wave function calculation. It is an interpretation of readily available quantities. Furthermore, the intramolecular partitioning of SAPT (ISAPT) energies is carried out through the use of embedding Fock-potentials.<sup>[33]</sup> The latter requires a charge analysis, artificial splitting of the protons of covalently bond atoms and the construction of Coulomb and exchange embedding potentials. This leads to issues such as spurious polarization effects in the linker bonds. In the case of local calculations, the only approximation required is the definition of the local orbital spaces.

In order to investigate the applicability of the DID plots, computations on the benzene dimer, substituted benzene dimers, methanol anisole derivatives and a coupled diamondoid were performed. Moreover, SCS-LMP2 was compared to the QCISD(T) or CCSDT high-level approaches at CBS limit. It was shown that SCS-LMP2 performed rather well in all cases. For the substituted benzene dimers the dispersion contributions were compared to the contributions obtained by F-SAPT showing a good agreement between the two approaches. In combination with previous publications, these results also confirm the application of the SCS-LMP2 method to large molecular systems as discussed in Chapters 5 and 6. In terms of the visualization technique the Voxel representation was found to be the most suitable one. It allows to combine the density information using an opacity gradient together with a color coding highlighting strong interaction regions. Moreover, it was found to be most sensitive for small dispersion contributions. Furthermore, the choice of the localization method was discussed. The use of Pipek-Mezey localization or IBOs is preferred to the Boys method since the latter keeps  $\sigma$ - $\pi$  space separation and, therefore, allows for a chemical intuitive interpretation.

After demonstrating that local orbital analysis using SCS-LMP2 and DIDs is well suited to investigate dispersive interactions, we moved on to study the latter in large supramolecular clusters. Synthetic supramolecular chemistry is able to supply a plethora of container systems of variable size, shape and chemical makeup as model systems helping to understand non-covalent binding in nature.<sup>[134–137]</sup> During the cooperation within the SPP1807 the Clever group was able to synthesize by metal-mediated self-assembly three different coordination cages obtained by the modular variation of a bis-monodentate pyridyl ligand. The latter itself forms in

---

the presence of  $[\text{Pd}(\text{CH}_3\text{CN})_4]-(\text{BF}_4)_2$  an interpenetrated double cage with three pockets filled by  $\text{BF}_4^-$  ions.<sup>[161]</sup> This system can be activated through the exchange of the outer pocket ions by  $\text{Cl}^-$  or  $\text{Br}^-$  to bind neutral guest molecules in the middle pocket. A modification of the bis-monodentate pyridyl ligand by the substitution of adamantyl groups leads to a sterically overcrowded coordination cage, which is able to bind mono- and bis-anionic guest molecules.<sup>[59]</sup> Local correlation calculations were carried out to determine the strength of dispersion interactions with respect to the host-guest binding. The functionalization by adamantyl residues leads to a huge contact surface between the cage and guest molecules within the pocket region. The calculations suggest that the main contributors to adamantyl interaction with guests are Pauli repulsion and dispersion, the two of opposite signs. The DID plots show that the main contributions arise from the interactions between the guests  $\pi$ -system and the pocket surface. Motivated by these results the Clever group is currently extending the scope of dispersion energy donors (DEDs) that can be implemented into the ligand backbones.

Next the uptake of neutral guest molecules in the interpenetrated host system was studied. The outcome curtails the scope of neutral guest molecules that can be encapsulated by the latter. In terms of size and shape, 6-rings and norbornene-type compounds are favoured. The opposite arrangement of heteroatoms in bridged or unbridged 6-ring guests is favourable, indicating that orientation of lone pairs towards both cationic Pd(II) centers leads to a stabilizing two-point attachment. This assumption was supported by X-ray structural analysis of the host-guest complex containing DABCO and a conformational sampling using electronic structure calculations. Local orbital analysis revealed a correlation between the experimental free binding energy and computed dispersion contributions. The DID profiles showed that close C-H contacts to the interior of the pocket dominate the dispersion profile. By a rather simple approach, with regard to the complexity of the system, free binding energies could be determined. In general, the computed values agree well with the experimental estimates. In particular, in the case of large guest molecules like cyclooctane, it became obvious that a more sophisticated approach is required to account for the cage relaxation by incorporating such molecules. Therefore, three additional schemes to determine thermodynamic quantities were proposed.

The study of non-covalent interactions was continued by the analysis of metallophilic interactions. The special case of aurophilic gold-gold interactions has already been discussed in great detail both experimentally and theoretically.<sup>[203,208–210]</sup> However, questions about the influence of the ligands have been raised more frequently in

recent years.<sup>[192,193]</sup> Furthermore, there were reports of mixed metallophilic interactions that were deemed as exceptionally strong. In order to investigate such effects three different compounds, two aurophilic<sup>[62]</sup> and one mixed metallophilic containing gold(I) and mercury(II)<sup>[64]</sup> served as a guideline. The first two systems ClAuR<sub>2</sub>bimy (R=Me, Et) deviate in the substitution of the benzimidazol-2-ylidene group. X-ray analysis revealed aurophilic contacts just in case of the methyl substituted system. To determine the dominant interactions for both complexes SCS-LMP2 calculations were performed. Through local orbital analysis the electron correlation energy was decomposed in fragments and the relative weight of ligands and gold centers stabilizing the complexes were identified. The calculations carried out undoubtedly show that even if favorable short-distance Au···Au contacts are observed in the crystal, this serves as no indication of the driving force. It is necessary to consider the latter relative weight. Even in the case of the methyl substituted carbene ligand, where the aurophilic bond is observed, the structure is determined by an optimal arrangement of the rings, a face-to-face orientation. In conclusion it was observed that both systems are mainly determined by  $\pi$ - $\pi$  interactions.

The same computational procedure was applied to the mixed metallophilic complex. The analysis of the fragments revealed no structure-inducing strong Au(I)···Hg(II) contacts. The ligand-ligand interaction adds up to nearly half of the intermolecular electron correlation energy. Even though there are 6 heavy d<sup>10</sup> metals interacting per dimer, the pure metal-metal interactions only play a small role. However, evidence was given that Au(I) and Hg(II) centers will interact stronger than in the case of Hg(II)/Hg(II). The results for the examined systems in combination with the previous work of our group lead to the conclusion that by considering the ratio  $\Delta E/\text{atom}$ , there are good reasons to highlight the strength of the metallophilic interaction. However, mistaking such an interaction for the main driving force in molecular aggregation should be avoided. The calculations carried out for the ClAuR<sub>2</sub>bimy complexes show that simple changes in the ligands can lead to crystals where even no aurophilic contacts are observed. In some cases, metals may just be found in close contact because the ligand interactions dictate it, respecting the monomer symmetry and the templating effect of the metal. The d<sup>10</sup> cations can contribute in stabilizing a molecular crystal, but the ligand composition seems a main factor.

The examination of large molecular systems as in Chapter 5 were hampered by the time spent on PAO-based local correlation methods. The rediscovery of PNOs has made the latest generation of local correlation methods much more efficient.<sup>[39,42,44]</sup>



---

To benefit from this progress and inspired by the local energy decomposition analysis (LED) of Neese *et al.*,<sup>[1]</sup> an energy decomposition analysis for PNO-LMP2 has been implemented into a development version of Molpro 2018.2.<sup>[182]</sup> The latter was probed on the S22 benchmark set for non-covalent interactions. The dispersion contributions converge rather well, regarding the basis set size. By adding augmented functions, already the double zeta basis set provides a good estimate of the dispersion interactions. These results are encouraging with respect to the analysis of large molecular clusters. Furthermore, it was shown that the dispersion decays with  $r^{-6}$  and that the results are in good agreement with the initial PAO-LMP2 implementation. The charge transfer energies show a stable pattern if non augmented-basis sets are applied. It was observed that, especially for dispersion dominated systems, the energies become too large by applying small basis sets with diffuse functions. This should be linked to the larger BSSE compared to the non-augmented basis sets. In particular, for the charge transfer energies it could be clearly shown that the PNOs have to be localized by a strong spatial criterion in order to obtain a stable scheme. Thereby, the Boys localization is favored over the PM-scheme, since the latter produces  $\pi$ -like PNOs which are often shared between fragments. Furthermore, the population analysis has to be robust with regard to the basis set size. Therefore, unstable schemes like Mulliken or Löwdin analysis should be omitted. With the aim to provide a deeper insight into the charge transfer contributions a technique related to the Dispersion Interaction Density has been implemented. The resulting plots can be directly compared to the dispersion interactions by using the same plotting settings. The results for the S22 benchmark set<sup>[65]</sup> suggest that the charge transfer contributions mainly compensate the overestimation of the polarization by HF. The improvement of the stability compared to the initial LED scheme is accompanied by a larger computational effort. The Boys localization as well as the natural population analysis are computationally more demanding than the PM localization and Mulliken analysis used in the LED. However, it was clearly shown that this is necessary to obtain stable results. Comparing the required computing times for PAO-LMP2 and PNO-LMP2 plus the additional EDA the PAO analysis was faster in all cases. Test calculations for large systems using massive parallelization over many processors still have to show whether the scheme brings a time advantage over PAO methods.



# List of Abbreviations

A-SAPT	atomic partitioned SAPT
AO	atomic orbital
bimy	benzimidazol-2-ylidene
BP	Boughton-Pulay
BSSE	basis set superposition error
CBS	complete basis set
CC	Coupled-Cluster
CI	Configuration Interaction
CT	charge transfer
DED	dispersion energy donor
DF	density fitting
DFT	density functional theory
DFT-D	dispersion corrected DFT
DID	Dispersion Interaction Density
DLPNO	domain based local PNO
ECP	effective core potential
EDA	energy decomposition analysis
EPM	electrostatic potential map
F-SAPT	fragment-wise partitioned SAPT
FB	Forster-Boys
GGA	generalized gradient approach
HF	Hartree-Fock
IAO	intrinsic atomic orbital
IBO	intrinsic bond orbital
L(MP2/CCSD)	local (MP2/CCSD)
LCAO	linear combination of orbitals
LDF	local density fitting
LED	local energy decomposition

## List of Abbreviations

---

LMO	local molecular orbital
MP $n$	$n$ th order Møller-Plesset perturbation theory
MPPT	Møller-Plesset perturbation theory
NCI	noncovalent interaction
NLMO	natural localized molecular orbital
NPA	natural population analysis
OSV	orbital specific virutal
PAO	projected atomic orbital
PM	Pipek-Mezey
PNO	pair natural orbital
PP	pseudopotential
RHF	restricted HF
RRHO	rigid rotor harmonic oscilator
SAPT	Symmetry-Aapted Perturbation Theory
SCS	spin-component scaling
Voxel	volumetric pixel

# Bibliography

- [1] W. B. Schneider, G. Bistoni, M. Sparta, M. Saitow, C. Riplinger, A. A. Auer, F. Neese, *J. Chem. Theory Comput.* **2016**, *12*, 4778–4792.
- [2] L. Pauling, *J. Am. Chem. Soc.* **1932**, *54*, 3570–3582.
- [3] L. Pauling, *J. Am. Chem. Soc.* **1929**, *51*, 1010–1026.
- [4] P. Drude, *Ann. Phys.* **1900**, *306*, 566–613.
- [5] L. Pauling, *The Nature of the Chemical Bond - An Introduction to Modern Structural Chemistry*, 3rd ed., Cornell university press Ithaca, NY, **1960**.
- [6] G. A. Jeffrey, *An introduction to hydrogen bonding*, 1st ed., Oxford university press New York, **1997**.
- [7] R. S. Mulliken, *J. Chem. Phys.* **1935**, *3*, 573–585.
- [8] A. E. Reed, L. A. Curtiss, F. Weinhold, *Chem. Rev.* **1988**, *88*, 899–926.
- [9] J. H. Williams, *Acc. Chem. Res* **1993**, *26*, 593–598.
- [10] K. C. Janda, J. C. Hemminger, J. S. Winn, S. E. Novick, S. J. Harris, W. Klemperer, *J. Chem. Phys.* **1975**, *63*, 1419–1421.
- [11] P. Hobza, H. L. Selzle, E. W. Schlag, *J. Am. Chem. Soc.* **1994**, *116*, 3500–3506.
- [12] D. A. Dougherty, *Science* **1996**, *271*, 163–168.
- [13] J. N. Israelachvili, *Intermolecular and surface forces*, 3rd ed., Academic press, **2011**.
- [14] D. R. Hartree, *Math. Proc. Camb. Philos. Soc.* **1928**, *24*, 89–110.
- [15] V. Fock, *Z. Phys.* **1930**, *61*, 126–148.
- [16] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.
- [17] J. Čížek, *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- [18] J. Čížek, *Adv. Chem. Phys.* **1969**, 35–89.

- [19] A. Szabo, N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*, 1st ed., Dover Publications, **1996**.
- [20] R. Ahlrichs, R. Penco, G. Scoles, *Chem. Phys.* **1977**, *19*, 119–130.
- [21] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864.
- [22] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.
- [23] J. P. Perdew, W. Yue, *Phys. Rev. B* **1986**, *33*, 8800.
- [24] S. Kristyán, P. Pulay, *Chem. Phys. Lett.* **1994**, *229*, 175–180.
- [25] L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, *J. Chem. Phys.* **2000**, *112*, 7374–7383.
- [26] T. M. Gilbert, *J. Phys. Chem. A* **2004**, *108*, 2550–2554.
- [27] S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- [28] E. Caldeweyher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2017**, *147*, 034112.
- [29] B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **1994**, *94*, 1887–1930.
- [30] T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, C. D. Sherrill, *J. Chem. Phys.* **2014**, *140*, 094106.
- [31] R. M. Parrish, C. D. Sherrill, *J. Chem. Phys.* **2014**, *141*, 044115.
- [32] R. M. Parrish, T. M. Parker, C. D. Sherrill, *J. Chem. Theory Comput.* **2014**, *10*, 4417–4431.
- [33] R. M. Parrish, J. F. Gonthier, C. Corminbœuf, C. D. Sherrill, *J. Chem. Phys.* **2015**, *143*, 051103.
- [34] S. Saebø, P. Pulay, *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.
- [35] W. Meyer, *Int. J. Quantum Chem.* **1971**, *5*, 341–348.
- [36] W. Meyer, *J. Chem. Phys.* **1973**, *58*, 1017–1035.
- [37] F. Neese, A. Hansen, D. G. Liakos, *J. Chem. Phys.* **2009**, *131*, 064103.
- [38] F. Neese, F. Wennmohs, A. Hansen, *J. Chem. Phys.* **2009**, *130*, 114108.
- [39] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 034106.
- [40] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
- [41] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2017**, *146*, 164105.

- [42] H.-J. Werner, G. Knizia, C. Krause, M. Schwilk, M. Dornbach, *J. Chem. Theory Comput.* **2015**, *11*, 484–507.
- [43] Q. Ma, H.-J. Werner, *J. Chem. Theory Comput.* **2015**, *11*, 5291–5304.
- [44] M. Schwilk, Q. Ma, C. Köppl, H.-J. Werner, *J. Chem. Theory Comput.* **2017**, *13*, 3650–3675.
- [45] Q. Ma, M. Schwilk, C. Köppl, H.-J. Werner, *J. Chem. Theory Comput.* **2017**, *13*, 4871–4896.
- [46] Q. Ma, H.-J. Werner, *J. Chem. Theory Comput.* **2017**, *14*, 198–215.
- [47] M. Schuetz, G. Rauhut, H.-J. Werner, *J. Phys. Chem. A* **1998**, *102*, 5997–6003.
- [48] L. Magnko, M. Schweizer, G. Rauhut, M. Schütz, H. Stoll, H.-J. Werner, *Phys. Chem. Chem. Phys.* **2002**, *4*, 1006–1013.
- [49] S. Grimme, C. Mück-Lichtenfeld, J. Antony, *Phys. Chem. Chem. Phys.* **2008**, *10*, 3327–3334.
- [50] N. Runeberg, M. Schütz, H.-J. Werner, *J. Chem. Phys.* **1999**, *110*, 7210–7215.
- [51] M. Andrejć, R. A. Mata, *Phys. Chem. Chem. Phys.* **2013**, *15*, 18115–18122.
- [52] J. A. Platts, J Grant Hill, *Mol. Phys.* **2010**, *108*, 1497–1504.
- [53] E. J. Fernández, A. Laguna, J. M. López-de Luzuriaga, M. Monge, P. Pyykkö, N. Runeberg, *Eur. J. Inorg. Chem.* **2002**, *2002*, 750–753.
- [54] A. Wuttke, M. Feldt, R. A. Mata, *J. Phys. Chem. A* **2018**, *122*, 6918–6925.
- [55] A. Wuttke, R. A. Mata, *J. Comput. Chem.* **2017**, *38*, 15–23.
- [56] T. H. Meyer, W. Liu, M. Feldt, A. Wuttke, R. A. Mata, L. Ackermann, *Chem. Eur. J.* **2017**, *23*, 5443–5447.
- [57] T. P. M. Goumans, A. W. Ehlers, K. Lammertsma, E.-U. Würthwein, S. Grimme, *Chem. Eur. J.* **2004**, *10*, 6468–6475.
- [58] R. M. Parrish, C. D. Sherrill, *J. Am. Chem. Soc.* **2014**, *136*, 17386–17389.
- [59] S. Löffler, J. Lübben, A. Wuttke, R. A. Mata, M. John, B. Dittrich, G. H. Clever, *Chem. Sci.* **2016**.
- [60] M. Krick, J. J. Holstein, A. Wuttke, R. A. Mata, G. H. Clever, *Eur. J. Org. Chem.* **2017**, *2017*, 5141–5146.
- [61] S. Löffler, A. Wuttke, B. Zhang, J. J. Holstein, R. A. Mata, G. H. Clever, *ChemComm* **2017**, *53*, 11933–11936.

- [62] H. M. J. Wang, C. Y. L. Chen, I. J. B. Lin, *Organometallics* **1999**, *18*, 1216–1223.
- [63] J. M. López-de Luzuriaga, M. Monge, M. E. Olmos, D. Pascual, T. Lasanta, *Chem. Commun.* **2011**, *47*, 6795–6797.
- [64] T. Lasanta, J. M. López-de Luzuriaga, M. Monge, M. E. Olmos, D. Pascual, *Dalton Trans.* **2016**, *45*, 6334–6338.
- [65] P. Jurečka, J. Šponer, J. Černý, P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
- [66] L. Hua, R. Zhou, D Thirumalai, B. Berne, *Proc. Natl. Acad. Sci.* **2008**, *105*, 16928–16933.
- [67] S. Grimme, *Chem. Eur. J.* **2012**, *18*, 9955–9964.
- [68] L. Ackermann, S. Barfüßer, C. Kornhaaß, A. R. Kapdi, *Org. Lett.* **2011**, *13*, 3082–2085.
- [69] P. R. Schreiner, L. V. Chernish, P. A. Gunchenko, E. Y. Tikhonchuk, H. Hausmann, M. Serafin, S. Schlecht, J. E. P. Dahl, R. M. K. Carlson, A. A. Fokin, *Nature* **2011**, *477*, 308–311.
- [70] H. C. Gottschalk, A. Poblitzki, M. A. Suhm, M. M. Al-Mogren, J. Antony, A. A. Auer, L. Baptista, D. M. Benoit, G. Bistoni, F. Bohle, et al., *J. Chem. Phys.* **2018**, *148*, 014301.
- [71] D. Bernhard, F. Dietrich, M. Fatima, C. Pérez, H. C. Gottschalk, A. Wuttke, R. A. Mata, M. A. Suhm, M. Schnell, M. Gerhards, *Beilstein J. Org. Chem.* **2018**, *14*, 1642–1654.
- [72] E. Schrödinger, *Ann. Phys.* **1926**, *385*, 437–490.
- [73] M. Born, R. Oppenheimer, *Ann. Phys.* **1927**, *389*, 457–484.
- [74] J. C. Slater, *Phys. Rev.* **1929**, *34*, 1293.
- [75] W. Pauli, *Z. Phys.* **1925**, *31*, 765–783.
- [76] F. Jensen, *Introduction to computational chemistry*, 2nd ed., John Wiley & Sons, **2007**.
- [77] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69.
- [78] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- [79] C. Hampel, H.-J. Werner, *J. Chem. Phys.* **1996**, *104*, 6286–6297.
- [80] P. Pulay, S. Saebø, *Theor. Chim. Acta* **1986**, *69*, 357–368.



- [81] H.-J. Werner, C. Köppl, Q. Ma, M. Schwilk in *Fragmentation - Toward Accurate Calculations on Complex Molecular Systems*, (Ed.: M. S. Gordon), John Wiley & Sons, **2017**, pp. 1–61.
- [82] S. F. Boys in *Quantum Theory of Atoms, Molecules, and the Solid State*, (Ed.: P. O. Löwdin), Academic Press, New York, **1966**, pp. 253–262.
- [83] J. Pipek, P. G. Mezey, *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- [84] J. W. Boughton, P. Pulay, *J. Comput. Chem.* **1993**, *14*, 736–740.
- [85] H.-J. Werner, K. Pflüger, *Ann. Reports in Comput. Chem.* **2006**, *2*, 53.
- [86] G. Knizia, *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- [87] W. Meyer, *J. Chem. Phys.* **1976**, *64*, 2901.
- [88] J. Yang, Y. Kurashige, F. R. Manby, G. K. L. Chan, *J. Chem. Phys.* **2011**, *134*, 044123.
- [89] J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, H.-J. Werner, *J. Chem. Phys.* **2012**, *136*, 144105.
- [90] M. Schütz, J. Yang, G. K.-L. Chan, F. R. Manby, H.-J. Werner, *J. Chem. Phys.* **2013**, *138*, 054109.
- [91] R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 1833–1840.
- [92] P.-O. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365–375.
- [93] A. E. Reed, R. B. Weinstock, F. Weinhold, *J. Chem. Phys.* **1985**, *83*, 735.
- [94] R. A. Mata, H.-J. Werner, *Mol. Phys.* **2007**, *105*, 2753–2761.
- [95] K. Kitaura, K. Morokuma, *Int. J. Quantum Chem.* **1976**, *10*, 325–340.
- [96] K. Morokuma, *J. Chem. Phys.* **1971**, *55*, 1236–1244.
- [97] E. D. Glendening, A. Streitwieser, *J. Chem. Phys.* **1994**, *100*, 2900–2909.
- [98] R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, M. Head-Gordon, *J. Phys. Chem. A* **2007**, *111*, 8753–8765.
- [99] W. Chen, M. S. Gordon, *J. Phys. Chem.* **1996**, *100*, 14316–14328.
- [100] P. Su, H. Li, *J. Chem. Phys.* **2009**, *131*, 014102.
- [101] K. Szalewicz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 254–272.
- [102] A. Heßelmann, G. Jansen, *Chem. Phys. Lett.* **2002**, *357*, 464–470.
- [103] A. Heßelmann, G. Jansen, *Chem. Phys. Lett.* **2002**, *362*, 319–325.
- [104] N. O. B. Lüttschwager, T. N. Wassermann, R. A. Mata, M. A. Suhm, *Angew. Chem. Int. Ed.* **2013**, *52*, 463–466.

- [105] S. Grimme, R. Huenerbein, S. Ehrlich, *ChemPhysChem* **2011**, *12*, 1258–1261.
- [106] S. Grimme, P. R. Schreiner, *Angew. Chem. Int. Ed.* **2011**, *50*, 12639–12642.
- [107] I. Mayer, *Int. J. Quantum Chem.* **1983**, *23*, 341–363.
- [108] E. Pastorczak, A. Prlj, J. F. Gonthier, C. Corminboeuf, *J. Chem. Phys.* **2015**, *143*, 224107.
- [109] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
- [110] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, 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, M. Wang, MOLPRO, version 2015.1, a package of ab initio programs, see <http://www.molpro.net>, Cardiff, UK, **2015**.
- [111] Schrödinger, LLC, The PyMOL Molecular Graphics System, Version 1.6, **2011**.
- [112] J. Ahrens, B. Geveci, C. Law, C Hansen, C Johnson, *Paraview - The visualization handbook* **2005**, *1*.
- [113] 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, M. Wang, MOLPRO, version 2012.1, a package of ab initio programs, see <http://www.molpro.net>, Cardiff, UK, **2012**.
- [114] H.-J. Werner, F. R. Manby, P. Knowles, *J. Chem. Phys.* **2003**, *118*, 8149–8160.
- [115] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- [116] M. S. Sinnokrot, C. D. Sherrill, *J. Phys. Chem. A* **2004**, *108*, 10200–10207.

- [117] J. G. Hill, J. A. Platts, H.-J. Werner, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4072.
- [118] T. Janowski, P. Pulay, *Chem. Phys. Lett.* **2007**, *447*, 27–32.
- [119] E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, K. S. Kim, *J. Phys. Chem. A* **2007**, *111*, 3446–3457.
- [120] O. Bludský, M. Rubeš, P. Soldán, P. Nacthigall, *J. Chem. Phys.* **2008**, *128*, 114102.
- [121] T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [122] R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.
- [123] F. Weigend, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285.
- [124] F. Weigend, A Köhn, C. Hättig, *J. Chem. Phys.* **2002**, *116*, 3175.
- [125] R. A. Mata, H.-J. Werner, *J. Chem. Phys.* **2006**, *125*, 184110.
- [126] A. E. Reed, F. Weinhold, *J. Chem. Phys.* **1985**, *83*, 1736.
- [127] C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.
- [128] S. E. Wheeler, K. N. Houk, *J. Am. Chem. Soc.* **2008**, *130*, 10854–10855.
- [129] H. C. Gottschalk, J. Altnöder, M. Heger, M. A. Suhm, *Angew. Chem. Int. Ed.* **2015**, *55*, 1921–1924.
- [130] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [131] L. C. Palmer, J. Rebek Jr, *Org. Biomol. Chem.* **2004**, *2*, 3051–3059.
- [132] M. D. Pluth, K. N. Raymond, *Chem. Soc. Rev.* **2007**, *36*, 161–171.
- [133] P. W. Snyder, M. R. Lockett, D. T. Moustakas, G. M. Whitesides, *Eur. Phys. J. ST* **2014**, *223*, 853–891.
- [134] J. W. Steed, J. L. Atwood, *Supramolecular chemistry*, 2nd ed., John Wiley & Sons, **2009**.
- [135] H. Amouri, C. Desmarets, J. Moussa, *Chem. Rev.* **2012**, *112*, 2015–2041.
- [136] P. D. Frischmann, M. J. MacLachlan, *Chem. Soc. Rev.* **2013**, *42*, 871–890.
- [137] S. Kubik in *Chemistry of Nanocontainers*, Springer, **2011**, pp. 1–34.
- [138] R. Chakrabarty, P. S. Mukherjee, P. J. Stang, *Chem. Rev.* **2011**, *111*, 6810–6918.

- [139] M. M. Smulders, I. A. Riddell, C. Browne, J. R. Nitschke, *Chem. Soc. Rev.* **2013**, *42*, 1728–1754.
- [140] M. Han, D. M. Engelhard, G. H. Clever, *Chem. Soc. Rev.* **2014**, *43*, 1848–1860.
- [141] A. J. McConnell, C. S. Wood, P. P. Neelakandan, J. R. Nitschke, *Chem. Rev.* **2015**, *115*, 7729–7793.
- [142] S. Freye, R. Michel, D. Stalke, M. Pawliczek, H. Frauendorf, G. H. Clever, *J. Am. Chem. Soc.* **2013**, *135*, 8476–8479.
- [143] M. Han, R. Michel, B. He, Y.-S. Chen, D. Stalke, M. John, G. H. Clever, *Angew. Chem. Int. Ed.* **2013**, *52*, 1319–1323.
- [144] C. J. Hastings, M. D. Pluth, R. G. Bergman, K. N. Raymond, *J. Am. Chem. Soc.* **2010**, *132*, 6938–6940.
- [145] T. Murase, Y. Nishijima, M. Fujita, *J. Am. Chem. Soc.* **2011**, *134*, 162–164.
- [146] W. Cullen, M. C. Misuraca, C. A. Hunter, N. H. Williams, M. D. Ward, *Nat. Chem.* **2016**, *8*, 231.
- [147] P. Howlader, P. Das, E. Zangrando, P. S. Mukherjee, *J. Am. Chem. Soc.* **2016**, *138*, 1668–1676.
- [148] M. Frank, J. Ahrens, I. Bejenke, M. Krick, D. Schwarzer, G. H. Clever, *J. Am. Chem. Soc.* **2016**, *138*, 8279–8287.
- [149] R. Custelcean, *Chem. Soc. Rev.* **2014**, *43*, 1813–1824.
- [150] C. Sgarlata, J. S. Mugridge, M. D. Pluth, B. E. Tiedemann, V. Zito, G. Arena, K. N. Raymond, *J. Am. Chem. Soc.* **2009**, *132*, 1005–1009.
- [151] M. Yoshizawa, J. Nakagawa, K. Kumazawa, M. Nagao, M. Kawano, T. Ozeki, M. Fujita, *Angew. Chem. Int. Ed.* **2005**, *44*, 1810–1813.
- [152] T. K. Ronson, C. Giri, N. Kodiah Beyeh, A. Minkinen, F. Topić, J. J. Holstein, K. Rissanen, J. R. Nitschke, *Chem. Eur. J.* **2013**, *19*, 3374–3382.
- [153] G. H. Clever, S. Tashiro, M. Shionoya, *Angew. Chem. Int. Ed.* **2009**, *48*, 7010–7012.
- [154] G. H. Clever, W. Kawamura, M. Shionoya, *Inorg. Chem.* **2011**, *50*, 4689–4691.
- [155] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [156] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [157] E. R. Johnson, A. D. Becke, *J. Chem. Phys.* **2006**, *124*, 174104.

- [158] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [159] F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- [160] M. Frank, M. D. Johnstone, G. H. Clever, *Chem. Eur. J.* **2016**, *22*, 14104–14125.
- [161] S. Löffler, J. Lübben, L. Krause, D. Stalke, B. Dittrich, G. H. Clever, *J. Am. Chem. Soc.* **2015**, *137*, 1060–1063.
- [162] S. Freye, J. Hey, A. Torras-Galán, D. Stalke, R. Herbst-Irmer, M. John, G. H. Clever, *Angew. Chem. Int. Ed.* **2012**, *51*, 2191–2194.
- [163] J. M. Dieterich, G. H. Clever, R. A. Mata, *Phys. Chem. Chem. Phys.* **2012**, *14*, 12746–12749.
- [164] W. Hehre, J. Yu, P. Klunzinger, L. Lou, et al., *Inc. Irvine* **2000**.
- [165] G. J. Kleywegt, T. A. Jones, *Acta Cryst. Section D: Biological Crystallography* **1994**, *50*, 178–185.
- [166] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.
- [167] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- [168] D. Andrae, U. Haeussermann, M. Dolg, H. Stoll, H. Preuss, *Theor. Chim. Acta* **1990**, *77*, 123.
- [169] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 91–100.
- [170] K. A. Peterson, D. Figgen, M. Dolg, H. Stoll, *J. Chem. Phys.* **2007**, *126*, 124101.
- [171] F. Weigend, *J. Comput. Chem.* **2008**, *29*, 167.
- [172] R. Gulde, P. Pollak, F. Weigend, *J. Chem. Theory Comput.* **2012**, *8*, 4062–4068.
- [173] C. Hättig, *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.
- [174] G. Hetzer, P. Pulay, H.-J. Werner, *Chem. Phys. Lett.* **1998**, *290*, 143–149.
- [175] F. Eckert, A. Klamt, *AIChE Journal* **2002**, *48*, 369–385.
- [176] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143–152.
- [177] D. Ardura, R. López, T. L. Sordo, *J. Phys. Chem. B* **2005**, *109*, 23618–23623.
- [178] R. Sure, S. Grimme, *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801.
- [179] R. Sure, S. Grimme, *J. Comput. Chem.* **2013**, *34*, 1672–1685.

- [180] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [181] C. Köppl, H.-J. Werner, *J. Chem. Theory Comput.* **2016**, *12*, 3122–3134.
- [182] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, 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, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. M. III, 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, M. Wang, M. Welborn, MOLPRO, version 2018.2, a package of ab initio programs, see <http://www.molpro.net>, Cardiff, UK, **2018**.
- [183] S. Miranda, J. Feldt, F. Pratas, R. A. Mata, N. Roma, P. Tomás, *Int. J. High Perform. Comput. Appl.* **2017**, *31*, 499–516.
- [184] A. T. Fenley, N. M. Henriksen, H. S. Muddana, M. K. Gilson, *J. Chem. Theory Comput.* **2014**, *10*, 4069–4078.
- [185] S. Grimme, *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514.
- [186] ChemShell, a Computational Chemistry Shell, see [www.chemshell.org](http://www.chemshell.org).
- [187] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. B. a title=.
- [188] M. Yoshizawa, J. K. Klosterman, M. Fujita, *Angew. Chem. Int. Ed.* **2009**, *48*, 3418–3438.
- [189] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, G. K.-L. Chan, *Science* **2014**, *345*, 640–643.
- [190] M. Cutini, B. Civalleri, M. Corno, R. Orlando, J. G. Brandenburg, L. Maschio, P. Ugliengo, *J. Chem. Theory and Comput.* **2016**, *12*, PMID: 27304925, 3340–3352.
- [191] J. Antony, S. Grimme, *Phys. Chem. Chem. Phys.* **2008**, *10*, 2722–2729.
- [192] E. O’Grady, N. Kaltsoyannis, *Phys. Chem. Chem. Phys.* **2004**, *6*, 680–687.
- [193] L. Ray, M. M. Shaikh, P. Ghosh, *Inorg. Chem.* **2008**, *47*, 230–240.
- [194] P. Pulay, *Chem. Phys. Lett.* **1983**, *100*, 151–154.
- [195] S. Saebø, P. Pulay, *Chem. Phys. Lett.* **1985**, *113*, 13–18.

- [196] J. Antony, S. Grimme, *J. Phys. Chem. A* **2007**, *111*, 4862–4868.
- [197] D. Figgen, G. Rauhut, M. Dolg, H. Stoll, *Chem. Phys.* **2005**, *311*, 227.
- [198] K. A. Peterson, C. Puzzarini, *Theor. Chem. Acc.* **2005**, *114*, 283.
- [199] M. Dolg, H. Stoll, A. Savin, H. Preuss, *Theor. Chim. Acta* **1989**, *75*, 173.
- [200] F. Weigend, A. Baldes, *J. Chem. Phys.* **2010**, *133*, 174102.
- [201] B. Assadollahzadeh, P. Schwerdtfeger, *Chem. Phys. Lett.* **2008**, *462*, 222–228.
- [202] J. Muñiz, C. Wang, P. Pyykkö, *Chem. Eur. J.* **2011**, *17*, 368–377.
- [203] P. Pyykkö, *Chem. Soc. Rev.* **2008**, *37*, 1967–1997.
- [204] R.-F. Liu, C. A. Franzese, R. Malek, P. S. Zuschowski, J. G. Angyan, M. M. Szczesniak, G. Chalasinski, *J. Chem. Theory Comput.* **2011**, *7*, 2399–2407.
- [205] P Zhou, J Huang, F Tian, *Curr. Med. Chem.* **2012**, *19*, 226–238.
- [206] S. E. Wheeler, J. W. Bloom, *J. Phys. Chem. A* **2014**, *118*, 6133–6147.
- [207] B. R. Beno, K.-S. Yeung, M. D. Bartberger, L. D. Pennington, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 4383–4438.
- [208] P. Pyykkö, *Angew. Chem. Int. Ed.* **2004**, *43*, 4412–4456.
- [209] P. Pyykkö, *Inorg. Chim. Acta* **2005**, *358*, 4113–4130.
- [210] H. Schmidbaur, A. Schier, *Chem. Soc. Rev.* **2012**, *41*, 370–412.