Charles University in Prague
Faculty of Science

Ph. D. study program: Modelling of Chemical Properties of Nano- and Biostructures

RNDr. Róbert Sedlák

Quantum-chemical study of noncovalent interactions

Dissertation Thesis

Supervisor:
Prof. Pavel Hobza, DrSc., FRSC, Dr. h. c.

Academy of Sciences of the Czech Republic
Institute of Organic Chemistry and Biochemistry, v.v.i.
Center of Excellence: "Controlling structure and function of biomolecules at the molecular level: theory meets experiment"

Prague, 2014

Univerzita Karlova v Praze
Přírodovědecká fakulta

Doktorský studijní program: Modelování chemických vlastností nano- a biostruktur

RNDr. Róbert Sedlák

Kvantově-chemická studie nekovalentních interakcí

Disertační práce

Školitel:
prof. Ing. Pavel Hobza, DrSc., FRSC, Dr. h. c.

Akademie věd, Česká republika
Ústav organické chemie a biochemie
Centrum Excelence: "Řizení struktury a funkce biomolekul na molekulové úrovni: souhra teorie a experimentu"

Praha, 2014

Declaration of Authorship

I, RÓBERT SEDLÁK, declare that this dissertation titled, "*Ab initio* QUANTUM CHEMICAL STUDY of NONCOVALENT INTERACTIONS" contains solely my original work. All the literature is properly cited, and I have not been yet awarded any other academic degree or diploma for this thesis or its substantial part.

Prague, September 2014 Róbert Sedlák

Acknowledgement

First, I would like to thank my family. I am also indebted to all my collaborators and members of Prof. Hobza‘s group for their help, cooperation, support and friendly atmosphere. Last but not least I would like to thank Professor Pavel Hobza for his kind guidance and everlasting support.

Contents

List of Tables

List of Figures

List of Abbreviations

HB	**H**ydrogen **B**ond
DHB	**D**i**H**ydrogen **B**ond
MP2	**M**øller-**P**lesset second order perturbation theory
ESP	**E**lectro**S**tatic **P**otential
NPA	**N**atural **P**opulation **A**nalysis
CSD	**C**ambridge **S**tructural **D**atabase
PDB	**P**rotein **D**ata **B**ank
XB	**H**alogen **B**ond
CC	**C**oupled **C**luster theory
CCSD	**C**oupled **C**luster theory with **S**ingle and **D**ouble excitations
CCSD(T)	**C**oupled **C**luster theory with **S**ingle, **D**ouble and perturbative **T**riple excitations
CBS	**C**omplete **B**asis **S**et
HF	**H**artree-**F**ock theory
BSSE	**B**asis **S**et **S**uperposition **E**rror
CPC	**C**ounter**P**oise **C**orrection
QCISD(T)	**Q**uadratic **C**onfiguration **I**nteraction theory with **S**ingle, **D**ouble and perturbative **T**riple excitations
SCS-MP2	**S**pin **C**omponent **S**caled **MP2**
SCS(MI)-MP2	**S**pin **C**omponent **S**caled **M**olecular **I**nteraction **MP2**
RMSD	**R**oot **M**ean **S**quare **D**eviation
SCS-CCSD	**S**pin **C**omponent **S**caled **CCSD**
SCS(MI)-CCSD	**S**pin **C**omponent **S**caled **M**olecular **I**nteraction **CCSD**
MP2C	**MP2 C**orrected method
MP3	**M**øller-**P**lesset third order perturbation theory
MP2.5	**M**øller-**P**lesset second order perturbation theory with the inclusion of third order correction
DFT	**D**ensity **F**unctional **T**heory

DFT-D	**DFT** with empirical **D**ispersion correction
SAPT	**S**ymmetry **A**dapted **P**erturbation **T**heory
DFT-SAPT	**SAPT** in which intramolecular correlation is treated fully by **DFT**
SQM	**S**emiempirical **Q**uantum **M**echanical theory
NPA	**N**atural **P**opulation **A**nalysis
NBO	**N**atural **B**ond **O**rbital
NAO	**N**atural **A**tomic **O**rbital
DMA	**D**istributed **M**ultipole **A**nalysis
CT	**C**harge-**T**ransfer
HOMO	**H**ighest **O**ccupied **M**olecular **O**rbital
LUMO	**L**owest **U**noccupied **M**olecular **O**rbital

Abstract

The aim of this thesis is to investigate strength and origin of the stabilization for various types of noncovalent interactions. As this knowledge could lead to a deeper understanding and rationalization of the binding phenomena. Further, to participate on the development of new noncovalent data sets, which are nowadays inevitable in the process of parametrization and validation of new computational methods.

In all the studies, different binding motifs of model complexes, which represent usually crystal structures, structures from unrelaxed scans or the local minima, were investigated. The calculations of the reference stabilization energies were carried out at *ab initio* level (e.g. CCSD(T)/CBS, QCISD(T)/-CBS). Further, the accuracy of more approximate methods (e.g. MP2.5, DFT-D or SQM methods) toward the reference method, was tested. In order to obtain the nature of the stabilization the DFT-SAPT decomposition was frequently utilized.

In the first part of the thesis, the importance and basic characteristics of different types of noncovalent interactions (e.g. halogen bond, hydrogen bond, $\pi\cdots\pi$ interaction *etc.*), are discussed. The second part provides the description of computational methods which were essential for our investigation. The third part of the thesis provides an overview for part of our research during my PhD studies, relevant to the topic of the thesis.

Abstrakt

Cílem této práce je určit sílu a původ stabilizace různých typů nekovalentních interakcí, a to z důvodu, že by jejich znalost mohla vést k racionalizaci a hlubšímu pochopení vazebních motivů. Dalším cílem je podílet se na vývoji nových nekovalentních datasetů, které jsou v současné době, v procesu parametrizace a testování nových výpočetních metod, nezastupitelné.
V rámci jednotlivých projektů byly zkoumány různé vazební motivy modelových komplexů, které představují obvykle krystalové struktury, struktury z neoptimalizovaných skenů nebo lokální minima. Referenční stabilizační energie byly počítány na *ab initio* úrovni (např. CCSD (T)/CBS nebo QCISD(T)/CBS). Následně byla testovaná přesnost výpočetně méně náročných metod (např. MP2.5, DFT-D nebo SQM metod) vůči referenční metodě. DFT-SAPT výpočty byly často využívány za účelem zjištění povahy stabilizace komplexů.
První část práce se zabývá významem a základními vlastnostmi různých typů nekovalentních interakci (např. halogenové vazby, vodíkové vazby atd.). Druhá část popisuje výpočetní metody, které byly důležité v rámci našich studií. Poslední část této práce má za účel poskytnout přehled o části našeho výzkumu, která souvisí s tématem této disertační práce a na niž jsem se podílel během mého doktorského studia.

Chapter 1

Introduction

1.1 Noncovalent Interactions

Noncovalent interactions have been the subject of intensive research during last couple of decades. The proper knowledge of these interactions is a prerequisite for designing wide range of materials as well as understanding many relevant biological phenomena. Noncovalent interactions, such as hydrogen bond, halogen bond, $\pi\cdots\pi$ stacking, etc., play an important role in processes like the molecular recognition, crystal packing, protein folding, stacking of nucleobases, etc.[1–6] Although these interactions are at least by an order of magnitude weaker than covalent interactions, their impact on structure and function of biomolecules and molecular crystals is fundamental. Noncovalent interactions can be divided, according to strength, into strong, medium and weak,[7] although the demarcation zones between these groups are blurred.[8] However, what applies in general is the following: The more specific and directional the interaction is, the stronger it is.
In the next subsections several fundamental types of noncovalent interactions are briefly described with special emphasis on halogen bonding.

1.1.1 Hydrogen Bond

Hydrogen bond (HB) is a classical example of strong, directional and specific noncovalent interaction. It plays an important role in structure determination of polymers and biomacromolecules. The HB is an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule (i.e. intramolecular HB or intermolecular HB, respectively), in which there is evidence of bond formation.[9] The forces responsible for stabilization of HB mainly include those of electrostatic origin. However, charge-transfer (CT) as well as dispersion forces also participate in the stabilization. The typical hydrogen bond is schematically depicted as X–H$\cdots$Y, where the dots represent the bond. X–H represents the hydrogen bond donor. The hydrogen bond acceptor (Y) may be an atom, an anion Y, or a molecule. In any case, the acceptor Y is an electron rich region such as, a lone electron pair or π electron density.

The formation of hydrogen bond is usually connected with several characteristic features. The length of the X–H bond usually increases on hydrogen bond formation leading to a red shift in the infrared X–H stretching frequency and an increase in the infrared absorption cross-section for the X–H stretching vibration. The greater the lengthening of the X–H bond in X–H$\cdots$Y, the stronger is the H-bond.[9] There are, however, certain hydrogen bonds in which the X–H bond length decreases, the so-called blue-shifting or improper H-bond.[10,11]

The strength of HB varies significantly from 1 kcal.mol^{-1} (e.g. C–H$\cdots\pi$ 1 - 1.5 kcal.mol^{-1}),[7] through few kcal.mol^{-1} (e.g. N–H$\cdots$O, O–H$\cdots$O 5 - 7 kcal.mol^{-1})[12,13] up to over 10 kcal.mol^{-1} (e.g. O–H$\cdots$O^{-}, F–H$\cdots$F^{-} 23 - 39 kcal.mol^{-1}).[7] There are few important factors which strongly influence the strength of the X–H$\cdots$Y hydrogen bond.

- The strength of the H$\cdots$Y bond correlates positively with the increase

of the proton donation ability of the atom X, i.e. electronegativity (e.g. proton donation ability of carbon atom depends on its hybridization state: $C(sp^3) < C(sp^2) < C(sp)$).[3]

- The arrangements of three atoms X–H$\cdots$Y is usually linear (180°), i.e. the more linear it is, the stronger is the HB.

In conventional HB, atoms directly bonded to H (X and Y) are typically highly electronegative (e.g. O, N, F). In the case of non-conventional HB (e.g. C–H$\cdots$O, N, F)[7,14–16] the situation is different. The non-conventional HBs are also important, especially in the absence of conventional HBs, despite the fact they are considerably weaker than conventional HB. This, at first glance surprising fact, can be understood as a consequence of high abundance of this interaction in (bio)macromolecules. The importance of the C–H$\cdots$O interaction in crystal engineering[17,18] or biological systems was already addressed.[19–23]

The N,O,C–H$\cdots\pi$ interactions represent another subset of HB, which can be associated with non-conventional HBs; so called π-HBs. The hydrogen atoms point toward the electron rich region and the relatively short distance (smaller than sum of vdW radii of H and $C(sp^2)$) is typically observed. This type of HBs is weaker than conventional HBs. Nevertheless, similarly to previously discussed interactions (C–H$\cdots$O, N, F), these also play crucial role in formation of molecular assemblies,[24] stability of the peptides[25] and proteins[26,27] and conformation of the small organic molecules.[28]

Comprehensive structural study done by Steiner and Koellner[29] revealed that about 1 out of 11 aromatic amino acids act as π acceptors for H bonding with the N, O, S–H donors. This relatively high frequency, although small compared to classical hydrogen bonds, supports the postulated role which these interactions play in stabilizing the secondary structure of proteins.[30] The often observed contraction of the N, O, C–H bond is opposite to what

is observed for classical HBs. This contraction leads to a shift of the bond stretching movement to higher frequencies in the infra-red spectrum; so called blue-shifting HBs.[10,11]

1.1.2 $\pi\cdots\pi$ Interactions

The $\pi\cdots\pi$ interactions are important in both the material sciences and biochemistry. They are abundant in proteins,[4,5,31] significantly influence protein conformation[4,5,32] and thermal stability of the proteins[33] and nucleic acids.[4,34] The majority of medicinal agents contain aromatic substituents and their recognition by proteins is likely dominated by $\pi\cdots\pi$ interactions.[35] They are crucially involved in protein-deoxynucleic acid complexes where interactions between aromatic residues and base pairs are seen in X-ray crystal structures.[36,37] The macroscopic properties of carbon-based materials are largely determined by $\pi\cdots\pi$ interaction. Graphene sheets, fullerenes, carbon nanotubes have already many applications and are still intensely studied.[38–40]

Because $\pi\cdots\pi$ interactions are prevalent across chemistry, a large body of experimental and theoretical work has focused on determining the gas phase structure of the prototype, benzene dimer.[41–48] The experimentally observed structure depends heavily upon the observation technique.[49] Parallel displaced (PD), and T-shaped (TS), structures are two most relevant conformers on the potential energy surface (see Figure 1.1).

Large scale ab initio electronic structure theory suggests that the PD and TS structures are nearly isoenergetic.[41–48] As the intermolecular potential surface is quite soft one must view the dimer as a dynamic system rather than one with a well defined structure. Survey by McGaughey *et al.*[34] on a nonhomologous set of proteins corrected for probability distribution bias and including only isolated dimer pairs, revealed that the aromatic side chain amino acids do have a preferred intermolecular structure. The preferred PD orientation is found to be more stable than a TS structure by 0.5-0.75 kcal.mol^{-1} for Phe-Phe dimers and by 1.0 kcal.mol^{-1} for the full set of dimers.[34]

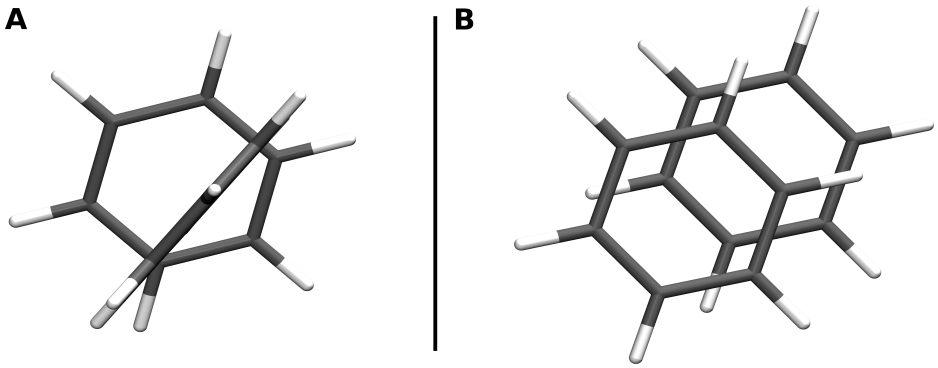

Figure 1.1: T-shaped (A) and parallel displaced (B) structure of benzene dimer (top view).

Grimme's study concernig specificity of the $\pi\cdots\pi$ stacking interaction revealed several interesting features: (i) the stacking interaction become significant between molecules with more than 10-15 carbon atoms, as the interactions between saturated fragments of smaller size is similarly strong;[50] (ii) the geometrical arrangement of the fragments determines the character of the interaction. The aromatic molecules in a TS orientation show a very similar decrease of the interaction energy with system size as the saturated molecules. Also the dependence of the intermolecular distances on system size is similar for TS and for stacked saturated dimers, but is different in the case of π-stacked dimers, for which a substantial shortening is observed for larger molecules;[50] (iii) π-stacking is dominated by the more favorable dispersion component compare with TS and stacking of saturated species. The electrostatic term favors stacking for saturated rather than π complexes. This effect is in the π-stacking arrangements minimized by parallel displacement;[50] (iv) Special nonlocal electron correlations between the π electrons in the two fragments at small interplane distances are responsible for the π-stacking.[50] In the aromatic T-shaped complexes, such electron correlations are insignificant because of the larger average distances. The special shape of the π orbitals allows a closer contact of the monomers (i.e. maximizing the attractive dispersion component) without too much increasing the Pauli exchange repulsion, consequently tighter binding leads to cooperative π effects. For geometrical reasons, this is not possible in the analogous

(stacking) arrangement of saturated molecules (or in the TS structures of aromatic molecules), and thus for these molecules interaction components are additive.

The quantitative description of $\pi\cdots\pi$ stacking intermolecular interaction is very demanding. As, the stabilization forces of these interactions are predominantly formed by dispersion, they can not be accurately described at the Hartree-Fock level. Further, the local or semilocal character of density functionals is able to reproduce only short-range and middle-range correlation effects.[51] The Møller-Plesset method (MP2) is the simplest method able to describe long-range correlation effects from the right reason. However, it is well known that MP2 overestimates the $\pi\cdots\pi$ stacking interaction.[52–55] Second, the dispersion is weak effect and at van der Waals distances is significantly compensated by Pauli repulsion, what can lead to error amplification. Moreover, for small and middle size basis set superposition error can be as big as dispersion. Therefore, use of more elaborate *ab initio* methods in combination with sufficiently large basis set is needed in order to obtain directly the correct description of this type of interaction.

1.1.3 Dihydrogen Bond

The crystal structure of hypophosphite $NH_4^+(H_2PO_2)^-$, published in 1934 by Zachariansen and Mooney,[56] was the first structure where rather strange behavior of hydrogen atoms was recognized and reported; hydrogens of the hypophosphite group act as H^- species with respect to ammonium. The N–H$\cdots$H–B interaction was suggested in liquid dimethylaminoborane $(CH_3)_2NH{\cdot}BH_3$, in 1964.[57] However, the H$\cdots$H contacts were firstly recognized as true hydrogen bonds in late 1960s by M. P. Brown and coworkers.[58,59] In another study by Brown *et al.* precise analysis of the infra-red spectra of the boron coordination compounds ($L{\cdot}BH_3$, L = Me_3N, Et_3N, Py, Et_3P and $Me_3N{\cdot}BH_2X$, X = Cl, Br, I) in presence of proton donors (MeOH, PhOH, p-F-C_6H_4OH) in CCl_4 led to conclusions which proposed formation of new type of hydrogen bonds, where BH_3 and BH_2 groups act as proton acceptors.[60] In 1995, Crabtree suggested for short N–H$\cdots$H–B contacts the term

"dihydrogen" bond.[61]
Nowadays, the dihydrogen bond is general, well-established, highly specific, directional, noncovalent interaction between σ-bonding electron pair of M–H bonds (M = B, Al, Ga, Li, Be, Xe, Ir, Mo, Mn, Os, Re, Ru, W) and proton donors X–H (X = F, O, N, C).[62–75]

At this place we list several properties of dihydrogen bonds

- the origin of the stabilization comes from electrostatic and dispersion
- the stabilization energy increases proportionally with proton donor's acidity;[76] it ranges from 1 to 7 kcal.mol^{-1}
- the H$\cdots$H distance varies between 1.7 and 2.2 A; it is significantly shorter than the sum of the van der Waals radii of two hydrogen atoms (2.4 A)[61]
- strongly bent type of conformation is prefered: the XH$\cdots$H–M angles amounts to 90 - 135°; the X–H$\cdots$HM angle amounts to 145 - 175° [76]

There were many experimental and theoretical studies published during late 1990s, which contributed to gradual understanding of dihydrogen bonding, explained conformational preference, high melting points or stability of the small molecules and complexes.[77–84]

Other studies presented the importance of H$\cdots$H interaction in crystal packing processes[63] or supramolecular chemistry.[85,86] Reactivity and selectivity control by dihydrogen bonding in solutions manifested through activation of the metal–H bond upon formation of dihydrogen bond and subsequent formation of η^2–H_2 complexes followed by elimination of molecular hydrogen and formation of new covalent bond demonstrate atypical reactivity of these species caused by unusual nature of this noncovalent interaction.[87–89] Furthermore, combining ability of boron hydrides in solid state to self-assemble into extended dihydrogen-bonded networks together with specific reactivity in solutions, makes them potentially powerful tool for rational assembly of new crystalline covalent materials.[90,91]

Moreover, nowadays the dihydrogen bonding is studied in connection with biomolecular chemistry and medicinal chemistry. The heteroboranes (substituted boran hydrides) interact with biomolecules through dihydrogen bonds.[92] Metallacarboranes, which bind to active site of the protein through dihydrogen bonds, were shown to act as potent and specific inhibitors of human immunodeficiency virus protease[93] and human carbonic anhydrase.[94]

1.1.4 Halogen Bond

1.1.4.1 Introduction to the Halogen Bond

The complexes of halogens (Cl_2, Br_2, I_2) with amines are known since 18^{th} and 19^{th} century.[95,96] Many other complexes of organic halides and N/O atoms as Lewis base moieties were crystallographically characterized in the middle of previous century.[97,98] Crystallographic survey by Murray-Rust[99] revealed that covalently-bonded halogen can interact attractively with nucleophile (*cf.* Figure 1.2 A) as well as electrophile (*cf.* Figure 1.2 B).

Figure 1.2: Schematic illustration of the interaction with the halogen atom X.

The interaction with nucleophile is called "halogen bond" (XB), however this term is known only since the 70's.[100,101] The significant potential hidden in halogen bond was demonstrated in various applications. Imakubo *et al.* and Amico *et al.* showed that halogen bonding can serve as a tool for predictable packing of molecules which can be used for designing the materials with desirable properties.[102,103] These works meant the beginning of new field which is now rapidly growing and has relevance to magnetics, optics and electronics.[104–107] Increased interest in halogen bonding followed after the recognition of its widespread occurrence in biological systems. This was observed after publishing Auffinger survey[108] on crystal structures from Protein Data

Bank. The role of halogen bonding in many biological and medicinal areas is now being investigated; among these are protein-ligand interactions, conformational stability, drug design, docking processes, etc.[104,108–110] Light halogens e.g. fluorine and chlorine are widely used substituents in medicinal chemistry. Originally, they were exclusively considered only as hydrophobic moieties or Lewis bases due to their electronegativities. Thus, ability of the compounds containing chlorine, bromine or iodine to form direct close contacts of the type R-X$\cdots$Y-R', where the halogen X acts as Lewis acid and Y is any electron donor moiety, is much in contrast with their original perception.

1.1.4.2 Driving Force of the Halogen Bond

Ability of the compounds containing chlorine, bromine or iodine to form direct close contacts of the type R-X$\cdots$Y-R', where the halogen X acts as Lewis acid and Y is any electron donor moiety, is driven by the σ-hole. The σ-hole is a region of positive electrostatic potential (ESP) of halogen atom,[111] which appears along the extension of the covalent R-X bond (*cf.* Figure 1.3).

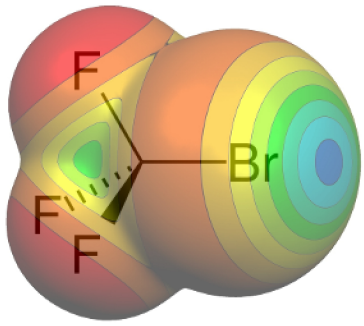

Figure 1.3: ESP (blue - positive, red - negative) of the CF_3Br molecule; σ-hole is blue region on the Br atom.

The σ-hole is usually surrounded by negative region of electrostatic potential (*cf.* Figure 1.3), thus the linear interactions of halogen species with nucleophiles (*cf.* Figure 1.2 A) and perpendicular interaction with electrophiles (*cf.* Figure 1.2 B) can be easily explained.[112–114] This arrangement of posi-

tive and negative regions of electrostatic potentials provide possibility to have attractive interaction also between two halogen atoms; so called "dihalogen bond".[115]

The concept of σ-hole is more general and can be used not only for halogen atoms, but also for Group IV-VI atoms. In fact many of the noncovalent interactions involving Group IV-VII atoms with negative sites can be explained based on electrostatic attraction between the σ-hole and particular negative site.[116] Thus, halogen bonding can be understood as a subset of the more general category of bonding: σ-hole bonding.
Moreover, a certain type of noncovalent interaction known as π-hole bonding can be also formally assign to same type of interaction as σ-hole bonding.[117] The π-hole is a region of low electron density (i.e. positive electrostatic potential) that is perpendicular to an atom in planar portion of the molecule.[117] This is in contrast with σ-hole where the "hole" appears along the extension of covalent bond to the atom. We will not discuss the issue of π-hole bonding any further. Therefore, the σ-hole along with π-hole bonding can be understood as two different forms of more general type of noncovalent interaction; so called hole-bonding.

The "hole" can be seen as a consequence of anisotropic distribution of electron density around the halogen atom.[118–120] What is the origin of this anisotropy? A free, ground-state atom has a spherically symmetric electronic charge distribution.[121] The electrostatic potential created around the atom by nucleus and electrons is positive everywhere,[122] contribution from dens nucleus dominates over that originated from dispersed electrons. When the atoms forms covalent bonds usually the electron density is polarized toward the bond region, leading to the decrease of electron density in the outer parts (along the extension of the bond) and increase in the region perpendicular to the bond axis. The final result of this polarization is that electrostatic potential in the regions around atoms perpendicular to the bond orientation becomes negative, while in the region around the elongation of the bond axis

becomes more positive, the σ-hole.
The anisotropic distribution of electron density of methyl halides has been nicely interpreted by Clark *et al.*[111] By utilizing natural population analysis (NPA),[123] the valence electron configuration of halogen atoms (X = Cl, Br, I) was approximated to $s^2p_x^2p_y^2p_z^1$ (where z is the direction of the C–X bond).

A convenient way of showing these features is by plotting the electrostatic potential on the surface of the molecule. The molecular surface is usually defined as 0.001 au (electrons per $bohr^3$) isodensity surface as proposed by Bader *et al.*[124] The electrostatic potential in the space can be explicitly evaluated using the eqn. 1.1.

$$V(\mathbf{r}) = \sum_{A=1}^{N} \frac{Z_A}{|\mathbf{R_A} - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}' d\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}, \tag{1.1}$$

where V($\mathbf{r}$) is potential created at any point $\mathbf{r}$ by nuclei and electrons of molecule, Z_A is charge of nucleus A located at $\mathbf{R}_A$ and $\rho(\mathbf{r})$ is molecule's electron density.

The electrostatic potential is real physical property, an observable, and can be determined both experimentally, by diffraction techniques and computationally.[125,126] Its sing in any point depends upon weather the effect of the nuclei (positive) or the electrons (negative) predominate. Nowadays, mapping of the electrostatic potential on isodensity surface, is heavily used as powerful tool in rational drug design (*cf.* Figure 1.4).

The size of the σ-hole can be characterized by the point on the isodensity surface which posses the most positive value of the potential; so called $V_{S,max}$ value. The $V_{S,max}$ value can be increased by increasing the atom's polarizability or decreasing the atom's electronegativity relatively to the remainder of the molecule. For example, the $V_{S,max}$ of the halogen atoms (X) in the methyl halides (CH_3X) increase in the order F < Cl < Br < I,

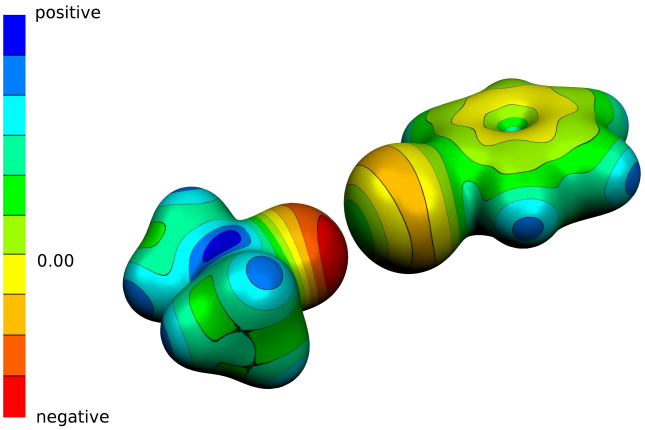

Figure 1.4: ESP of the bromobenzene$\cdots$acetone complex.

as halogens X become more polarizable and less electronegative. Further, the CF_3 group, which has greater electron-withdrawing power than CH_3 group, makes the $V_{S,max}$ value of the halogen X more positive in the CF_3X than in the CH_3X molecules.

As already mentioned, the existence of σ-hole is not inherent only for halogens but also for atoms of Group IV-VI. Moreover, it is common that Group IV-VI atoms have more than one σ-hole. When the atom has two or more σ-hole the one on the extension of the bond with more electron-withdrawing group is more positive. Specifically, the Se atom in $Se(CH_3)CN$ molecule has two holes. The one produced by NC–Se bond has $V_{S,max}$ value of 35 kcal.mol^{-1}, compared to 17 kcal.mol^{-1} for the H_3C–Se.[127] The tetravalent Group IV atoms have entirely positive surface. This can happen for Group V-VII atoms, when the rest of the molecule is highly electron-withdrawing e.g. Cl–CN, $Se(CN)_2$ and $As(CN)_3$.[127] Nevertheless, the σ-holes are still the most positive parts of the surfaces.
The situation is opposite when the Group V-VII atoms is much more electronegative than the bonding partner, e.g. CH_3–F or $(CH_3)_3N$. The covalent bond is in this case polarized toward more electronegative atom and neutralize the σ-hole. The σ-hole will be negative although less negative than its surroundings. Accordingly, these atoms can not participate on σ-hole inter-

actions. For some time it was believed that fluorine can not halogen bond at all.[127] However, it is now well established that fluorine can have positive σ-hole.[128–130] In next paragraph we will discuss σ-hole bonding considering only Group VII atoms, in more details.

1.1.4.3 Halogen Bonding and Protein$\cdots$Ligand Interactions

The diversity of interactions present in the single protein-ligand complex is enormous. The structure of the complex represents the balance between the attractive and repulsive interactions. Molecular recognition relies on the existence of specific interactions and the role of structure-based drug design is to identify and optimize such interactions between partner molecules. In order to identify the interactions and understand the energetic contribution of the particular interaction to the binding free energy the combination of statistical analysis of large groups of publicly available X-ray structures (CSD[131], PDB[132]) and experimental and theoretical studies on model systems is usually performed.[2]

In 1984 the survey of the CSD for short I$\cdots$N, O, S contacts revealed that these contacts are also formed in biologically relevant systems.[133] In protein-ligand environments, halogen bonds can be formed between halogenated ligand and any accessible Lewis base in the binding pocket.[134] The backbone carbonyl oxygen is the most observable Lewis base involved in halogen bond in protein binding sites.[108,135] Additionally, XBs can be formed involving amino acid side chain groups e.g. hydroxyls in serine, threonine and tyrosine; carboxylate groups in aspartate and glutamate; sulfurs in cysteine and methionine; nitrogens in histidine; and the π surfaces of phenylalanine, tyrosine, histidine, and tryptophan.[1]

The importance of halogen bond in drug discovery and drug development process can be justified by statistical analysis of organohalogens during this process. It is worth mentioning that in the Thomas Reuters Pharma, more

halogenated drugs (34%) are in clinical trials or at the stage of pre- and registered than that in the drug launched (26%), implying that halogenations have been much appreciated nowadays.[136] This holds, despite the fact that natural products contain little halogens in their structures. Another indicator which reflects the importance of the halogen bond is the ratio of heavy organohalogens to organofluorines (X/F ratio) at different stages of drug development. In general, the larger the ratio is, the more prevalent the halogen bond is. Specifically, the X/F ratio is 0.9, 0.9, 0.8, 1.0, 1.4 and 1.7 for drug discovery stage, clinical phase I, clinical phase II, clinical phase III, pre- and registered, and launched stages, respectively.[136] The content of heavy organohalogens increases steadily from clinical phase II (44.0%) to launched phase (63.3%).[136] This reveals that mostly the moieties with fluorines are excluded during drug development process. Thus, we conclude that the halogen containing compounds (i.e. prerequisite for the halogen bond formation) play important role in drug development process.

In following paragraphs we will describe the behavior of the energetics connected with halogen bond interaction. The geometrical preferences, relative strength, tunability of the halogen bond will be discussed in more details.

1.1.4.3.1 Strength and Tunability of Halogen Bond

The strength of halogen bond can be estimated via theoretical quantum chemical calculations. The theoretical method such as Couple Cluster[137–141] (CCSD(T)) or perturbation theory[142] (MP2) can be used for benchmark estimation of the strength of the halogen bonds. The typical model system of halogen bonded complex consists of carbonyl oxygen moiety (e.g. acetone or N-methylacetamide) and small organohalide molecule (e.g. halogenmethyl or halogenbenzene). The Table 1, taken from the work of Wilcken *et al.*,[1] summarize the CCSD(T) and MP2 benchmark calculations done for the set of small model complexes (BB in the first column stands for the N-

methylacetamide, "backbone" model system; CBS in the second column stands for the Complete Basis Set, see section 2.2.1).

complex	method	ΔE (kJ.mol^{-1})	X···O dist. (Å)	σ-hole angle (°)	work
$CH_3Cl\cdots OCH_2$	CCSD(T)/CBS	-4.9	3.26	166.8	Riley and Hobza [143]
$CH_3Br\cdots OCH_2$	CCSD(T)/CBS	-7.1	3.29	171.2	Riley and Hobza [143]
$CH_3I\cdots OCH_2$	CCSD(T)/CBS	-9.7	3.30	172.9	Riley and Hobza [143]
$PhBr\cdots OC(CH_3)_2$	CCSD(T)/CBS	-12.4	3.10	178.9	Kolář and Hobza [144]
PhCl···BB	MP2/aug-cc-pVDZ	-5.4	3.08	173.6	Hardegger *et al.* [145]
PhCl···BB	MP2/aug-cc-pVDZ	-7.5	3.13	180.0	Jorgensen and Schyman [146]
PhBr···BB	MP2/aug-cc-pVDZ-PP	-10.0	3.03	173.6	Hardegger *et al.* [145]
PhI···BB	MP2/aug-cc-pVDZ-PP	-14.5	3.08	173.6	Hardegger *et al.* [145]
PhBr···BB	MP2/aug-cc-pVDZ-PP	-12.1	3.05	180.0	Jorgensen and Schyman [146]
PhI···BB	MP2/aug-cc-pVDZ-PP	-16.9	3.09	180.0	Jorgensen and Schyman [146]
PhCl···BB	MP2/TZVPP	-5.6	3.12	171.2	Wilcken *et al.* [147]
PhBr···BB	MP2/TZVPP	-9.0	3.04	177.4	Wilcken *et al.* [147]
PhI···BB	MP2/TZVPP	-14.2	3.02	175.6	Wilcken *et al.* [147]
PhI···BB	CCSD(T)/CBS	-17.6	3.06	175.2	Wilcken *et al.* [147]

Table 1.1: Benchmark data for halogen-bonded model complexes.

The interaction distances are significantly shorter than the sum of the van der Waals radii of partner atoms for all complexes. The average halogen···oxygen distance equals to 3.18 (3.27) Å, 3.10 (3.37) Å and 3.06 (3.50) Å for Cl, Br, and I, respectively. Numbers listed in parenthesis represent sum of the van der Waals radii. Almost all considered complexes interact in ideal linear conformation (i.e. C_{ar}–X···O angle is close to 180°. The strength of the halogen bond do not overcome the strength of the typical hydrogen bond in water dimer complex, which amounts to approximately 5 kcal.mol^{-1}.[148] The stabilization energy of the C_{ar}–X···O halogen bond is estimated at 1.3 - 1.8, 2.2 - 2.9 and 3.4 - 4.2 kcal.mol^{-1} for X = Cl, Br and I. However, they bonds should suffer less from desolvation penalties than typical polar groups. Thus, halogen bond represents specific potential tool, which can be used for en-

hancement of the affinity between partner molecules.

The effect of halogen substitution on particular scaffold was investigated in several experimental studied.[145,149,150] These studies revealed that systems under investigation show increase of binding affinity, by two order of magnitude, after substituting H for I.[145,149,150]
Based on the fact that halogen bond is mainly driven by electrostatic attraction between electron-deficient σ-hole and electron reach Lewis base, size of the halogen bond can be controlled by introducing electron-withdrawing or electron-donating substituents on the halogenenated molecule. Riley *et al.*[151] investigated the effect of substitution of hydrogen atoms for fluorine atoms. The authors found that specified substitutions increase the size of σ-hole, what implies the increase of the halogen bond strength. Specifically, in the case of pentuple substitution (i.e. 1,2,3,4,5-pentafluorohalobenzenes) the strength of the halogen bond is doubled, with respect to mono-halobenzene.[151] Further, the introduction of two fluorine molecules (i.e. 1,3-difluoro-5-chloro-
benzene) provide halogen bond of the very similar strength as bromobenzene. The same relation holds also when comparing 1,3-difluoro-5-bromobenzene and iodobenzene.[151] The later knowledge is particularly valuable, since iodine groups are perceived as rather problematic. Hence, the substitution of the iodo-scaffold for the 1,3-difluoro-5-bromo one can serve as possible solution, when one needs to avoid iodine groups and at the same time retain binding affinity. However, one should bare in mind that listed fluoro-substitutions do not influence only the size of σ-hole, but they often significantly modify the entire electrostatic potential of the molecule. This usually leads to considerable changes in binding preferences of particular molecule.

Figure 1.5 (adopted from Wilcken's work) presents the electrostatic potential around the iodobenzene molecule and its fluoro-derivates. The electrostatic potential of the iodobenzene contains the σ-hole region surrounded by negative belt, see blue "ring" around C–I bond, and the π clouds over and bellow benzene ring (*cf.* Figure 1.5 a) In the case of perfluorinated iodoben-

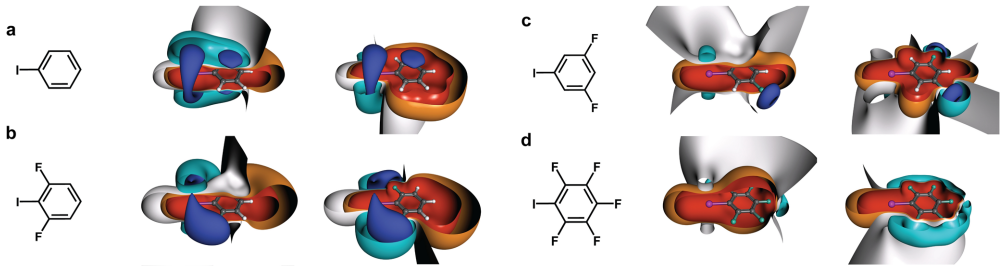

Figure 1.5: ESP of the iodobenzene and its fluoro-derivates.

zene potential looks substantially different and listed changes are evident. On the one hand, both negative regions of potential disappear, the only region where is the potential negative are areas close to the fluorine atoms (*cf.* Figure 1.5 d). On the other hand, the areas of positive potential at the region of original σ-hole are significantly increased. Double substitution of fluorines to ortho- position, in the case of 1,3-difluoro-2-iodobenzene (*cf.* Figure 1.5 b) leads to enlargement of the σ-hole. However, the σ-hole become shielded from the sides by the negative clouds of potential caused by presence of fluorine atoms. These simple examples illustrate that introducing of electron-withdrawing groups is rather complex issue,[151] specially, when more complex substituents are considered.[152]

Another possibility how to influence the strength of the halogen bond is by introducing changes to core scaffold that the halogen is attached to. Such study was performed by Hardegger *et al.*,[145] where 4 commonly used core scaffolds: benzene, pyrrole, pyridine and indole were tested on, how they influence the strength and accessibility of the σ-hole. The iodobenzene was considered as reference moiety. In the case of 3-iodopyridine and 7-iodo-1H-indole molecules, the σ-hole become larger and more accessible for Lewis base partners, with respect to iodobenzene. Whereas, in the case of the 3-iodo-1H-pyrrole molecule, the situation was opposite. Here, the iodine's σ-hole become almost entirely surrounded be the negative cloud of electrostatic potential; implicate almost zero capability of 3-iodo-1H-pyrrole to participate in halogen bonding.[1,145] These data demonstrate that the core scaffold has drastic impact on the ability of the molecule to create halogen bond and

that local properties of electronic structure of the core scaffold are important.[1,145,153]

1.1.4.3.2 Geometrical Preferences of Halogen Bond

Halogen bonds between druglike ligands and proteins almost never adopt optimal geometry because the halogen bond formation is distracted from its ideal conformation by other types of interactions which appear at ligand-protein interface. Therefore, it is important to identify with respect to which intermolecular coordinates is the strength of the halogen bond most sensitive. In general there are three parameters which influence the strength of halogen bond substantially.

- The distance $d_{X\cdots O}$ between halogen atom (X) and atom of Lewis base carrying lone electron pair (O) (*cf.* Figure 1.6 A)
- The angle $\alpha_{C-X\cdots O}$ defined between vector of the C–X bond and the X$\cdots$O contact (*cf.* Figure 1.6 A)
- The spacial orientation of halogen atom with respect to the Lewis base moiety; defined by angle $\alpha_{Y-O\cdots X}$ and dihedral angle $\delta_{Z-Y-O\cdots X}$ (*cf.* Figure 1.6 B)

There have been several theoretical studies focusing on relations between these three parameters and strength of the interaction, performed for different model systems.[143–147] The preferences of all three internal coordinates were studied also by Wilcken *et al.*[1] The dependence of the stabilization energy with respect to the distance $d_{X\cdots O}$ and angle $\alpha_{C-X\cdots O}$ was investigated using the chloro-, bromo-, iodobenzene$\cdots$N-methylacetamide complexes as model systems.[1] In the case of the spacial orientation of the halogen with respect to partner molecule, complexes of iodobenzene with N-methylacetamide and dimethylsulfide were utilized.[1]
Here are listed the most important conclusions regarding energetic and geometric preferences of particular intermolecular coordinates.[1]

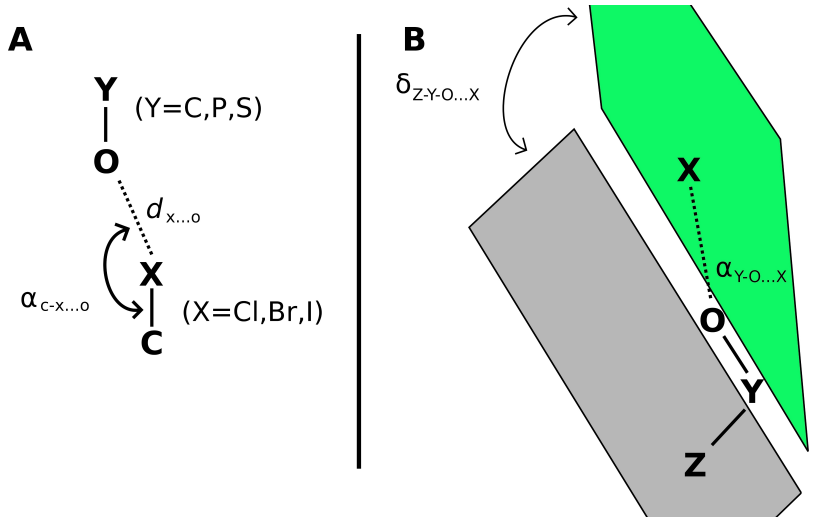

Figure 1.6: Main intermolecular coordinates of the halogen bond.

distance $\mathbf{d}_{X\cdots O}$

- the stabilization energy is reduced by half, with respect to optimal value, after increasing the X$\cdots$O distance by 1 Å.
- decrease of stabilization upon elongation of the X$\cdots$O contact can be attributed to decrease in overlap between σ-hole and lone electron pair
- at the distances close to 5 Å iodobenzene still shows substantially better interaction energies than benzene
- reduction of the distance by 0.35 Å decreases stabilization by 50% for all three heavy halides
- no attraction interaction occurs bellow 2.5 Å
- better stabilization than for benzene can be achieved for distances equal or larger than 2.92 Å, 2.74 Å and 2.63 Å for Cl$\cdots$O, Br$\cdots$O, and I$\cdots$O contact, respectively
- iodine tolerates displacements from the equilibrium distance significantly better than chlorine or bromine

- the advantage for chloro-, bromo-, iodobenzene over benzene is approximately 0.5, 1.0 and 2.75 kcal.mol^{-1} for the formation of halogen bond in optimal geometry

angle $\alpha_{C-X\cdots O}$

- significant decrease of stabilization is observed upon increase in deviation from 180°, for all halogens
- deviations between 25° and 30° decreases stability by 50%
- deviations exceeding 40° are connected with no attractive interaction, thus halogen bonds should be considered only for geometries where angle is 140° or larger

angles $\alpha_{Y-O\cdots X}$ **and** $\delta_{Z-Y-O\cdots X}$

- dimethyl sulfide molecule (methionine model system) as Lewis base: strong preference for orientations of the halogenated moiety perpendicular to the plane of the methionine model system are/were observed. Only such orientation of the interacting molecules provides attractive interaction. This preference is caused by presence of two σ-holes in the molecular plane along the extension of the C–S bonds and two strongly negative areas above and bellow the plane
- N-methylacetamide (backbone carbonyl model system) as Lewis base: provide almost no constraints for orientation of partner molecule, except from the orientations where sterical clashes between backbone and halogenated scaffold occure
- electrostatic potential around oxygen atom of backbone carbonyl model system is negative and quit uniform/isotropic; maximum reduction of stabilization energy is only about 30% of the optimal value

Despite restrictions which need to be met in order to obtain attractive halogen bond interaction in the case of methionine sulfur atom, the interacting potential of this Lewis base moiety for halogen bonding is still appreciable/significant. This is due to the fact that in many cases sidechains of the aminoacids are flexible, thus sulfur atom can adapt to ligand, facilitating attractive perpendicular orientation with partner molecule.[1] On the other hand, backbone is often fixed within more rigid secondary structure elements. Thus, the fact that backbone oxygen tolerate broad range of halogen-bonding geometries has important implications for molecular design.[1]

Chapter 2

Methods

2.1 Noncovalent Data Sets

The important role of data set in method development is indisputable and is known for quit a long time. As already noted, probably the most important reason of having reliable test set is its necessity for parametrization and validating of lower level computational methods. For example, the ability to reproduce the energetics of the G2 data set[154] or extensions to it has become the *de facto* standard for measuring the accuracy of a new computational method. As variety of new methods were developed in recent years, because of the fast development of new computer hardware and software. The need for high-quality data sets is enormous and development in this direction is remarkable, especially in last decade. For example, General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions (GMTKN30) data set,[155] Noncovalent Interaction Energy data set (NCIE53),[156–158] etc. have recently been introduced. Before brief description of some of the most commonly used noncovalent data sets, we list the most important properties which should be cover by any benchmark database set: (i) balanced representation of all important bonding motifs, and (ii) easy extendability in the future for new structural motifs.

2.1.1 JSCH2005 Data Set

Jurecka, Sponer, Cerny, Hobza presented in 2005 set (JSCH2005)[159] of Møller Plesset/Complete Basis Set (MP2/CBS) and Coupled Cluster Single Double and perturbative Triple/Complete Basis Set (CCSD(T)/CBS) interaction energies and geometries for more than 100 DNA base pairs, amino acid pairs, and other model complexes. Extrapolation to the CBS limit was done using the two-point Truhlar's[160] and Helgaker's[161–163] extrapolation scheme utilizing different basis sets (aug-cc-pVDZ + aug-cc-pVTZ, aug-cc-pVTZ + aug-cc-pVQZ and cc-pVTZ + cc-pVQZ). CCSD(T) correction terms, were determined in relatively smaller basis sets (6-31G**(0.25,0.15) and cc-pVDZ). Two types of complex geometries were considered, optimized and experimental ones. It was shown that utilizing these combination of basis sets for extrapolation can lead to rather large relative errors (up to 10-15%). However, we should bare in mind that presented results were, at the time of publication, considered as rather very accurate. The JSCH2005 data set, could be preferably used for first screening (verification) of the methods rather that for parametrization, as it is relatively large (over 100 complexes).

2.1.2 S22 Data Set

The S22 data set is part of the JSCH2005 data set.[159] It consists of 7 H-bonded, 8 disperse-bounded and 7 mixed noncovalent complexes. S22 set has become the most popular interaction energy database. It has been used for the parametrization and validation of many different computational methods (see bellow). However, recently new data sets has been developed (see bellow), as some shortcomings of the S22 set has been reported. Specifically, several studies showed that accuracy of the presented interaction energies has been already overcome. These errors were evaluated to few tenths of kcal.mol^{-1} (not exceeding 0.5 kcal.mol^{-1}), when compared to the most accurate CCSD(T)/CBS values.[164–167] Another problem of the S22 set is that it is heavily weighted toward nucleic acid-like structures, containing many base pair-like (cyclic) hydrogen bonds and many stacked aromatic

complexes. Consequently, there are several binding motifs that are strongly under-represented (e.g. single hydrogen bonds and aromatic-aliphatic dispersion interaction) or practically missing (e.g. such as aliphatic-aliphatic dispersion interaction). The single-hydrogen bond binding motif was introduced in the data set, through four new complexes. Specifically, the methanol dimer (C_1), the methanol $\cdots$ formaldehyde (C_1), the methyl amide dimer (α) and the methyl amide dimer (β).[168] The new augmented S22 data set was named S26.[168] Merz and co-workers and Grafova and co-workers independently enhanced the S22 data set, by the introduction of dissociation curves, in order to increase the robustness of the data set.[169,170] Resulted data sets are denoted as S22+[169] and S22x5,[170] respectively.

2.1.3 S66 Data Set

S66 data set and its "dissociation curve" extension (S66x8) were presented by Rezac and Hobza in 2011.[171] It was designed with special emphasis to eliminate drawbacks of the S22 data set. The most significant enhancements with respect to the S22 set are following: (i) inclusion of 3 times larger amount of complexes, (ii) balance representation of all main types of noncovalent interactions, (iii) use of same basis set and geometry optimization protocol for all complexes. Therefore, the S66 data set as upgraded successor of the S22 set has potential to become also extensively used by computational chemists.

2.1.4 Other Noncovalent Data Sets

Many other noncovalent data sets are available. For example ten out of 30 data sets, which are included in Grimme's GMTKN30[155] superdata set are noncovalent data sets. Zhao's and Truhlar's superdatabase includes NCIE-53[156–158] category, which is actually group of eight individual noncovalent data sets. Several new data sets, which include much larger complexes, contrary to all previous sets, have been presented (e.g. S12L,[172] L7[173]).
Noncovalent data sets, which were exclusively designed in our laboratory (e.g.

S22,[159] S66,[171] P26,[174] A24,[175] X40x10[176]) are available through the BEGDB web site (www.begdb.com).[177]

2.2 Interaction Energy Calculations

2.2.1 Extrapolation to the Complete Basis Set Limit

It well known that convergence of the energy with respect to the size of the basis set is relatively slow and errors resulting from the use of incomplete basis set are typically the most significant ones. Majority of the extrapolated schemes are based on the fact that the convergence of the energy toward the complete basis set depends on the basis set cardinal number (X). Here we will present only the two point extrapolation schemes as it is known that the inclusion of third set of energies of low quality (e.g. double-ζ basis) especially for small systems can spoil the quality of the resulting energy. The most commonly used method for basis set extrapolation to the complete basis set was proposed by Helgaker.[161–163] This scheme treats Hartree-Fock (HF) and correlation part of the energy separately (see. eqn. 2.1 and 2.2)

$$\Delta E_X^{HF} = E_X^{CBS} + A\ exp(-\beta X), \tag{2.1}$$

where E_{HF} is basis set limit for HF energy, A is pre-exponential factor, X is the cardinal number of the basis set (2, 3, 4 for double-, triple-, quadruple-ζ, respectively) and β is set to 1.54 (for extrapolation from triple- and quadruple-ζ basis sets) or 1.43 (for extrapolation from double- and triple-ζ basis sets).

$$\Delta E_X = E_{CBS} + B/X^{\beta}, \tag{2.2}$$

where E_{CBS} is basis set limit of the correlation energy, B is pre-power factor, X is the cardinal number of the basis set and β is set to 3.

Truhlar developed extrapolation method,[160] which similar to the Helgaker's, treats HF and correlation parts of the energy separately. Specifically, the correlation energy is also extrapolated through the eqn. 2.2, but with the β coefficient being 2.2 and 2.4 for MP2 and Coupled Cluster Single and Double (CCSD) or CCSD(T) calculations, respectively. Further, the HF energy is, contrary to the Helgaker's method, treated also through eqn. 2.2, but with the β coefficient being 3.4.
Both schemes, Truhlar's as well as Helgaker's, were designed to utilize systematically improved basis sets (e.g. Dunning's augmented or non-augmented sets). Extrapolation scheme, of slightly different kind was developed by Kim.[178–180] Kim's method is designed for an extrapolation of the interaction energy. It relies on the fact that two sets of interaction energies: corrected and uncorrected for the basis set superposition error (for BSSE correction see section 2.2.2) converge to the same value. Thus, it is possible to use data from different basis sets. The CBS interaction energy is defined by eqn. 2.3

$$\Delta E_{CBS} = 1/2(\delta_X \epsilon_{X+1} - \delta_{X+1}\epsilon_X)/(\delta_X - \delta_{X+1}), \tag{2.3}$$

where δ_X and ϵ_X (and δ_{X+1} and ϵ_{X+1} analogously) are defined by eqn. 2.4 and 2.5

$$\delta_X = \Delta E_X^b - \Delta E_X^n \tag{2.4}$$

$$\epsilon_X = \Delta E_X^b + \Delta E_X^n, \tag{2.5}$$

where ΔE_X^b and ΔE_X^n are BSSE corrected and uncorrected interaction energies in basis set with cardinal number X.

Nowadays, it is still hard to determine the accuracy of any of the above mentioned extrapolated methods, because there is not enough of statistical data. It is well known that performance of any of the basis set strongly depends on method applied, size and geometry of the molecule. For small

and medium-size complexes, the aug-cc-pVDZ basis set provides quit converged energies at HF level, but not at the MP2 level. It is not recommended to use the MP2/aug-cc-pVDZ energies for extrapolation procedure, because these can spoil the quality of the CBS energy. However, when applied for extended complexes, the MP2 energies in aug-cc-pVDZ basis should be accurate enough. This is a consequence of the overlap of the basis functions. Moreover, applying the aug-cc-pVDZ basis for extended clusters is often the only feasible choice mostly because of two reasons: (i) prohibitive computational cost of the calculation, and (ii) possible convergence problems caused by linear dependencies of the basis set. In general, all three above mentioned extrapolation procedures perform similarly, with differences in final extrapolated energies usually not exceeding 0.1 kcal.mol^{-1}.

Another possibility how circumvent the problem of slow convergence of (correlation) energy is by employing explicitly correlated methods. Nowadays, medium-size systems can be investigated by utilizing methods such as R12-/F12-MP2[181–188] and R12/F12-CCSD.[189–197] The main advantage of explicitly correlated methods, compared to conventional methods, is their significantly faster convergence toward CBS limit. As they are able to provide results of CBS quality already in the aug-cc-pVTZ and heavy-aug-cc-pVTZ basis sets.[198]

2.2.2 Supermolecular Approach

Within supermolecular approach the interaction energy of a binary complex A$\cdots$B (ΔE(A$\cdots$B)) is defined by eqn. 2.6

$$\Delta E(A\ \cdots B) = E(A\ \cdots B) - E(A) - E(B), \tag{2.6}$$

where E(A$\cdots$B) stands for the total electronic energy of the complex and E(A) and E(B) are the electronic energies of the partner molecules.
On the one hand, the supermolecular approach has several practical features: (i) it is applicable to any type of molecular cluster, (ii) it yields wave function which can be used for derivation of various properties of the cluster,

and (iii) accurate interaction energies can be obtained, provided that sufficiently large basis set is used and major part of correlation energy is included. On the other hand, the main drawback of the supermolecular approach is its basis set inconsistency which lead to basis set superposition error. The BSSE is purely mathematical artifact due to the fact that basis sets of different size are used for energy evaluations of cluster and interacting molecules. In the case of the complex calculation the electronic energy of the monomers is artificially lowered due to the borrowing of the basis set from partner subsystem. In 1970 the Boys and Bernardi introduced counterpoise correction (CPC), which eliminates the BSSE.[199] The principle of the method is very simple: subsystems are not treated in their own basis set but in the basis set of whole cluster i.e. all three energy calculations (cluster and subsystems) are performed with the basis set of same size.

Values of E(A$\cdots$B) and the sum of E(i) in the eqn. 2.6 typically differ only by several kcal.mol^{-1}, while the value of E(A$\cdots$B) is many order of magnitude larger than that of ΔE(A$\cdots$B). Therefore, in order to obtain accurate interaction energy, all the energies at right hand side (rhs) in the eqn. 2.6 must be calculated with extremely high accuracy. Because the electronic energy of the molecule can be expressed as the sum of HF and correlation energy, the interaction energy can be expressed as the sum of Hartree-Fock interaction energy and correlation interaction energy (see eqn. 2.7).

$$\Delta E = \Delta E^{HF} + \Delta E^{corr} \tag{2.7}$$

Nowadays, the accurate evaluation of the ΔE^{corr} part of the interaction energy represent the most challenging part of the calculation from the computational point of view. It is well known that the different types of noncovalent interaction require different approaches from the computational point of view. For example, the quantitative description of $\pi\cdots\pi$ stacking intermolecular interaction is very demanding. The $\pi\cdots\pi$ stacking interactions are dominantly formed by dispersion, which is pure correlation effect, hence not covered at the HF level of theory (see eqn. 2.7). The local or semi-local

character of density functionals is able to reproduce only short-range and middle-range correlation effects.[51] Further, the MP2 method is the simplest method able to describe long-range correlation effects from the right reason. However, it is well known that MP2 overestimates the $\pi\cdots\pi$ stacking interaction.[52–55] Moreover, for small and middle size basis set superposition error can be as big as dispersion. Therefore, use of more elaborate *ab initio* methods in combination with sufficiently large basis set is needed in order to obtain directly the correct description of this type of interaction.
The situation is not so complicated for hydrogen bonded complexes, as it is known that MP2 provides accurate description of hydrogen-bonding. In next sections different supramolecular methods, which are used for the description of noncovalent interactions, will be described.

2.2.2.1 Reference Calculations

Reference calculations, despite their enormous computational and time demands, continues to attract significant attention by computational chemists. This can be (at least partially) understood as a consequence of the fact that variety of new computational methods, which potentially improve the accuracy of the results and/or computational efficiency, have been developed. After realizing that most of these methods include at least one empirical parameter (see sections 2.2.2.2), which was (were) usually obtained by fitting to the set of benchmark data i.e. data set (see section 2.1), causality of our original statement become clear. As, high-quality (reference) Quantum Mechanical (QM) calculations are necessary for obtaining suitable data sets.

The CCSD(T) method, among computational chemists also known as "golden standard", due to its outstanding accuracy/computational cost ratio, is the most commonly used method for utilizing reference calculations. The CCSD(T) method is based on the concept of Coupled Cluster (CC) theory, which uses exponential form of the wave operator and its expansion into the clusters of excitation operators.[137–141] The CC methods are size-extensive and their convergence toward full configuration interaction value is faster com-

pared to the methods with same asymptotic scaling with respect to system size. Moreover, CC methods can be systematically improved upon inclusion of higher excitation operator.

Before discussing the CCSD(T) in more details, we should mention about the Quadratic Configuration Interaction Single Double and perturbative Triple method (QCISD(T)).[200] As it is probably the only alternative to the CCSD(T) method, when considering methods suitable for reference calculations. It was shown that the QCISD(T) method provides very similar results as the CCSD(T) as far as single-reference character of the wave function is maintained.[201,202] The main advantage of QCISD(T) over CCSD(T) is it's slightly less expense, as some of the integrals are not evaluated in QCISD(T), contrary to CCSD(T).[203] Thus, QCISD(T) can become the method of choice, especially when feasibility of the CCSD(T) method starts to become very problematic.

2.2.2.1.1 Affordability of the CCSD(T)

From year to year, the affordability of the CCSD(T) method is increased because of the fast progress in computational hardware and software. The scaling of the CCSD method is $N_o^2 N_v^4$ and $N_o^3 N_v^3$, where N_o stands for the number of correlated occupied and N_v the number of active virtual orbitals. The inclusion of perturbative triple excitation into the calculation, increase the scaling to $N_o^3 N_v^4$. Nowadays, system of size approximately 30 atoms can be routinely treated, however systems with more than about 50 atoms are still significantly impractical. So far, the largest published systems for which regular CCSD(T) calculation were performed have approximately 70 atoms: coronene dimer[204] and guanine-cytosine step from DNA.[205]

2.2.2.1.2 Convergence of the CCSD(T)

The CCSD(T) method is especially successful for assessing ground-state energies and properties for system with single-reference character. The overall accuracy of the CCSD(T)/CBS values goes beyond "chemical accuracy" (1 kcal.mol^{-1}) but usually hardly reaches "subchemical accuracy" (0.1 kcal.mol^{-1}).[206–208] The largest error of the method is attributed to the incompleteness of the basis set, thus the convergence behavior of the CCSD(T) method toward CBS will be discussed firstly.
The CCSD(T) interaction energy can be partitioned into three terms, see eqn. 2.8

$$\Delta E^{CCSD(T)} = \Delta E^{HF} + \Delta E^{MP2} + \Delta\Delta E^{CCSD(T)}, \tag{2.8}$$

where ΔE^{HF} is the HF interaction energy, ΔE^{MP2} is the MP2 correlation interaction energy and $\Delta\Delta E^{CCSD(T)}$ is the higher order correlation correction to interaction energy. The $\Delta\Delta E^{CCSD(T)}$ term is defined as difference between interaction energies calculated at the CCSD(T) and MP2 level ($\Delta E^{CCSD(T)}$ - ΔE^{MP2}). All three term on the rhs of eqn. 2.8 exhibit different speed of convergence. The convergence properties of the ΔE^{HF} and the ΔE^{MP2} components were already discussed in the section 2.2.1. The rate of convergence of the $\Delta\Delta E^{CCSD(T)}$ term was investigated in several studies.[209–212] These studies revealed, that the convergence of this term is very fast for the vast majority of complexes, however, complexes with slow convergence has also been reported (e.g. benzene$\cdots Na^+$, benzene$\cdots$water).[210,211] As shown by Pitonak *et al.*, for $\pi\cdots\pi$ complexes, the use of small basis set (6-31G*(0.25) can lead to high underestimation of the $\Delta\Delta E^{CCSD(T)}$ (up to 10-20%) and use of aug-cc-pVTZ set brings 1-3% agreement with CBS value.[213,214] This relatively fast rate of convergence of the $\Delta\Delta E^{CCSD(T)}$ term, can be understood as a consequence of the error compensation between convergence of the CCSD and (T).[213,214]
The importance of the $\Delta\Delta E^{CCSD(T)}$ term is different for H-bonded, $\pi\cdots\pi$ and other types of complexes. For the $\pi\cdots\pi$ complexes the inclusion of the

term usually leads to qualitative/quantitative agreement. While, for the H-bonded or charge-transfer (CT) complexes neglect of the term has just marginal effect on the accuracy. Moreover, for the complexes stabilized dominantly by electrostatic neglect of the term can lead to more balanced description.[211]

2.2.2.1.3 Approximations in the CC Theory

The magnitude of errors resulted from neglecting higher order excitations within CC expansion, are not so pronounced as errors resulted from incompleteness of the basis set, as long as we are not approaching subchemical accuracy, where errors resulted from both types of approximation might be of similar importance. As the effect of higher order excitations is sensitive to truncation/incompleteness of the basis set used, quantification of the errors caused by their neglect is difficult, therefore there are only several studied regarding this issue in literature. Based on several studies which focused on this issue, following conclusions can be drown: (i) for $\pi\cdots\pi$ complexes the effect of approximating triples by perturbation theory amounts to roughly 1% of the $\Delta\Delta E^{CCSD(T)}$ term, for H-bonded complexes same effect can amounts up to 30% of the $\Delta\Delta E^{CCSD(T)}$ term.[215]; (ii) for $\pi\cdots\pi$ complexes the effect of neglecting quadruples is about 5-10% of (T) contribution, being repulsive.[216–218] As, higher order correlation terms which are not included in the CCSD(T) are repulsive, the CCSD(T) provides upper bound to stabilization energy.[216]

Quit recently Rezac and Hobza examined the accuracy of the estimated CCSD(T)/CBS binding energies for the set of twenty four small complexes.[175] This study revealed that the average relative error of the CCSD(T)/CBS energies with respect to CCSDT(Q)/CBS energies, which were considered as reference, amounts to 1%.[175] Further, as long as sufficiently large basis sets are utilized (e.g. at least heavy aug-cc-pVQZ for HF part, cc-pVTZ and cc-pVQZ for MP2 part and aug-cc-pVDZ for ΔCCSD(T) term) the overall error in the CCSDT/CBS interaction energy should not be larger than 3%.

Another quit recent study by Rezac and Hobza, focused on the performance of the CC expansions which go beyond the CCSD(T) method (e.g. CCSD[T], CCSD(TQ), CCSDT(Q), CCSDTQ).[219] These methods were tested toward six small noncovalent complexes in small and medium size basis sets and following conclusions were drawn: (i) The CCSDT(Q) interaction energies reproduce the CCSDTQ results with negligible errors; (ii) The inclusion of quadrupole excitations (CCSDT $\rightarrow$ CCSDT(Q)) changes the interaction energies by about 10 cal.mol^{-1} on average in aug-cc-pVDZ basis; (iii) The CCSD(T) and CCSD[T] methods provide more accurate interaction energies, with respect to the CCSDT(Q) reference, than more expensive CCSDT method.[219]

2.2.2.2 Post-Hartree-Fock Methods

Necessary condition for any commonly used computational method used for evaluation of the interaction energy, is its ability precisely and efficiently account for the correlation energy. Thus methods which go beyond Hartree-Fock are used for investigation of noncovalent interactions. In last decades, many of new computational approaches has been developed, in order to accurately and/or efficiently treat the issue of correlation energy. In next few sections some of these approaches will be described.

The simplest method which account for correlation energy from the right reason is the MP2.[142] As already mentioned, it is known that MP2 significantly overestimates stabilization for the $\pi\cdots\pi$ complexes.[52–55] This is due to the fact that the dispersion energy is within MP2 described only at uncoupled Hartree-Fock (UCHF) level.[54,55] Thus several different approaches, how to deal with aforementioned drawback of the MP2 method, have been developed (see bellow).

2.2.2.2.1 Spin-Scaled MP2

In the spin component scaled MP2 method (SCS-MP2), the idea of individual scaling of the antiparallel (singlet,S) and parallel (triplet,T) spin components of the correlation energy within MP2 was introduced (see eqn. 2.9).[220]

$$Er_{corr,scaled} = p_S E_S + p_T E_T, \tag{2.9}$$

where E_S and E_T are singlet and triplet components of the energy and $p_S = 6/5$ and $p_T = 1/3$ are respective scaling parameters. The singlet parameter p_S was derived from theory, whereas, p_T was fitted to a set of reference data (reaction and atomization energies).[220]

Several studies on noncovalent complexes revealed that SCS-MP2 significantly eliminates the problem of plain MP2 (overestimation of the $\pi\cdots\pi$ complexes), only for nonpolar and slightly polar $\pi\cdots\pi$ complexes.[221–225] One should bare in mind, SCS-MP2 parameters were derived regardless noncovalent interaction. However, considering different types of noncovalent complexes, the performance of both methods (SCS-MP2 and MP2) is comparable.[226,227] Consequently, following Grimme's strategy, several new spin-scaled MP2 methods such as: SCSN-MP2,[227] SSS(MI)-MP2[223] or SCS(MI)-MP2[223] were developed. All these methods were designed with special emphasis on accurate and balance description of wide range of noncovalent interactions. Therefore, spin scaling coefficients within these methods were derived by fitting against the set of reference interaction energies. In particular, SCSN-MP2, SSS(MI)-MP2 as well as SCS(MI)-MP2 were optimized for best performance on nucleic acid base pairs and S22 data set,[159] respectively.

SCSN-MP2 and SSS(MI)-MP2 consider only one, same spin scaling coefficient (p_T), whereas, SCS(MI)-MP2 considers both of them (p_S and p_T). The increased accuracy, compared to SCS-MP2 and MP2, is most pronounced in the case of SCS(MI)-MP2. Specifically, the Root Mean Square Deviation (RMSD) for the S22 data set of SCS(MI)-MP2, SCS-MP2 and MP2 in the cc-pVTZ amounts to 0.31, 1.45 and 0.99 kcal.mol^{-1}.[223] Moreover, per-

formance of SCS(MI)-MP2 for H-bonded complexes is still not significantly better than this of MP2, but is clearly improved over this of SCS-MP2.[223]

In complete analogy with MP2, the individual scaling of the singlet and triplet spin components of the correlation energy can be utilized for CCSD method (*cf.* eqn. 2.9). Thus, SCS-CCSD[228] and SCS(MI)-CCSD[229] methods can be understood as counterparts to SCS-MP2 and SCS-MI-MP2. SCS-CCSD as well as SCS(MI)-CCSD were optimized toward same training sets as SCS-MP2 and SCS(MI)-MP2 (see above). Several studies demonstrated ability of both methods to provide accurate interaction energies. In particular, error of SCS-CCSD for several configurations of the benzene dimer is bellow 0.2 kcal.mol^{-1}, whereas errors for the SCS-MP2 approach almost 1.0 kcal.mol^{-1}. Further, SCS(MI)-CCSD was shown to provide stabilization energies for the S66 data set in excellent agreement with the CCSD(T) reference.[171]

2.2.2.2.2 MP2C

Replacing the UCHF dispersion from MP2 interaction energy by some other, more sophisticated, form of dispersion (e.g. scaled coupled HF dispersion from time-dependent HF) represents rather different way of treating the issue with MP2 overestimation; see eqn. 2.10.

$$\Delta E = \Delta E^{MP2} - E_{disp}^{UCHF} + E_{disp}^{M}, \tag{2.10}$$

where ΔE is corrected MP2 interaction energy, ΔE^{MP2} is original supermolecular MP2 interaction energy, E_{disp}^{UCHF} is dispersion from UCHF theory and E_{disp}^{M} is time-dependent HF or time-dependent density functional theory (abbreviated as TD-DFT). This idea was originally proposed by Cybulski and Lytle.[54] MP2C, proposed by Pitonak and Hesselmann,[230] evaluates E_{disp}^{M} term by utilizing TD-DFT method, specifically by LHF/xALDA (localized HF method[231] combined with the exchange-correlation kernel from the adiabatic local density approximation (ALDA) in its exchange-only variant).

The excellent accuracy of the method was confirmed not only for S22 data set but also for medium-size and large systems.[173,230]

2.2.2.2.3 Scaled MP2

MP2.5 method takes advantage of the fact that for the $\pi\cdots\pi$ stacking complexes, both MP2 (overestimated) and MP3 (underestimated) deviate from accurate CCSD(T) results by almost the same amount.[232] This empirical observation holds because of following: the fourth-order effect of triples is always attractive and its value is roughly 50% of the MP3 third-order correction.[232] MP2.5 interaction energy, is defined by eqn. 2.11

$$\Delta E^{MP2.5} = \Delta E^{MP2} + 1/2E^{(3)} = \Delta E^{MP2} + 1/2(\Delta E^{MP3} - \Delta E^{MP2}), \quad (2.11)$$

where $\Delta E^{MP2.5}$ and ΔE^{MP2} are MP2.5 and MP2 interaction energies and $E^{(3)}$ is third-order correction which is defined as follows $E^{(3)} = \Delta E^{MP3} - \Delta E^{MP2}$. Accuracy of the MP2.5 method is in principle never considerably worse than accuracy of the MP2. Moreover, MP2.5 can be significantly better, especially for complexes for which the third-order correction presents an important part of the interaction energy (e.g. $\pi\cdots\pi$ complexes). Very practical feature of the scaled third-order correction is that its basis set convergence toward CBS is very similar to that of CCSD(T), which allows us to use same scheme (see eqn. 2.8). MP2.X method is just more empirical generalization of MP2.5, which is done by introducing scaling parameter for third-order correction (i.e. in eqn. 2.11, 1/2 is replaced by scaling parameter).[233] The optimum value of the scaling parameter has been optimized against the S22 data set, for various types of basis sets. Although MP2.5 and MP2.X represent only an empirical correction to approximate higher order correction terms which are included in the CCSD(T) interaction energy, their performance for various types of noncovalent complexes is quite accurate.[171,173]

2.2.2.3 Density Functional Theory Based Methods

It is well known that standard density functionals (LDA, GGA, meta-GGA, hybrids) are not able to account for long-range, van der Waals (dispersion) interaction, because of their local (semi local) dependence on electron density. Thus, providing relatively poor description of noncovalent interactions, especially, those dominated by dispersion. At this point it should be stressed that DFT itself is exact (provides the exact solution of the Schrodinger equation), including long-range correlation dispersion. Many different approaches were developed to eliminate the shortcoming caused by locality of density functionals:

- atom-atom pairwise empirical dispersion correction (e.g. Grimme's schemes,[234–236] Jurecka's scheme,[237] etc.)
- truly nonlocal density functionals (e.g. vdW-DF,[238] LC-BOP+ALL,[239,240] vdW-DF-09,[241] VV09,[242] VV10[243])
- reparametrization of current functionals (e.g. M05,[244] M05-2X,[245] M06,[246,247] M06-2X[246,247])
- double hybrid functionals (e.g. PW2P-LYP,[248] B2P-LYP[249])
- long-range corrected functionals (e.g. wB97,[250] wB97X,[250] CAM-B3-LYP[251])

In next paragraphs several variants of atom-atom pairwise empirical dispersion corrections will be briefly reviewed.

2.2.2.3.1 Atom-Atom Empirical Dispersion Corrections

A most straightforward way how to include dispersion in DFT, can be done by simple addition of the dispersion energy to the plain DFT result. Dispersion energy term can be within this approach expressed by following atom-atom pairwise functional form:

$$E_{disp}[\rho] = -\sum_{i \leq j} \sum_{n \geq 6} C_n^{ij}(\rho) R_{ij}^{-n} f_{damp}^n(R_{ij}, \rho), \tag{2.12}$$

where i and j are atom labels, $C_n^{ij}[\rho]$ are, in general, density-dependent dispersion coefficients of particular atomic pairs (ij), R_{ij} is the interatomic separation and $f_{damp}(R_{ij},\rho)$ is, in general , density-dependent damping function. Equation 2.12 can be derived from the second-order perturbation theory expression for dispersion energy between isolated molecules.[252] Nowadays, variety of methods, which utilize above mentioned functional form for dispersion energy (see eqn. 2.12), are commonly used. Forms of the empirical dispersion correction utilized by majority of methods differ mostly in following three aspects: (i) how many terms are included in eqn. 2.12 (i.e. n=6,(8,10)), (ii) which type of damping function is utilized, and (iii) degree of approximations used to derive $C_n^{ij}[\rho]$ coefficients.

Lets make small comment on dispersion energy in the DFT, before presenting particular correction schemes in more details. As already noted the long-range correlation energy, is the energy component not recovered by recent local or semilocal density functionals. However, we should keep in mind that these functionals can cover, at least to some extent, the medium-range correlation energy at the van der Waals distances, where intermolecular overlap is not negligible. Therefore, eqn. 2.12 includes damping function, which must be considered from two main reasons : (i) to avoid double counting and resulting overestimation of the interaction, and (ii) because overlap invalidates multipole expansion of the dispersion.

So it is obvious that damping function is crucial, especially in the range of small and medium distances. Strictly speaking, as the term "dispersion" originally referred to the long-range forces between systems with negligible overlap, it should not be used for corrections specified by eqn. 2.12. Other different terms as "damped dispersion" and "overlap-dispersive interaction" were suggested by Pitonak *et al.*[253] and Grimme,[254] respectively. Despite this inconsistency in definitions, most of the chemist name this correction simply "dispersion", as it reflects the nature of this interaction resulting from the long-range correlation of instantaneous electric multipole moments

of fluctuating electronic density.

2.2.2.3.2 Grimme's Corrections

In 2004 Grimme presented his first generally oriented dispersion correction in DFT, named DFT-D.[234] The DFT-D scheme utilizes dispersion correction in following form:

$$E_{disp}[\rho] = -s_6 \sum_{i \leq j} C_6^{ij} R_{ij}^{-6} f_{damp}^{6}(R_{ij}), \tag{2.13}$$

where s_6 is global scaling parameter, i and j are atom labels, C_6^{ij} are dispersion coefficients of particular atomic pairs (ij), R_{ij} is the interatomic separation and $f_{damp}(R_{ij})$ is damping function. Equation 2.13 represents slightly different form for correction than eqn. 2.12, as it introduces global scaling parameter (*cf.* eqn. 2.12 and 2.13). Further, the DFT-D includes only one term from expansion (i.e. n = 6, *cf.* eqn. 2.13) and utilizes isotropic and density independent C_6^{ij} coefficients. Moreover, employees modified Fermi type of damping function, where no scaling of vdW radii is introduced (for more details see work by Jurecka *et al.* and references therein.[237]). Despite, above noted simplification it provides very good results across a wide range of weakly bonded complexes of different nature.[234] This can be seen as a consequence of fitting procedure, as the global scaling parameter (s_6) was introduced in the eqn. 2.13. Parameters in empirical correction (s_6 and two coefficients in damping function) were fitted against the set of 18 non-covalent complexes, for the B-LYP, BP86 and PBE functionals. It should be noted that applicability of this correction is rather restricted, as it can be utilized only for molecular systems composed only of carbon, hydrogen, nitrogen, oxygen and fluorine atoms.[234]

In 2006 Grimme come up with his second-generation (D2) of damped pairwise atom-atom dispersion correction in DFT (DFT-D2).[235] He reparametrized

the B97[255,256] functional (GGA functional) by explicit inclusion of dispersion correction (B97-D). Similar to the first-generation of correction, this version also includes only the first term from the expansion (n = 6, *cf.* eqn. 2.13), uses global scaling parameter (s_6), empirical vdW radii and exactly same type of damping function. However, C_6^{ij} coefficients are calculated in less empirical way.[235] The B97-D functional brought some improvement compared to the DFT-D scheme. Specifically, we could say that in general B97-D is more robust, as, it does not overestimate (in average) the interaction energies of the H-bonded complexes.[234]

In 2010 Grimme proposed his third-generation of dispersion correction within the DFT (DFT-D3).[236] The DFT-D3 is a significantly revised compared to previous D2 versions. Firstly, it includes an additional terms from dispersion expansion (n = 6, 8; *cf.* eqn. 2.12). Secondly, it utilizes improved damping function along with that the global scaling parameter was removed. Finally, introduces geometry dependent C_n^{ij} coefficients, which are calculated in more rigorous way using Casimir-Polder formula.[262] The last point is the most important, as it incorporates awareness of the chemical environment into the scheme. Moreover, within the D3 approach, the three-body dispersion term (Axilrod-Teller-Muto)[257,258] was also introduced. Let us note here, that the importance of the many-body terms for large systems is not negligible.[259,260]

All these methodological improvements led to the significant refinement compared to the previous DFT-D2 method.[236] It was also shown that the error of the DFT-D3 binding energies (with respect to CCSD(T) values), for small- and medium-size complexes at van der Waals distances, is roughly about 10%.[261] Minor improvement of results was achieved upon using Becke-Johnson damping[263–265] function, which is now a recommended damping function within the DFT-D3 scheme.[266] Many others studies revealed that DFT-D3 provide quit accurate results for different types of noncovalent complexes.[171,173,176]

2.2.2.3.3 Jurecka's Correction

In 2007 Jurecka *et al.*[237] suggested very similar pair-wise atom-atom empirical correction as Grimme in 2004 (DFT-D). Jurecka's scheme utilizes also Fermi type damping function, however in slightly different form, where the vdW radii are scaled instead of the C6 coefficients (for more details see Jurecka's work[237]). This in fact removed the main disadvantage of Grimme's DFT-D scheme: the unphysical scaling of the C_6^{ij} coefficients (*cf.* eqn. 2.13) at long ranges.[237] Grimme adopted this type of damping in his DFT-D3 method.

2.2.2.3.4 Other Correction Schemes

Becke and Johnson proposed semiclassical ("heuristic") model of exchange hole dipole moment (XDM), in which instantaneous dipole moment of (electron and its exchange hole) generate dispersion interaction between nonoverlapping systems.[263,264] This method includes in the correction all three term from dispersion expansion (n=6, 8, 10). Further, use simple rational damping function, which employees effective vdW radii and utilize density dependent C_n^{ij} coefficients, which are derived based on XDM model.[263–265] It was shown that average error of binding energies determined by this method is in range 10-20%.[267]

The method of Tkatchenko and Scheffler (labeled vdW-TS),[268] similar to XDM and Grimme's D3 schemes, is capable of reflecting the bonding environment by utilizing the density-dependent C_6^{ij} coefficients as well as the density-dependent damping function. However, the vdW-TS scheme restricts itself to the first term from dispersion expansion (*cf.* eqn. 2.12) contrary to either XDM or DFT-D3 methods. Further, it utilizes Fermi type damping function, same as Jurecka's scheme.[237] The performance of the vdW-TS scheme for binding energies of the S22 data set, coupled with PBE exchange-correlation functional, was found to be equally good or better than the Grimme's DFT-

D3 scheme.[269]

In 2009 Sato and Nakai introduced their Local Response Dispersion (LRD) method.[270] LRD method, sum all three terms from the expansion (i.e. up to n = 10, *cf.* eqn. 2.12) and uses the density-dependence and anisotropy of C_n^{ij} coefficients Moreover, applies system- and n-order dependence of the damping function. The performance of the LRD method in combination with long-range correction scheme for generalized gradient approximationed,[271] for binding energies of the S22 data set, was found to be excellent.[236]

2.2.2.4 Semiempirical Methods

Semiempirical Quantum Mechanical (SQM) methods were originally developed to reproduce thermochemical properties. They were parametrized toward covalent bonding, therefore, are in principle able to properly describe quantum effects. As SQM methods are computationally cheap they can be used for description of extended systems, for which the ab initio methods are infeasible. Nowadays, they are used also for description of noncovalent interactions. Since, they were not parametrized against noncovalent complexes, their applicability in this direction is more problematic. Consequently, development of SQM methods was oriented toward elimination of this shortcoming.

Performance of AM1,[272] and PM3[273,274] methods on H-bonded complexes was improved, compared with the original MNDO[275,276] method. This was achieved by the introduction of additional core-core term together with the parametrization against H-bonded complexes. First attempt to cover the dispersion in the SQM methods explicitly through empirical R^{-6}, R^{-8}, R^{-10} terms were done by Martin and Clark.[277] Since than many different SQM methods augmented with the empirical corrections for the repulsion and dispersion energy appeared (e.g. OM1,[278] OM2,[279] OM3,[280] etc.), but none

of them is accurate enough for different applications in noncovalent complexes. Surprisingly, all SQM methods provide surprisingly good geometries for all types of complexes, however, the interaction energies for dispersion-bounded and H-bonded complexes are substantially underestimated.[281]

PM6 is one of the most accurate SQM methods.[282] It is a NDDO-based method improved by the Viotyuk's core-core diatomic interaction term and Thiel's d orbital approximation.[278–280] PM6 method in combination with MOZYME algorithm[283] can be applied for large molecular systems (up to several thousands of atoms). Despite all these advantages, the PM6 method (like other SQM methods) still lacks the ability to accurately describe noncovalent interactions. Specifically, interactions driven by dispersion energy and hydrogen bonds.

To improve the performance of the PM6 method up to the edge of the chemical accuracy (1 kcal.mol^{-1}), three modifications, all developed in our laboratory, were introduced: (i) an empirical dispersion term, which improves the description of complexes stabilized by the dispersion ("D"),[236,237,284,286] (ii) an additional electrostatic term, which improves the description of hydrogen-bonded complexes ("H"),[284–286,288] and (iii) an electrostatic term, which improves the description of halogen-bonded complexes ("X").[289] Because of continues improvements each of this modifications has appeared in few version. All the resulting methods were named in following way: PM6-DdHh(X).[284,285] Where, lowercase letters d and h specify the version (generation) of the particular modification; e.g. PM6-D3H4 stands for PM6 method augmented with third generation of dispersion term (d=3, Grimme's D3 correction)[236] and fourth generation of hydrogen bond correction (h=4).[286] The first PM6-DH method provides very good estimates of interaction energies for dispersion-bound complexes, but the description of H-bonded complexes was still not completely satisfactory.[284] Subsequently, the PM6-DH2 method which includes second generation of the hydrogen bond correction with torsional dependence, was introduced. Introduction of the enhanced

hydrogen bond correction term resulted in overall higher accuracy, especially for hydrogen bonded complexes. Specifically, the mean unsigned error for the whole S26[168] data set was reduced from 0.54 (0.67) kcal.mol^{-1} in the case of PM6-DH to 0.36 (0.19) kcal.mol^{-1} in the case of PM6-DH2 (numbers in parenthesis correspond to subset of hydrogen bonded complexes from S26 set).[285] In 2013, the PM7 method was introduced by Stewart.[287] Parametrization of the intermolecular interaction term for PM7 was modified with accord to ideas explored in PM6-DH2,[285] PM6-DH+,[288] and PM6-D3H4 studies. The significant improvement over PM6 method can be seen for heats of formation of organic solids, geometries and barrier heights for simple reactions of the type catalyzed by enzymes.[287] However, considering accuracy in description of the noncovalent interactions, the PM7 method does not outperform PM6 methods with corrections e.g. PM6-DH2(X) or PM6-D3H4(X).[290] So far, the best performing SQM methods with respect to noncovalent interactions is the PM6-D3H4(X).[286,289] The overall accuracy of the PM6-D3H4(X) method, tested toward the 13 noncovalent data sets, was shown to be higher than that of any other SQM method from PM-x family.[290]

2.2.3 Intermolecular Perturbation Theory

The first quantum mechanical theory of intermolecular interactions was developed in 1930s by London and co-workers.[291–294] It is based on standard Rayleigh-Schrodinger (RS) perturbation theory, where unperturbed Hamiltonian is defined as sum of isolated (noniteracting) Hamiltonians and the perturbation is intermolecular potential. In particular, RS perturbation equations correspond to the Hamiltonian partitioning H = H^0 + V, where H is total nonrelativistic Coulomb Hamiltonian of the dimer, $H^0 = H^0_A + H^0_B$ is the sum of monomer Hamiltonians and the perturbation V = H - H^0 consists of all the intermonomer Coulomb interactions i.e. Coulomb interaction between particles (electrons and nuclei) of monomer A and monomer B. In London's method the exact intermolecular potential V is represented by its multipole expansion.[295,296] By definition this expansion of the interaction

energy is convergent only in small part of configuration space when electrons are localized on monomers. Thus London method is valid only asymptotically for large intermolecular distances[297] and its multipole expansion is divergent for finite intermolecular distances.[298–300]

When the exact form of intermolecular potential V is used instead of its multipole expansion, resulting expansion is referred to as polarization approximation.[301] The extra component which is included in polarization expansion compared with the London method originate is the penetration effect. Inclusion of the penetration effects removes divergence character of the multipole expansion i.e. all the energy terms in polarization theory have well defined values for finite intermolecular distances. However, despite significant improvement with respect to London's theory the intermolecular potential obtained from polarization theory is still not even qualitatively correct. Specifically, the polarization theory within finite order is not able to account for electron exchange effects resulting from tunneling of electrons between interacting molecules, even for the systems like H_2^+ or H_2.[302] This is due to the fact that the unperturbed (zero-order) wave function Φ^0_{AB} is defined as product of the ground-state wave functions of isolated subsystems H^0_A and H^0_B: $\Phi^0_{AB} = \Psi^0_A \Psi^0_B$.

Thus, it is clear that wave function in such a form is not antisymmetric with respect to intermolecular electron permutations and thus violates Pauli exclusion principle. Specifically, polarization theory converges for the case of the ground state helium dimer,[303] however, it may converge to the energy which is forbidden by the Pauli exclusion principle.[304] Further, divergence problems of polarization theory shows up when complexes, where at least one of the monomer carries more than two electrons. This divergence can be viewed as a consequence of the fact that physical ground state of the complex is submerged in the continuum of the Pauli-forbidden states, which appear after Penning-type ionization in the physically inaccessible part of the Hilbert space.[305,306] It was shown by Patkowski *et al.* that the Pauli-forbidden continuum of the Li$\cdots$H complex (the simplest system for which

the physical ground state is submerged in the Pauli-forbidden continuum) starts 0.6 Hartree below the ground state energy of H^0.[307] Such a drastic impact of the perturbation V on the unperturbed Hamiltonian H^0 demonstrate that V can not be considered as small perturbation; consequently convergence problem of the perturbation expansion can be expected.

Many perturbation schemes were proposed in order to reproduce correctly the intermolecular potential i.e. account for the polarization as well as the exchange part of the interaction. All these schemes rely on the observation that the symmetry projected n^{th}-order correction of wavefunction $A\Psi^n_{AB}$ is always better approximation to the wavefunction Ψ_{AB} than just Ψ^n_{AB} i.e. $\|A\Psi^n_{AB} - \Psi_{AB}\| \leq \|\Psi^n_{AB} - \Psi_{AB}\|$, where $\|\cdot\|$ is the L^2 norm in the Hilbert space. Thus, one can expect that introducing the symmetry projections into the energy and wavefunction expansions will lead to enhanced convergence of the perturbation expansion. The perturbation schemes, which apply symmetry projection, are referred to as the symmetry-adapted perturbation theories (SAPT).

The SAPT schemes can be divided into two main group. The first group, called "weak symmetry forcing", contains for example: symmetrized Rayleigh-Schrodinger,[308] Murrell-Shaw-Musher-Amos[309,310] and Hirschfelder-Silbey schemes.[311] These methods diverge in the presence of Pauli-forbidden continuum.[307] Nevertheless, these expansions provide quit accurate interaction energies in low order and very accurate ones if the perturbation corrections are summed up until they start to grow in absolute value.[307,312]

The second group, called "strong symmetry forcing", includes: Eisenschitz-London-Hirschfelder-van der Avoird[293,313,314] or Amos-Musher[315] scheme. In these scheme the unphysical component of the wave function is projected out, so these expansions may converge despite the existence of Pauli-forbidden continuum. Unfortunately, these methods suffer from different problem, unlike the methods from the first group. In finite order they are unable to re-

cover even the first terms in the $1/R^n$ power expansion of the interaction potential i.e. have wrong asymptotic behavior.[297] Jeziorski-Kolos expansions[316] also belongs to the second group of schemes, however predicts correctly a few lowest order term from the $1/R^n$ expansion. This partial asymptotic correctness improves convergence for the Jeziorski-Kolos scheme over Eisenschitz-London-Hirschfelder-van der Avoird and Amos-Musher schemes in low orders. The Jeziorski-Kolos scheme can be modified in order to increase number of correctly recovered asymptotic terms, by simple enhancement of the wave function form which enters the Bloch equation.[312] However, this modification do not lead to improved convergence in regions of the van der Waals minimum. Therefore, new SAPT expansion which effectively copes with both problems of aforementioned SAPT schemes (convergence in the presence of Pauli-forbidden continuum as well as correct asymptotic behavior) appeared in literature.[317] This goal is achieved by splitting the attractive Coulomb terms in intermolecular potential into regular, long-range and singular, short-range parts. For more details concerning this approaches see work by Patkowski *et al.*[317] All above mentioned SAPT schemes were tested/developed for the Li$\cdots$H complex (molecule). It is believed that this molecule mimic to some extend the noncovalent interaction and captures all essential complications affecting the SAPT treatment.[317] Therefore, it is generally assumed that conclusions written on the basis of these model studies are transferable to larger systems. However, applications of these methods to larger systems would require approximate solution of the intramonomer correlation energy problem and development of the many-body versions of the theory. Nowadays, the symmetrized Rayleigh-Schrodinger formalism is used for investigation of the many-electron systems, mainly because implementation of this type of method for many-body system is significantly less complex than that of any from the "strong symmetry forcing" methods (e.g. Hirschfelder-Silbey method).

The interaction energy is within the SAPT method defined as the sum of the physically meaningful terms. By definition it is free from basis set superposition error.[318] The electrostatic, induction, dispersion and exchange

interaction energy components, which arise from low orders of perturbation theory, provide insight into origin of the interaction and help to interpret the interaction based on the monomer properties.

Despite all the problems such as considerable formal complexity, difficulties concerning convergence or wrong asymptotic behavior, the intermolecular potential obtained from SAPT are fairly accurate. Hence, nowadays the SAPT methods represent in many cases an attractive alternative to the conventional variational methods.

2.2.3.1 Density Function Theory Symmetry Adapter Perturbation Theory

One of the most essential practical drawbacks of the many-body SAPT method is its extremely high computational cost. In particular, the many-body SAPT method contains terms which formally scale like N^7, where N is a measure of the size of the molecular system. What is incidentally same scaling possessed by the CCSD(T) method. To be more precise, the second-order exchange-dispersion component of the interaction energy (see further) includes terms scaling like $n_{occ}^3 n_{virt}^4$ and $n_{occ}^4 n_{virt}^3$.[319] Therefore, the many-body SAPT method can not be used for investigation of interactions in larger molecular clusters. Thus, method such as Density Functional Theory SAPT (DFT-SAPT) have been developed.[320–324]
The monomer correlation energy is within the DFT-SAPT method treated through less expensive DFT method, contrary to regular many-body SAPT. The scaling of the DFT-SAPT method is N^6 with respect to system size.[319] However, this enhanced scaling properties of the DFT-SAPT compare to the many-body SAPT needs to be improved further, in order to develop method which is computationally feasible for investigation of the extended complexes. This can be achieved through density fitting (also called resolution of identity) procedure. Resulting method is abbreviated as DF-DFT-SAPT and it scales as N^5 with the system size.[319] For more technical details about this method see work by Hesselmann *et al.*[319]

The DFT-SAPT interaction energy is given as the sum of the first- (E_1) and second-order (E_2) perturbation energy terms and δHF energy terms (see eqn. 2.14)

$$E_{int}^{DFT-SAPT} = E_1^{Pol} + E_1^{Ex} + E_2^{I} + E_2^{Ex-I} + E_2^{D} + E_2^{Ex-D} + \delta HF, \quad (2.14)$$

where E_1^{Pol} first-order polarization (Coulomb) term, E_1^{Ex} is first-order exchange term, E_2^{Ex-I} and E_2^{Ex-D} are second-order exchange-repulsion terms, E_2^{I} is second-order induction term, E_2^{D} is second-order dispersion term and δHF is Hartree-Fock correction term.

The first-order Coulomb energy (E_1^{Pol}), second-order induction (E_2^{I}) and dispersion (E_2^{D}) terms from eqn. 2.14 could be in principle calculated exactly.[321–323] The first-order Coulomb energy is determined from following formula $E=\langle\Psi_A^0\Psi_B^0|\hat{V}_{AB}|\Psi_A^0\Psi_B^0\rangle$, where $\hat{V}_{AB}$ is intermolecular Coulomb operator and Ψ_A^0, Ψ_B^0 are wavefunctions of unperturbed monomers. The first-order polarization energy has a very transparent physical interpretation. It simply represents the energy of the Coulombic interaction of the monomers' charge distributions. For this reason it is referred to as the electrostatic energy (see further). At large intermonomer distances R the electrostatic energy can be represented as a sum of classical electrical interactions between the permanent multipole moments of unperturbed monomers (referred later as long-range electrostatic energy). However, one should emphasize that the electrostatic energy contains also important short-range terms due to the mutual penetration (charge overlap) of monomers' electron clouds. Further, the second-order induction energy term is obtained from a solution of the coupled-perturbed Kohn-Sham equations and the second-order dispersion energy from a solution of the frequency-dependent coupled-perturbed Kohn-Sham equations, also called time-dependent density functional theory. Thus, it would be possible to calculate exactly the first and second order terms in the complete basis set limit, provided that the exact xc potential and the exact frequency-dependent xc kernel would be utilized. But, one should bear in mind that exact potential and kernel are not known in general. Regarding exchange components of the interaction

energy. These components do not share the feature of "potential exactness" and can only be approximated within the DFT-SAPT theory. Despite this inconsistency, Hesselmann's and Jansen's studies revealed that all the approximations within DFT-SAPT method work very well,[321–323] as long as balanced xc functional with asymptotic correction is employed (e.g. PBE0AC,[321,325–328] LPBE0AC).[319,329,330] The LPBE0AC and PBE0AC functionals were shown to provide accurate estimations of interaction energies for various types of noncovalent interactions.[319,321,332–348]

However, one should bare in mind that these functionals in combination with aug-cc-pVDZ basis set were shown to give accurate first-order terms as well as induction and exchange terms,[320–322] while dispersion contribution is underestimated approximately by 10-20%.[323]

Further progress in speeding up the DFT-SAPT calculation was introduced by Hesselmann, who employed the empirical form for the dispersion component of the interaction energy in a similar fashion as Grimme.[331] The empirical correction was designed to mimic the effects of both the E_2^D as well as the E_2^{Ex-D} terms. Further, the parameters included in the correction were fitted toward accurate CCSD(T)/CBS interaction energies for the S22 data set.

2.3 Atomic Charge

Atomic charge is not a physical observable and it is well known that there is no unique way how to assign a charge to the atom in the molecule. The intrinsic drawback of the atomic charge concept is the neglect of the anisotropy of the electron density around atom in molecule, i.e. it automatically assumes spherically-symmetrical charge distribution around atom. Consequently phenomena as halogen bonding can not be explain by atomic charge approach since covalently bonded halogens carry mostly negative charge. Despite all these facts the concept of atomic charges is commonly used among chemists

mainly due to its easy interpretation.

There are many approaches which can be used for derivation of the atomic charges. They can be divided in the following groups

- methods based on the partitioning of the wave function
- methods based on the partitioning of the electron density
- methods based on the reproduction of the electrostatic potential

In the following lines only some of the approaches which are based on partitioning of the wave function will be discussed in more details.

The most commonly used method within wave function partitioning methods is the Mulliken (Gross) Population Analysis.[349] The main entity within the Mulliken population analysis theory is the population matrix $\mathbf{P}$. The element of the $\mathbf{P}$, $\mathbf{P}_{\mu\nu}$, is for a closed shell system defined by eqn. 2.15

$$\mathbf{P}_{\mu\nu} = \mathbf{D}_{\mu\nu}\mathbf{S}_{\mu\nu}, \tag{2.15}$$

where $\mathbf{D}_{\mu\nu}$ is the element of the density matrix, $\mathbf{S}_{\mu\nu}$ is the element of the overlap matrix. The element of the density matrix and the overlap matrix are defined by eqns. 2.16 and 2.17, respectively.

$$\mathbf{D}_{\mu\nu} = 2\sum c_{\mu i}c^*_{\nu i} \tag{2.16}$$

$$\mathbf{S}_{\mu\nu} = \int \mu\nu \tag{2.17}$$

The Gross orbital product for atomic orbital ν, GOP_ν, is defined as the sum of all terms of the $\mathbf{P}_{\mu\nu}$ over μ. The sum of all the Gross orbital products equals to the total number of electrons. The Gross atom population GAP_A, is defined as the sum of the GOP_ν over all orbitals ν which belong to an atom

A. Finally, the Mulliken atomic charge assigned to the atom A, q_A, is defined as the difference between the atomic number Z_A, and the Gross atom population: $q_A = Z_A$ - GAP_A. Mulliken population analysis is easy to implement and requires negligible computational effort. This approach, however, suffers from two serious drawbacks. First one is an equal division of the off-diagonal terms in population matrix (i.e. an overlap populations) between two basis function (*cf.* eqn. 2.15). This drawback can be reduced by dividing the overlap population between the corresponding orbital population in the ratio of later.[350] This modified approach of the overlap population splitting, although only partially, incorporates the electronegativity preference into the population analysis. Second deficiency concerns the direct dependence of the Mulliken charges on the construction of the basis set. This dependence is due to the fact that partitioning scheme divides molecular wave function among atomic orbitals, each of which is fully assigned to an atom at which it is centered. So, one could imagine that employing population analysis for the molecule where the basis set is constructed in a way that all basis functions are centered only on the one atom from molecule would actually lead to no electron redistribution. Therefore, Mulliken charges are in principle ill defined and use different types of the basis sets can lead to drastically different results.[351] Despite all the difficulties, atomic charges obtained from Mulliken population analysis can be used for assessing qualitative trends, provided that the same type of basis set was utilized for all atoms.

Another method which is also based on splitting of the wave function is the Natural Population Analysis (NPA).[352,353] This method utilize the Natural Atomic Orbitals (NAOs) as a basis for determining the natural atomic charges. The NAOs are obtained after rather tedious process. Firstly, natural atomic "preorbitals" are obtained by diagonalizing of the individual atomic blocks of the density matrix, which were previously partitioned into atomic blocks. Because of this partitioning of the density matrix, it is clear, that "preorbitals" are orthogonal to each other only within same atomic block. Further, "preorbitals" are divided into two groups: strongly occupied orbitals and weakly occupied orbitals. Subsequently, both groups are orthogonalized

in three steps:

- strongly occupied pre-NAOs from each atomic block are made orthogonal to all the strongly occupied pre-NAOs from other atomic blocks by an occupancy weighted symmetry orthogonalization procedure,[352,354]

- weakly occupied pre-NAOs are done orthogonal to strongly occupied pre-NAOs in the same blocks by Schmidt orthogonalization

- same as first step but applied to weakly occupied pre-NAOs.

The occupancy weighted symmetry orthogonalization procedure is designed in such a way that it minimizes the occupancy-weighted mean-squared deviations of the final NAOs from the parent nonorthogonal pre-NAOs i.e. obeys maximum-resemblance property. The atomic charges obtained from the NPA are much less basis set dependent than the Mulliken charges.

There exist many other alternatives for obtaining atomic charges which are more intuitive and also more physically relevant than Mulliken or NPA methods. Among such methods, which are routinely used, belong Bader's population analysis[355] or restrained electrostatic potential derived charges.[356] These methods are based on the splitting of the electron density among atoms and on the reproduction of the molecular electrostatic potential, respectively.[355,356]

2.4 Distributed Multipole Analysis

Distributed Multipole Analysis method (DMA) is a tool for the description of the molecular charge distribution by using the local multipoles at a set

of sites within molecule.[357,358] The method itself provides comprehensive view on the nature of the electron distribution, as well as simple way for calculating multipole moments of arbitrary order to high accuracy.

The starting point of the DMA method is expansion of the electron density ($\rho(\mathbf{r})$) in terms of the primitive Gaussian basis functions $\chi_i(\mathbf{r})$, which are defined by eqn. 2.18

$$\chi_i(\mathbf{r}) = N_i x_i^{a_i} y_i^{b_i} z_i^{c_i} exp[\zeta(\mathbf{r})^2], \tag{2.18}$$

where $\mathbf{r}_i$ = **r** - $\mathbf{A}_i$ is the electron position relative to the center of the primitive Gaussian $\mathbf{A}_i$, ζ_i is the exponent and N_i is the normalization factor. Method takes advantage of the well known property of the Gaussian basis function: "Product of two Gaussian functions each located on the different center (e.g. points **A** and **B**), is Gaussian function located on the third center (e.g. point **C**, which is located at the junction of the points **A** and **B**). Therefore, the method can be applied only to a wave function expressed in the Gaussian form. The key entity for the DMA method, a multipole moment (Q_{lk}) is defined by eqn. 2.19

$$Q_{lk} = \int R_{lk}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r}, \tag{2.19}$$

where R_{lk} is spherical harmonic and $\rho(\mathbf{r})$ is electron density defined by eqn. 2.20

$$\rho(\mathbf{r}) = \sum_{ij} \mathbf{D_{ij}}\chi_i(\mathbf{r})\chi_j(\mathbf{r}), \tag{2.20}$$

where $\mathbf{D}_{ij}$ is density matrix. Thus, the accuracy of the calculated multipole moments depends solely on the quality of the wave function. The multipole moment arising from the population contribution of the particular pair of primitive Gaussian function (χ_i,χ_j) is centered to the point **C** (see above). At this point, algorithm takes advantage of the fact that multipole moments

with respect to a particular origin can be expressed in terms of those with respect to another origin.[359] Because of this feature, all the multipole moments centered at the specific set of points (sites), can be expressed through the multipole moments centered at the different points (e.g. **C** points). Sites are usually located at the position of atom nucleus and middles of the chemical bonds. The crucial step in the DMA algorithm is the division of all the multipole moments among site points. This is done in two different ways. First, if the sum of the exponents of the Gaussian primitives is larger than specified threshold e.g. 4 (i.e. population is sufficiently localized) the particular multipole moment contribution is fully assigned to the nearest site.[360] Second, if the sum of the exponents of the Gaussian primitives is smaller than specified threshold (i.e. population is diffused among several sites) the particular multipole moment contribution is apportioned among few nearest sites using a grid-based analysis.[360] The whole algorithm is very fast and efficient, and gives an excellent representation of the molecular charge distribution.[360]

It should be stressed that the above mentioned algorithm is not unique in two aspects. First one is the choice of origins for the multipole expansions (i.e. sites). Each atom is a natural center for its own contribution, but there is no natural center for bond contributions. The second one arises from the fact that the expansion of the wave function is not unique. The most commonly used type of the wave function is multicenter. So, it is possible to represent a basis function on one center in terms of basis functions on another. This could cause, within the distributed multipole analysis, shifting of the contribution from one atom onto another atom or bond. Both of these deficiencies can be almost entirely eliminated by using the basis set which is reasonably balanced for each atom.

Distributed multipoles provide a much more accurate representation of the charge density than a single-point multipole expansion, even for small molecules.[357] Further, distributed multipole expansions has better convergence than the single-point multipole expansions.[357,358] Therefore, potentials repre-

sented by distributed multipoles, are more suited for calculation of the electrostatic energy than conventional single-site potentials, which converge adequately only for quit large distances.[357]

At the end of this section let us stress that concept such as: atomic charge, distributed multipole analyses, molecular orbitals are quantities arising from mathematical models. These quantities are used by computational chemists as tools for interpretations of specific phenomena, however, they are not physical observables (i.e. they are not real).

Chapter 3

Projects

3.1 L7 Data Set

3.1.1 Introduction

Most of the data sets presented so far in literature share the same restriction: only medium-sized systems (less than ∼30 atoms) are included. The only exception is the Grimme's S12L data set, but in this case the reference interaction energies are obtained indirectly: from experimental binding free energies measured in water environment.[172] It is a tacit assumption that the accuracy of methods parameterized for small complexes is preserved for larger ones. This, however, may not be the case if a method works well near the van der Waals minimum but is deficient for distant interactions because larger molecules have more long-range dispersion terms. Further, the three-body terms become more important for extended complexes. The potential accumulation of errors with increasing system size is not yet fully understood. Thus there is a need to test the accuracy of recent methods on larger systems. In this study, we have provided benchmark data for following large noncovalent complexes: "CBH", the octadecane dimer in stacked parallel conformation (representative of aliphatic dispersion dominated interaction); "GGG", a stacked guanine trimer arranged as in DNA (representative of the aromatic

stacking $\pi\cdots\pi$ dispersion interaction with implicit account for the three-body interaction, the binding energy of one of the outer guanine monomers is evaluated); "C3A", a stacked circumcoronene$\cdots$adenine dimer (representative of strong aromatic dispersion interaction with implicit account for three-body interaction); "C3GC", a stacked circumcoronene and Watson-Crick hydrogen-bonded guanine-cytosine dimer (representative of a strong aromatic dispersion interaction with implicit account for H-bonding-stacking nonadditivity, the binding energy of circumcoronene and guanine-cytosine base pair is calculated); "C2C2PD", a parallel displaced stacked coronene dimer (representative of strong aromatic dispersion interaction); "GCGC", a stacked Watson-Crick H-bonded guanine-cytosine dimers arranged as in DNA (representative of strong aromatic dispersion interaction with implicit three- and four-body interactions, the binding energy of two guanine-cytosine base pairs is evaluated); and "PHE", an amyloid fragment, a trimer of phenylalanine residues in mixed H-bonded-stacked conformation (representative of "mixed-character" interaction with implicit account for many-body interactions. The binding of one of the "outer" residues is evaluated). Structures of all complexes are shown in Figure 3.1.

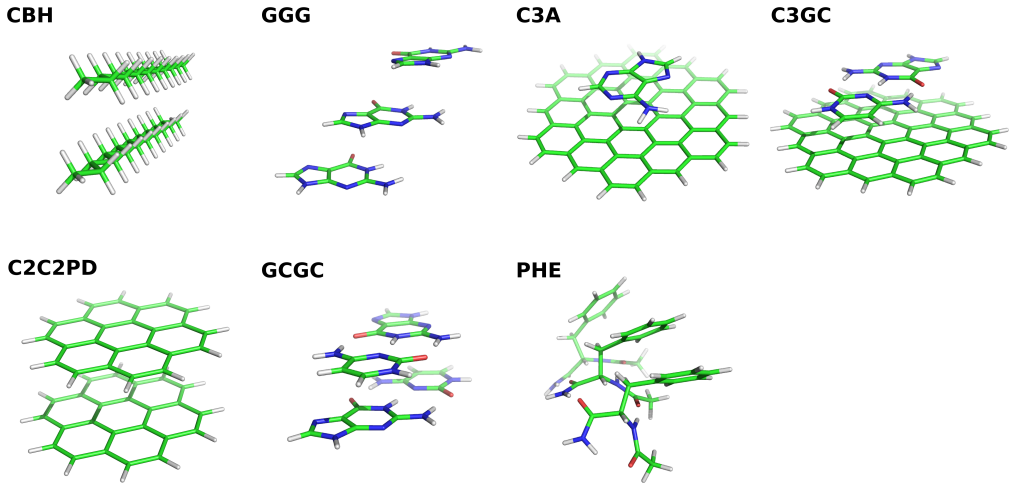

Figure 3.1: Structures of the complexes from the L7 data set.

Their size ranges from 48 to 112 atoms and they are intentionally selected to be mostly dispersion dominated. The motivation is simple: it is to as-

semble a set of noncovalent complexes, the accurate description of which is a challenge for the contemporary computational chemistry. We believe that the complexes included in the data set are representative of the most important motifs dominated by dispersion in biological chemistry.
Further, we have tested the performance of Wave Function Theory (WFT), DFT and SQM methods on above mentioned set of seven large molecular complexes. To be more specific, the group of WFT methods is represented by: SCS-MP2, SCS(MI)-MP2, MP2C, MP2.5 and MP2.X methods.[220,223,230,232,233] The tested methods from the DFT family were: B3-LYP-D3, B-LYP-D3, TPSS-D3, PW6B95-D3, M06-2X-D3 and TPSS-D;[236,237] following semiempirical methods were also examined: PM6-D, PM6-DH2, PM6-D3H4 and SCC-DF-TB-D.[284–286,361,362]

3.1.2 Strategy of Calculation

3.1.2.1 Structures and Geometries

The structures of CBH, C3A, C3GC, and PHE systems were determined at the TPSS-D/TZVP level[237] with no constrains. The structures of GCGC and C2C2PD complexes were taken from Pitonak *et al.*[205] and Janowski *et al.*,[363] respectively. The GCGC geometry was taken from crystal X-ray data, and the C2C2PD structure was optimized at the QCISD(T) level of theory. The GGG geometry was extracted from the 1ZF9 PDB structure[364] with hydrogen atoms added manually. Subsequently, the coordinates of the hydrogen atoms only were optimized at the TPSS-D/TZVP level.[237]

3.1.2.2 Stabilization Energies

The reference QCISD(T)/CBS interaction energies were approximated by utilizing the eqn. 2.8. The use of augmented basis set for extended cluster fragments (or monomers) is problematic, if not impossible. The reason is numerical instability caused by overcompleteness of the atomic basis

set. The overcompleteness problem can be overcome in the most straightforward way by excluding the linear dependent basis functions from the basis set. Doing so, the numerical problems usually disappear, but the error with respect to the complete basis set is unpredictable and discontinuities may occur on the potential energy surface. One of the methods of eliminating this problem, as used in this work, is to scale the ΔE(MP2/CBS) term obtained from extrapolation using non augmented Dunning's[365–367] cc-pVXZ (XZ) basis sets. Specifically, the first estimation of the MP2/CBS was obtained by the Helgaker two point extrapolation scheme in the cc-pVDZ and cc-pVTZ basis sets (see eqn. 2.1 and 2.2). Secondly, energy obtained in the first step was subsequently scaled, because of the lack of augmented functions in the extrapolation procedure. The particular scaling coefficient was considered for each complex individually and was defined as the ration between MP2/CBS values obtained by extrapolation from Dunning's augmented and non augmented basis sets of double- and triple-zeta quality. The two set of calculations, in augmented and non augmented basis sets, were performed for the following group of four medium-size binary complexes: coronene$\cdots$adenine, guanine dimer, phenylalanine-fragment dimer and octane dimer, which mimic the complexes in L7 data set. For more details about the scaling procedure, see Methods part in Attachment **A**. The ΔQCISD(T) correction term was evaluated in rather small 6-31G*(0.25) basis set, because of extreme computational demand. The exponents of the diffuse d functions, used in this modified 6-31G* basis set, were changed from their original value of 0.8 to 0.25.[368,369] This basis set and the 6-31G**(0.25,0.15) basis were designed for the treatment of noncovalent interactions. They have been extensively validated for hydrogen bonded and stacked DNA base pairs in Hobza's group,[370] and surprisingly, good performance was demonstrated also for more diverse data sets containing noncovalent interactions.[233,371]

The MP2C, MP2.5 and MP2.X methods were evaluated at the CBS limit, which was approximated by the eqn. 2.11. In the case of spin scaled variants of MP2 (SCS-MP2 and SCS(MI)-MP2) the CBS limit was approximated by utilizing the extrapolation scheme specified in the DiStasio *et al.*

paper.[223] The DFT methods were evaluated in the best possible basis set, according to the original papers. Specifically, all the DFT functionals with the Grimme's D3 correction were tested in the def2-QZVP in combination with Becke-Johnson and Zero damping. Jurecka's TPSS-D method was utilized in the TZVP and M06-2X in the def2-QZVP basis set. Regarding SQM methods, these calculations were done in subminimal basis set.

3.1.2.3 Statistical Evaluation

The accuracy of all tested methods with respect to the reference QCISD(T) energies was quantified with following statistical indicators: root mean square deviation (RMSD), mean unsigned error (MUE), mean signed error (MSE), and maximum unsigned error (MAX). RMSD and MUE characterize the overall accuracy of a method. MSE provides information about systematic errors. MAX identifies the worst described entity and thus it measures the robustness of the method. Table 3.1 lists all four statistical indicators, respectively, at different level of theory.

Statistical measures are calculated with respect to reference interaction energies. The total binding energies for the noncovalent complexes investigated vary significantly, from 2 to 32 kcal.mol^{-1} (*cf.* Table 2 in Attachment **A**). Hence, the relative (percentage) values marked with prefix "r", defined as 100 x (ΔE_{method} - $\Delta E_{reference}$)/$\Delta E_{reference}$ are more appropriate for the comparison across the whole data set.

3.1.3 Results and Discussion

3.1.3.1 Scaling of the MP2/CBS

The MP2/CBS interaction energies for four medium-size model complexes, obtained from extrapolating the augmented and non augmented cc-pVXZ, X = D, T basis sets to the complete basis set limit are listed in Table 1 of Attachment **A**. The extrapolated values of stabilization energies follow the expectations that the former are larger than the later ones, thus leading

method/basis set//statis. measure	RMSD	MUE	MSE	MAX
QCISD/CBS	3.50	2.86	2.86	6.46
MP2.X/CBS	2.34	1.87	1.87	4.61
MP2.5/CBS	0.79	0.61	0.61	1.56
MP2C/CBS	1.83	1.45	1.38	3.48
MP3/CBS	10.19	7.79	7.79	17.74
MP2/CBS	8.78	6.57	-6.57	14.77
SCS-MP2/CBS	2.37	2.12	0.58	3.18
SCS(MI)-MP2/CBS	4.20	3.00	-2.47	7.35
B3-LYP-D3/def2-QZVP	0.95	0.70	-0.24	1.90
B-LYP-D3/def2-QZVP	1.60	1.39	-1.16	2.36
TPSS-D3/def2-QZVP	1.87	1.66	-1.29	3.16
PW6D95-D3/def2-QZVP	2.15	1.83	-1.83	4.42
M06-2X-D3/def2-QZVP	2.17	1.72	1.72	3.81
M06-2X/def2-QZVP	5.33	4.81	4.81	7.57
TPSS-D/TZVP	2.95	2.34	1.36	5.67
PM6-D3H4/SMB	3.92	3.16	1.28	6.83
PM6-DH2/SMB	3.35	2.20	-0.64	8.18
PM6-D/SMB	3.00	2.27	0.04	6.62
SCC-DF-TB-D/SMB	4.33	3.78	2.81	6.84

Table 3.1: Set of statistical measures (in kcal.mol^{-1}) for the L7 complexes.

to the scaling coefficients above one. The variation of the scaling coefficients, as already noted, is surprisingly small, in orders of only few percent (1-5%). This is most likely the consequence of the basis set saturation already at the non augmented level due to the extended size of the investigated (medium-size) complexes.

3.1.3.2 Interaction Energies

3.1.3.2.1 MP2, MP3, MP2.5 and MP2C

The performance of the plain MP2 and MP3 methods is poor, see the RMSD, MUE, and MAX values in Table 3.1. Typical feature of the MP3 method is underestimation of the aromatic dispersion interaction (MSE about 7.8 kcal.mol^{-1}, *cf.* Table 3.1), contrary to MP2 (MSE about -6.6 kcal.mol^{-1}). However, both MP2 and MP3 describe hydrogen bonded complexes very well. The error cancellation between MP2 and MP3 is responsible for substantially higher accuracy of MP2.5, with RMSD of only about 0.8 kcal.mol^{-1} (4%), *cf.* Table 3.1 and Figure 3.2. A statistical evaluation of the MP2.X performance gives 2.3, 1.9, and 4.6 kcal.mol^{-1} and 13% for RMSD, MUE, MAX, and rRMSD, respectively (*cf.* Table 3.1), which is quite good. MP2.X outperforms both scaled-MP2 variants. However, its performance is worse than MP2.5, in spite of the optimization of the mixing parameter (*cf.* Table 3.1). This is most likely a consequence of parameterization toward substantially smaller molecular clusters in the S66 data set. The MP2C method shows a real improvement over MP2. In terms of the relative errors, it is the second best performer among the WFT methods with an rRMSD of about 8% (1.8 kcal.mol^{-1}, *cf.* Table 3.1 and Figure 3.2) and an rMAX of 14% (3.5 kcal.mol^{-1}) for the coronene dimer.

3.1.3.2.2 Scaled MP2

Introduction of empirical scaling of the spin components in the MP2 substantially improves the description for π stacking. RMSD and relative RMSD val-

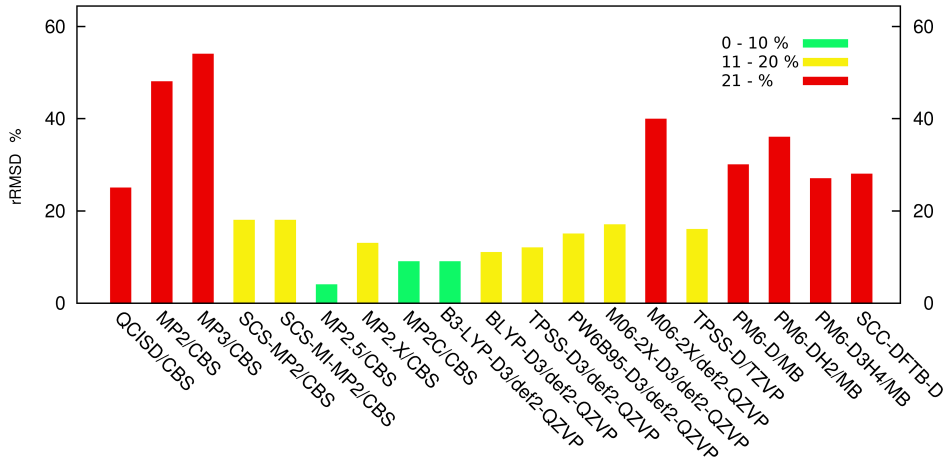

Figure 3.2: Relative RMSD (in %) for all investigated methods.

ues for MP2, SCS-MP2, and SCS(MI)-MP2 are 8.8 (48%), 2.4 (18%), and 4.2 (18%), respectively (*cf.* Table 3.1 and Figure 3.2). However, the performance for the C2C2PD complex is still inadequate, the error of - 7.4 kcal.mol^{-1} (-30%) at the SCS(MI)-MP2 level is enormous. According to the results in Table 3.1 SCS(MI)-MP2 performs worse than the original SCS-MP2. This is surprising, considering that SCS(MI)-MP2 was, unlike SCS-MP2, parameterized toward the best performance on interaction energies. Comparing the performance of both methods for dispersion dominated complexes in the S66 set, the opposite conclusion is obtained. The relative RMSD of SCS(MI)-MP2 and SCS-MP2 are 18% and 26%, respectively,[171] which indicates that the size of the investigated molecular cluster has a significant effect on overall accuracy of the method. It is known that SCS-MP2 underestimates the bonding between aliphatic species.[221,372] The relative error, roughly 30% for the aliphatic CBH complex, is among the largest from all the methods within the group (*cf.* Table 3 and Figure 3 in Attachment **A**).

3.1.3.2.3 DFT

The performance of the group of functionals augmented with the Grimme's D3 correction[236] utilizing the Becke-Johnson (B-J) damping[266] is better in comparison with the Zero damping. Hence, we will discuss almost exclusively

the DFT-D3 methods using the B-J damping. The M06-2X-D3 method will be discussed in combination with Zero damping, because there are no B-J damping parameters for the D3 correction. The best performing functional is the B3-LYP followed by the B-LYP, TPSS, PW6B95, and M06-2X. Corresponding RMSD in kcal.mol^{-1} (rRMSD in %) are 1.0 (9), 1.6 (11), 1.9 (12), 2.2 (15), and 2.2 (17) (*cf.* Table 3.1 and Figure 3.2). The M06-2X-D3 method in combination with zero damping procedure leads to substantially smaller error than the plain M06-2X functional. All the statistical indicators are reduced, two to three times, after applying D3 correction (*cf.* Table 3.1). This result serves as the proof that functionals which were fitted to reproduce dispersion near van der Waals minimum (for example M06-2X) are not able to cover long-range dispersion interaction. This feature was demonstrated by Grimme.[261]

The TPSS-D/TZVP method of Jurecka *et al.*[237] performs slightly worse than tested density functionals with D3 correction but clearly better than the M06-2X/def2-QZVP method of Zhao and Truhlar, with RMSD of 3.0 kcal.mol^{-1} (16%) and 5.3 kcal.mol^{-1} (40%), *cf.* Table 3.1 and Figure 3.2, respectively. The direct comparison of the TPSS-D/TZVP with the TPSS-D3/def2-QZVP reveals that the latter method is slightly more accurate. However, this comparison is not completely fair because of the unequal quality of the used basis sets. The def2-QZVP basis set is of better quality than TZVP which could be the reason for the higher accuracy.

DFT methods describe H-bonding (or electrostatic dominated interactions in general) fairly well. Empirically corrected DFT-D and DFT-D3 methods share a similar feature. Signed errors for the H-bonded PHE complex (ranging from 0.2 to 2.4 kcal.mol^{-1}, that is 1-9%) are significantly lower than for other complexes. Dispersion dominated interactions are described less accurately. The largest signed errors of Grimme's (DFT-D3) methods are: 4.4 kcal.mol^{-1} (18%) in the case of C2C2PD complex for PW6D95 functional and -2.1 kcal.mol^{-1} (8%) in the case of GCGC complex for B-LYP

functional. Description of aliphatic dispersion in the CBH complex, is even worse: the signed errors range from -2.4 kcal.mol^{-1} (21%) for B-LYP-D3 up to 6.4 kcal.mol^{-1} (57%) for M06-2X (*cf.* Table 3 in Attachment **A**).

3.1.3.2.4 SQM

The PM6 method,[282] corrected for dispersion (PM6-D)[284] gives an RMSD of 3.0 kcal.mol^{-1}, see Table 3.1. Corrections for hydrogen bonding (PM6-DH2, PM6-D3H4)[285,286] do not, however, increase the accuracy further (RMSD of 3.4 and 3.9 kcal.mol^{-1}, respectively). The correction for hydrogen bonding is most likely the source of undesirable errors, and it is responsible for slightly worse correlation with the benchmark data for complexes stabilized by dispersion interaction. On the other hand, errors decrease significantly for the PHE complex upon inclusion of the hydrogen bonding correction: the signed error drops from 3.0 (PM6-D) to 0.9 (PM6-DH2) and 0.3 (PM6-D3H4) kcal.mol^{-1-1}, respectively (*cf.* Table 3 in Attachment **A**). The performance of SCC-DF-TB-D method[361,362] is comparable with that of PM6-D3H4 (*cf.* Table 3.1).[286] The rRMSD and rMAX values for SQM methods vary between 26-36% and 46-74%, respectively. Nevertheless, the absolute errors of the SQM methods are comparable with some of more sophisticated methods such as M06-2X/def2-QZVP, MP2/CBS, MP3/CBS, or SCS(MI)-MP2 (*cf.* Table 3.1).

3.2 Charge-Transfer Complexes

3.2.1 Introduction

A charge-transfer (CT) complex, sometimes also labeled as an electron-donor-acceptor complex, is a complex of two (or more) systems where one is an electron donor (high ionization potential) and the other an electron acceptor (low electron affinity). CT complexes are frequent, there are organic as well as inorganic donors and acceptors, and CT complexes play a role also in the biolog-

ical[373–376] and material sciences.[377–380] Where does the stabilization of these complexes come from? Quantum mechanically, this is described as a resonance between the non-bonded state D$\cdots$A and the dative state D$^+\cdots$A$^-$. The isolated electron donor as well as the electron acceptor are mostly electroneutral, but after electron transfer the donor possesses a partial positive charge while the electron acceptor possesses a partial negative charge. These charges attract each other and the resulting electrostatic attraction represents an important energy contribution. Thus, an important characteristic of CT complexes is the amount of the electron transferred between the donor and acceptor. The electron transfer from the donor to the acceptor is usually small but may become important when the donor's and/or acceptor's abilities improve. The total electron transfer is deduced from the total charge of each subsystem and depends strongly on the type of atomic charges used. The most frequently used Mulliken charges are known to be extremely basis-set dependent,[349] and therefore the use of other charges, like the natural (NPA) ones, is recommended.[352,353] Contradictory to other types of noncovalent complexes (like H-bonded or stacked complexes) little is known about the nature of their stabilization. The question arises whether these complexes are mainly stabilized by CT energy or by combination of different energy terms like electrostatic, dispersion, and CT. Here, the CT7/04 database of Truhlar should be mentioned.[156] Seven small CT complexes covering weak and medium CT complexes (with a stabilization energy between 1 and 11 kcal.mol^{-1}) were investigated at the W1 level (CCSD(T)/CBS level), but no energy decomposition was performed.

The CT energy contribution can be approximated by the second-order perturbation theory analysis of the Fock matrix in the Natural bond orbital (NBO) basis,[352] see eqn. 3.1.

$$E^{CT} \simeq E^{(2)}_{\sigma\sigma*} = -2\frac{\langle\sigma|F|\sigma*\rangle^2}{\epsilon_{\sigma*} - \epsilon_\sigma}, \tag{3.1}$$

where F is the Fock matrix element between the σ and σ^* NBO orbitals, and $\epsilon_{\sigma*}$ and ϵ_σ are the energies of the σ^* and σ NBO orbitals, respectively. A high

degree of stabilization is expected when the energy difference between the σ orbital of the donor and the σ^* orbital of the acceptor is small and simultaneously when the Fock matrix element (reflecting the overlap of the σ and σ^* orbitals) is large. The perturbation E_2^{CT} CT energies are overestimated and cannot be compared with the other energy terms taken e.g. from perturbation theory (see bellow). However, the E_2^{CT} energies can be compared within different CT complexes and provide information on the importance of the CT energy term.

Obtaining CT energy from the DFT-SAPT treatment is far from straightforward. As, the DFT-SAPT energy is defined by eqn. 2.14 and it is evident that a genuine CT energy term is missing. According to the SAPT theory and computational implementation of the method the CT energy is included in second-order induction, exchange-induction and in δHF term. The δHF term, which is determined as the difference between the supermolecular HF interaction energy and the sum of the electrostatic, exchange, induction and exchange-induction energies determined at the HF-SAPT level in dimer basis set. The second-order induction energies cover not only the classical multipole-induced multipole energy, but also the CT energy, as long as dimer basis set is used. This arises from the fact that the sums, in this case, cover not only the occupied and unoccupied orbitals of the same subsystem, but also the occupied orbitals of the one subsystem and the unoccupied orbitals of the partner subsystem. We can thus conclude that the sum of the second-order induction and the nonperturbative δHF terms, labeled in this study as $E^{"CT"}$, represents the upper bound to the CT energy.
Stone and Misquita proposed scheme for obtaining the CT energy within the SAPT approach.[381] This scheme is based on the idea that the difference between the induction energy term calculated in the dimer and monomer basis set is the CT energy. However, this approach is strongly basis set dependent (for more details see ref. 381). The applicability of the DFT-SAPT analysis for CT complexes is questionable, mainly because the perturbation SAPT theory is not well-defined for CT complexes for which the intermolecular Hamiltonian is expected to be non-negligible with respect to unperturbed

Hamiltonian (see section 2.2.3). Consequently, the convergence of the perturbation expansion is expected to be slow. Therefore, it is believed that the SAPT theory up to the second-order can provide reliable data for systems without any or with only "small" electron transfer.

The different character of the perturbation E_2^{CT} energy and SAPT perturbation energies should be mentioned. Whereas the former is based directly on the concept of electron transfer between the donor and acceptor, the latter treats the issue of electron transfer indirectly. From this point of view, it is impossible to compare these energies directly with one another. On the other hand, the E_2^{CT} energies ,as well as CT energies from SAPT, can be compared separately within different complexes, allowing us to estimate the relative importance of the CT energy term.

To shed more light into the character of stabilization for the CT complexes, we investigated in this study the structure and stabilization of small CT complexes of various strengths (from weak through medium to strong) formed by various small electron donors (C_2H_4, C_2H_2, NH_3, NMe_3, HCN, H_2O) and small acceptors (F_2, Cl_2, BH_3, SO_2). The aim of the study is 3-fold: (i) to elucidate the nature of stabilization in CT complexes, that is, to find the dominant energy term, (ii) to test the performance of the MP2 method, and (iii) to investigate the applicability of the DFT-SAPT treatment. The nature and stabilization of the present CT complexes will be compared with the same (benchmark) values determined for H-bonded, stacked and mixed complexes of the S22 data set.[159] Further insight into the nature of the stabilization of CT complexes can be gained by passing from the correlated calculations (MP2 and CCSD(T)) to the Hartree-Fock ones. It must be kept in mind that CT energy (contrary to dispersion energy) is at least partially covered already at the HF level. In the present study, we have investigated small, mostly pseudo linear (contact) CT complexes.

3.2.2 Strategy of Calculation

3.2.2.1 Structures and Geometries

The structure of all of the complexes investigated was determined by the counterpoise-corrected gradient optimization performed at the MP2/cc-pVTZ level. It was shown that these geometries are close to those obtained at the CCSD(T)/CBS level.[382] The isolated subsystems were optimized at the MP2/cc-pVTZ level. All of the complexes studied (and their subsystems) were also optimized at the HF/cc-pVTZ level.

3.2.2.2 Stabilization Energies

The CCSD(T)/CBS stabilization energy was constructed as the sum of the MP2/CBS stabilization energy and the $\Delta\Delta$CCSD(T) correction term. The former energy was extrapolated (Helgaker extrapolation) from the MP2 energies evaluated at the MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ and MP2/aug-cc-pVQZ levels. The ΔCCSD(T) correction term was determined with the aug-cc-pVDZ basis set. In the case of the $NH_3\cdots F_2$ complex, the CCSD(T)/CBS interaction energy was also evaluated by a direct extrapolation of the CCSD(T) energies calculated with the aug-cc-pVDZ and aug-cc-pVTZ and aug-cc-pVTZ and aug-cc-pVQZ basis sets. The deformation energies were systematically not considered.
The DFT-SAPT calculations were performed with the PBE0AC functional and aug-cc-pVTZ basis sets. For the $NH_3\cdots F_2$ complex, the DFT-SAPT calculations were also performed with the smaller aug-cc-pVDZ basis set, and for the same complex the MP2-SAPT calculations with the aug-cc-pVDZ and aug-cc-pVTZ basis sets were also performed.

3.2.2.3 Atomic Charges and Frontier Orbitals

The atomic charges were approximated by the Mulliken and NPA charges and in both cases they were determined at the HF/6-31G*, HF/cc-pVDZ,

and HF/cc-pVTZ levels, respectively. The HOMO and LUMO energies of all of the subsystems were calculated at the HF/cc-pVTZ level of theory.

3.2.3 Results and Discussion

3.2.3.1 Structure and Geometry

The structures of all complexes studied are depicted in Figure 3.3, where selected intermolecular coordinates are also shown. Investigating optimized structures, we found that in the case of BH_3-containing complexes the intermolecular distance becomes very small, almost at the range of covalent bonding. The theoretical estimates of the $B\cdots N$ distance in $BH_3\cdots NH_3$ and $BH_3\cdots NMe_3$ complexes (1.658 and 1.643 Å) agree well with the gas-phase experiments (1.627 and 1.65 Å).[383,384]

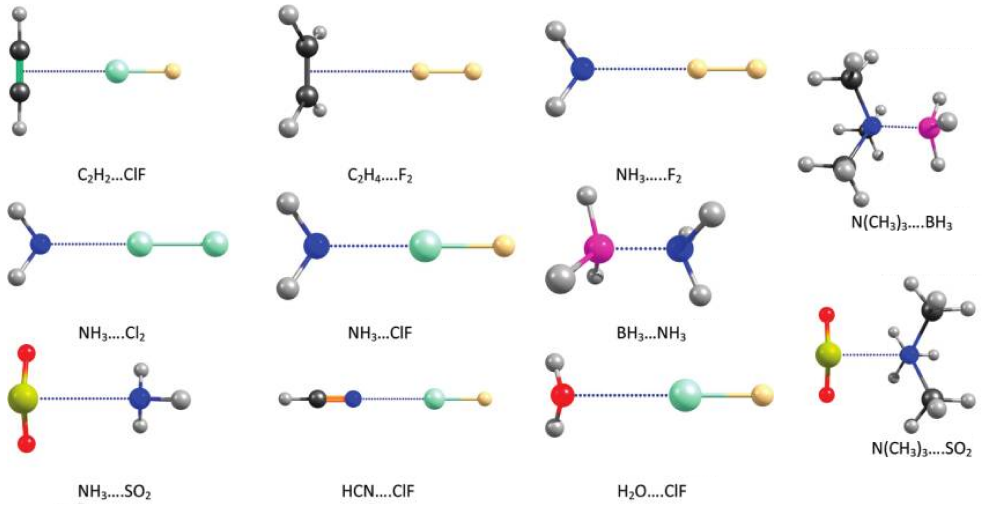

Figure 3.3: MP2/cc-pVTZ optimized structures of all the CT complexes.

3.2.3.2 Charge-Transfer and Frontier Orbitals

Considering the energies of the HOMO orbitals electron donors can be ordered as follows: $H_2O < HCN < NH_3 < C_2H_2 < C_2H_4 < NMe_3$; while the acceptor ability increases in the following order: $F_2 < ClF < BH_3 < Cl_2$

< SO_2, based on the energies of LUMO orbitals.
Table 3.2 lists the amount of the electron transferred from the donor to acceptor calculated at the HF level in the cc-pVTZ basis sets. An extremely large value of CT (more than 0.3 e) was obtained in complexes 10 and 11 with a BH_3 acceptor, and a very large CT (more than 0.1 e) in the $NH_3 \cdots ClF$ and $NMe_3 \cdots SO_2$ complexes. A moderately large CT was calculated in the $NH_3 \cdots Cl_2$ and $H_2O \cdots ClF$ complexes, whereas in all of the remaining complexes it is below 0.05 e. The complexes with CT higher than 0.3 e are characterized by very small intersystem distance.

#	complex	NPA	Mulliken	E_2^{CT}	HF/ cc-pVTZ	MP2/ CBS	CCSD(T)/ CBS
1	$C_2H_4 \cdots F_2$	0.007	0.002	1.80	-0.06	-1.29	-0.95
2	$NH_3 \cdots F_2$	0.014	0.008	2.88	-0.33	-1.74	-1.57
3	$C_2H_2 \cdots ClF$	0.031	0.028	8.24	-1.16	-4.74	-3.82
4	$NH_3 \cdots Cl_2$	0.063	0.071	14.04	-2.23	-5.24	-4.48
5	$HCN \cdots ClF$	0.025	0.028	2.34	-2.71	-5.39	-4.58
6	$H_2O \cdots ClF$	0.037	0.050	9.55	-3.40	-5.31	-4.86
7	$NH_3 \cdots SO_2$	0.030	0.047	7.43	-4.70	-5.40	-5.52
8	$NH_3 \cdots ClF$	0.149	0.142	39.84	-5.29	-12.37	-10.69
9	$NMe_3 \cdots SO_2$	0.185	0.212	58.71	-5.08	-15.74	-14.41
10	$NH_3 \cdots BH_3$	0.378	0.459	-	-35.21	-45.24	-43.91
11	$NMe_3 \cdots BH_3$	0.357	0.510	-	-41.82	-57.17	-54.89

Table 3.2: HF/cc-pVTZ, MP2/CBS and CCSD(T)/CBS interaction energies (in kcal.mol^{-1}); NPA and Mulliken CT (in e) and E_2^{CT} (in kcal.mol^{-1}) for all of the CT complexes evaluated at the HF/cc-pVTZ.

3.2.3.3 Stabilization Energies

The CCSD(T), MP2, and HF stabilization energies are presented in Table 3.2, whereas the DFT-SAPT energies and the corresponding energy contributions are collected in Table 3.3. Investigating the CCSD(T)/CBS stabilization energies, we found that the weakest complexes are those with F_2 as an electron

acceptor (the stabilization energies were below 2 kcal.mol^{-1}), which agrees with the ranking of F_2 as the worst electron acceptor. Moderately strong CT complexes with a stabilization energy between 3 and 11 kcal.mol^{-1} have ClF, Cl_2, and SO_2 (complex with NH_3) as an acceptor. Finally, the complexes with the BH_3 acceptor and also the $NMe_3 \cdots SO_2$ complex are strong CT complexes with a stabilization energy exceeding 14 kcal.mol^{-1}. The stabilization energy of the complexes with the BH_3 acceptor is extremely high, more than 43 kcal.mol^{-1}. The latter complexes are in fact systems with a dative bond. All of the complexes investigated are mostly stronger than similar complexes without a CT energy component. Allow us to mention as an example CT complexes with NH_3 as a donor. Out of the five CT complexes studied, only $NH_3 \cdots F_2$ can be labeled as weak, all the others as moderately strong or strong. The CCSD(T)/CBS stabilization energy of the ammonia dimer is 3.17 kcal.mol^{-1},[159] which is less than that of the four strongest CT complexes with a NH_3 donor. Similarly, the CCSD(T)/CBS stabilization energy for ethane dimer is 1.53 kcal.mol^{-1},[159] which is less than one-half of stabilization energy of the present acetylene$\cdots$ClF complex (3.82 kcal.mol^{-1}). On the basis of calculated stabilization energies, we can state that NH_3 is a stronger donor than both hydrocarbons as well as HCN and water. Following expectations, NMe_3 is a stronger donor than NH_3. Consequently, the donors can be ordered in the following sequence: $NMe_3 > NH_3 > H_2O > HCN > C_2H_2 > C_2H_4$. When comparing the stabilization energies of various acceptors, we found the following sequence: $F_2 < Cl_2 < SO_2 <$ ClF $< BH_3$. Comparing these orders with those based on HOMO/LUMO values, we have found important differences. This arises from the fact that the present ordering is based on the total stabilization energies and not only on the HOMO/LUMO energies. For 9 of the 11 complexes investigated, we have determined the E_2^{CT} CT energies (*cf.* Table 3.2). The E_2^{CT} energies could not be evaluated for the complexes with BH_3, as the analysis identify the respective complexes as one unit. The E_2^{CT} energies are very large, between 1.8 and 59 kcal.mol^{-1}, and they are larger than the interaction energies. When we order the electron donors and electron acceptors, we obtain identical orders to those above. This provides rather strong evidence that

the CT term by itself is dominant or very important stabilizing term.

Investigating the single energy terms in the variation interaction energy, we found that the MP2/CBS term forms the dominant energy contribution, with the CCSD(T) correction term being much smaller (forming on average about 8% of the MP2/CBS value). The only exception represent the "stacked" complexes ethylene$\cdots F_2$ and acetylene$\cdots$ClF, where the CCSD(T) correction term is larger. A similar situation occurred with the H-bonded complexes, where the $\Delta\Delta$CCSD(T) correction terms were also rather small.[159] The MP2/CBS and CCSD(T)/CBS interaction energies for all complexes investigated are visualized in Figure 3.4. (For labeling of x-axis, see Table 3.2). This figure clearly shows that for CT complexes of this type ("contact" structure) the expensive $\Delta\Delta$CCSD(T) correction term can be omitted and reliable stabilization energies are safely obtained already at the MP2/CBS level. The resulting error should be less than 10%.

Further, the Table 3.2 shows that all of the complexes are stable even

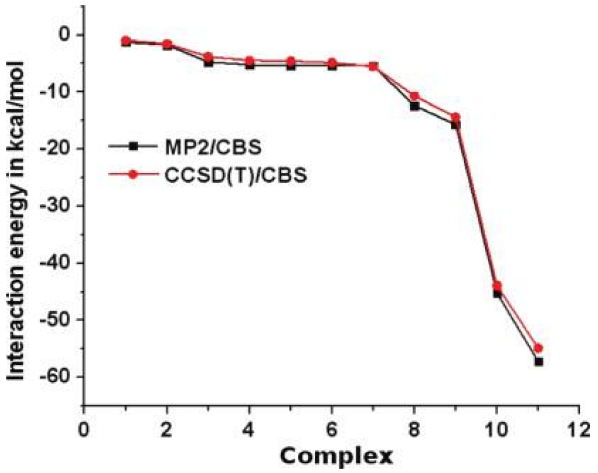

Figure 3.4: MP2/CBS and CCSD(T)/CBS interaction energies for all of the CT complexes.

at the HF level. Investigating values, we found that the HF/MP2 ratio (calculated at the cc-pVTZ level) is largest for the strong CT complexes, where it reaches more than 50%. On the other hand, this ratio considerably decreases (to less than 20%) when passing to weak CT complexes. Nevertheless, the HF

results clearly demonstrate that the CT term is covered already at the HF level.

Investigating the DFT-SAPT energies (*cf.* Table 3.3), we can state the following:
(1) The first-order polarization (Coulomb) energy is systematically the largest (attractive) term, and this term is always larger than the SAPT stabilization energy. The same is true for H-bonded complexes of the S22 set; for stacked and mixed complexes, the dispersion energy is dominant and the electrostatic term is mostly smaller than the SAPT stabilization energy.
(2) The dispersion energy and the δHF term are comparable. In most cases, the dispersion energy is slightly more attractive and only in the case of the $NH_3 \cdots ClF$ complex is the δHF term larger. The complexes containing the BH_3 as an acceptor are different and here the induction energy is much larger than the dispersion and δHF terms. The induction energy is larger than (or comparable to) the dispersion energy also for the $NH_3 \cdots ClF$ and $NMe_3 \cdots SO_2$ complexes. In the S22 set, the δHF term is always smaller (or even much smaller) than the dispersion energy, and the induction energy is systematically the smallest term. The rather large δHF and induction terms are thus characteristic features of CT complexes and reflect the importance of the CT energy term. The above-mentioned comparison confirms the different character of the complexes investigated and also a different nature of the stabilization when compared with H-bonded or stacked complexes.
(3) Five complexes (4, 8, 9, 10, and 11) are characteristic with a substantial CT (more than 0.05 e). These complexes have the sum of induction and δHF term ($E^{"CT"}$) significantly more attractive than the dispersion term (*cf.* Table 3.3).
(4) Stabilization energies for complexes 2 and 8, having NH_3 as an electron donor, differ substantially. While the former complex represents a typical vdW complex the later one can be characterized as a complex with strong halogen bond. The very positive σ-hole of ClF molecule is responsible for large electrostatic energy term. Further the CT energy in this complex is

also very large.
(5) Comparing the DFT-SAPT interaction energies with the CCSD(T)/CBS ones, we found that in most cases they are roughly comparable, with the average error of 7%. Only in case of $NH_3 \cdots F_2$ the error is significantly larger, about 20%. Not considering this complex, the average error of DFT-SAPT is only about 5%. In the S22 set, the CCSD(T)/CBS stabilization energies were mostly larger than the DFT-SAPT/CBS ones; only in 8 (out of 22) complexes was the opposite true, but the difference was systematically very small (mostly lower than 3%). Evidently, an average error is with present CT complexes comparable to that found for H-bonded and stacked complexes.

#	complex	E_1^{Pol}	E_1^{Ex}	E_2^{Ind}	δHF	$E^{"CT"}$	E_2^{Disp}	$E_{tot/CBS}$
1	$C_2H_4 \cdots F_2$	-1.40	2.54	0.04	-0.74	-0.70	-1.39	-1.04
2	$NH_3 \cdots F_2$	-2.86	4.33	-0.17	-0.93	-1.10	-1.53	-1.27
3	$C_2H_2 \cdots ClF$	-6.67	10.85	-1.10	-2.81	-3.91	-4.03	-4.09
4	$NH_3 \cdots Cl_2$	-11.49	16.81	-2.15	-3.04	-5.19	-4.34	-4.55
5	$HCN \cdots ClF$	-7.21	9.53	-1.35	-1.67	-3.02	-3.35	-4.30
6	$H_2O \cdots ClF$	-8.76	11.08	-1.47	-1.91	-3.38	-3.32	-4.66
7	$NH_3 \cdots SO_2$	-10.22	11.79	-1.78	-1.92	-3.70	-3.61	-5.99
8	$NH_3 \cdots ClF$	-28.81	43.88	-8.06	-8.94	-17.00	-7.53	-10.15
9	$NMe_3 \cdots SO_2$	-40.31	66.66	-13.84	-12.80	-26.64	-14.58	-15.75
10	$NH_3 \cdots BH_3$	-77.22	111.16	-48.82	-10.27	-59.09	-17.83	-44.20
11	$NMe_3 \cdots BH_3$	-89.51	126.53	-54.22	-10.37	-64.59	-23.64	-52.63

Table 3.3: DFT-SAPT interaction energies components (in kcal.mol^{-1}) for all of the CT complexes.

This finding is surprising because we expected that in case of strong CT complexes and especially with very strong CT complexes with very large CT, the DFT-SAPT interaction energy will differ more from the benchmark CCSD(T)/CBS values. Evidently this is not the case and perturbative DFT-SAPT analysis can be applied even for complexes having a nearly dative bond. This is evidently due to inclusion of δHF term which is, for strong CT

complexes, exceptionally large. An agreement of stabilization energies determined by both techniques is, for these complexes (10 and 11), surprisingly good. Only in one case was the CCSD(T)/CBS stabilization energy considerably larger, namely, with the $NH_3 \cdots F_2$ complex. Here the DFT-SAPT is underestimated by about 20%, and we have not been able to find the reason for such a failure. However, we must conclude that despite the existence of a rather large relative difference between both interaction energies the absolute error of 0.3 kcal.mol^{-1} is certainly not critical.

3.3 On the Nature of the Stabilization of Benzene$\cdots$Dihalogen and Benzene$\cdots$Dinitrogen Complexes

3.3.1 Introduction

The complexes of benzene with dihalogens can be seen as the CT complexes. As benzene is a prototypical aromatic electron donor, whereas dihalogens (F_2, Cl_2, and Br_2) are known to be good electron acceptors. The electron-donating orbitals in the former are π molecular bonding orbitals, whereas the electron-accepting orbitals in the latter are σ^* antibonding molecular orbitals. When moving from F_2 to Br_2 (or I_2), the electron-acceptor ability increases. An increase in the donor ability of benzene can be obtained, for example, from a substitution of hydrogen with an electron-donating group, such as a methyl group. Moving, however, from F_2 to Br_2 yields not only the mentioned increase in acceptor ability, but also an increase in the quadrupole moment and polarizability. Consequently, the electrostatic quadrupole-quadrupole (Q-Q) and dispersion energies are expected to increase as well. A discussion of the dominant stabilization term in the complexes mentioned becomes even more complicated when realizing that dihalogens, including F_2, exhibit a σ-hole, which is a prerequisite for the existence of a halogen bond. In the par-

ticular case of the C-X$\cdots\pi$ interaction (X=F, Cl, and Br), the halogen bond is of a strength comparable to that of a hydrogen bond and its existence can at least partially explain the stabilization of benzene$\cdots$dihalogen complexes. We have investigated the structure and stabilization of benzene, mesitylene, and hexamethylbenzene complexes with dihalogens (F_2, Cl_2, Br_2) and, for the sake of comparison, also with dinitrogen (N_2). The aim of the study was to elucidate the nature of stabilization in the complexes investigated, specifically to find the role of CT, electrostatic interaction, and halogen bonding, therefore, the DFT-SAPT method was utilized. The energy connected with the CT was estimated from the second-order perturbation theory analysis of the Fock matrix in the NBO basis,[352] or from the DFT-SAPT (see section 2.2.3.1).

In general E_2^{CT} energies are overestimated with respect to CT estimates from the SAPT theory ($E^{"CT"}$ energy); this was confirmed by results in our previous study (*cf.* section 3.2) where we investigated medium and strong CT complexes. For four $NH_3\cdots X$ complexes (X=F_2, Cl_2, SO_2, ClF), we found that the E_2^{CT} NBO CT energy was 2.4 times larger than the sum of the induction and δHF terms.

3.3.2 Strategy of Calculation

3.3.2.1 Structures and Geometries

The structures of all of the complexes were determined at the counterpoise-corrected MP2/cc-pVTZ level of theory. It was shown that these geometries were close to those obtained at the CCSD(T)/CBS level.[382] The geometries of the isolated systems were determined at the MP2/cc-pVTZ level. All of the complexes and the isolated systems were also optimized at the HF/6-311+G* level of theory.

3.3.2.2 Stabilization Energies

The CCSD(T)/CBS stabilization energy was constructed as the sum of the MP2/CBS stabilization energy and a CCSD(T) correction term ($\Delta\Delta$CCSD(T)), see eqn. 2.8. The former energy was extrapolated (Helgaker extrapolation)[161–163] from the MP2 energies evaluated at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels. The CCSD(T) correction term was determined in the aug-cc-pVDZ basis set.

DFT-SAPT calculations were performed with the PBE0AC functional in aug-cc-pVTZ basis sets.[327,328,365–367] The estimate of the CT energy, labeled as ($E^{"CT"}$), was considered as a sum of the induction (E_2^{Ind}) and δHF term. The DFT-SAPT/CBS value was based on calculations with the aug-cc-pVTZ basis set. Only the dispersion contribution was constructed differently. The E_2^D term was replaced by its CBS estimate, labeled as E_2^{Disp}. The E_2^{Disp} value was constructed from E_2^D/aug-cc-pVDZ and E_2^D/aug-cc-pVTZ terms utilizing the Helgaker extrapolation scheme.[161–163]

3.3.2.3 Frontier Orbitals, Electric Quadrupole Moments and Polarizabilities

Charge transferred between the electron donor and acceptor was deduced from the atomic charges of both subsystems. The atomic charges were approximated by NPA charges calculated at the HF/6-311+G* level.[352] The HOMO and LUMO energy, electric quadrupole moments, and polarizabilities of all of the subsystems were calculated at the HF/6-311+G* level of theory. The electrostatic multipole-multipole interaction energy, (abbreviated as Q-Q, because the quadrupole-quadrupole contribution was the most important one), was determined by utilizing distributed multipole moment analysis.[360] The distributed multipole moments for all of the molecular sites were calculated up to hexadecapole, based on the PBE0/aug-cc-pVDZ wave function.

3.3.3 Results and Discussion

3.3.3.1 Isolated Subsystems

3.3.3.1.1 Frontier Orbitals

The HOMO and LUMO energies of all of the donors and acceptors are provided in Table 3.4. According to our expectations, the electron donors can be ordered as follows: benzene < mesitylene < hexamethylbenzene, while the electron-acceptor ability increases in the following order: $N_2 < F_2 < Cl_2 < Br_2$.

molecule	HOMO [eV]	LUMO [eV]	Q_{zz} [D.A]	α [Bohr^3]	ESP [au]
F_2	-	1.986	0.33	6.3	0.0173
Cl_2	-	0.569	1.90	20.8	0.0419
Br_2	-	-0.231	3.24	33.1	0.0493
N_2	-	3.344	-1.00	10.0	-0.0140
C_6H_6	-9.162	-	-6.44	63.7	-
$C_6H_3(CH_3)_3$	-8.572	-	-5.47	98.9	-
$C_6(CH_3)_6$	-8.033	-	-4.33	132.1	-

Table 3.4: Molecular properties of investigated molecules calculated at the HF/6-311+G*.

3.3.3.1.2 Multipole Moments, Polarizabilities and ESP

The first nonzero multipole moment for all of the subsystems corresponds to the quadrupole moment and their values are collected in Table 3.4. Benzene belongs to the group of aromatic compounds with a negative Q_{zz} component of the quadrupole moment (the z axis coincides with the main rotational axis of the symmetry of the molecule C_6, which is oriented perpendicular to the plane of the carbon ring). The absolute value of the Q_{zz} component of dinitrogen is considerably smaller than that of benzene. In the case

of dinitrogen the z axis has same orientation as the main rotational axis of the diatomic C_∞, which in this case coincides with the bond). The signs of the dinitrogen schematic quadrupole moment are not surprising because the molecule possesses a triple bond and has a lone electron pair at each nitrogen (|N≡N|). When moving to dihalogens, the sign of the quadrupole is changed.[385] Dihalogen possesses a single bond and each halogen has three lone electron pairs, (| $\overline{\underline{F}}$ - $\underline{\overline{F}}$|); we expected that the sign of their quadrupole moments should be same as in the case of dinitrogen. It should be mentioned here that quantum mechanical calculations predicted quadrupole moments of all systems in agreement with experiment results. An interpretation of the different signs of the dihalogen and dinitrogen quadrupole moments can be rationalized by performing the natural population analysis (NPA). The NPA provides in the case of dihalogens following occupancies of the valence p-type natural atomic orbitals (NAOs). The occupancy of the p_z orbital is smaller (1.05 , 1.05 , 1.04 for difluorine, dichlorine and diiodine) than p_x and p_y (equals to 2.00 for both direction for all three dihalogens), whereas the opposite is true for dinitrogen. The occupancy of the p_x and p_y orbitals equals to 0.99 and p_z to 1.36. Dihalogens thus lack electrons in the z axis, whereas dinitrogen has a surplus of electrons in this axis. Figure 3.5 shows the ESP of all the dihalogens and, for comparison, also for dinitrogen. A substantial difference is evident at first glance. Whereas the ESP at the top (cusp point) of the dihalogen molecules is positive (i.e., each molecule exhibits a σ-hole), it is negative in the case of dinitrogen. Correspondingly, the Q_{zz} component of the quadrupole moments in these systems should be opposite, that is, positive for dihalogens and negative for dinitrogen.

Investigating the subsystem polarizabilities (*cf.* Table 3.4), we recognize a substantial increase upon the methylation of benzene and hexamethylbenzene possesses twice the polarizability of benzene. Similarly, there is a polarizability increase when moving from F_2 to Br_2; the polarizability of dibromine is more than five times greater than that of difluorine. The polarizability of difluorine and dinitrogen are similar. These values indicate that the dispersion energy in benzene$\cdots$dihalogen complexes significantly increases upon

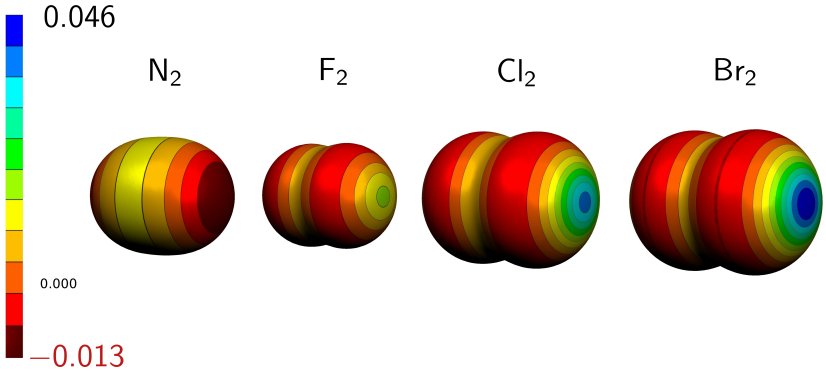

Figure 3.5: ESP of dihalogens and dinitrogen.

moving from F_2 to Br_2 and is comparable for complexes of benzene with difluorine and benzene with dinitrogen.

3.3.3.2 Complexes

3.3.3.2.1 Structure and Stabilization Energies

The structures of all five conformers considered for benzene$\cdots$diatomic complexes are depicted in Figure 3.6.

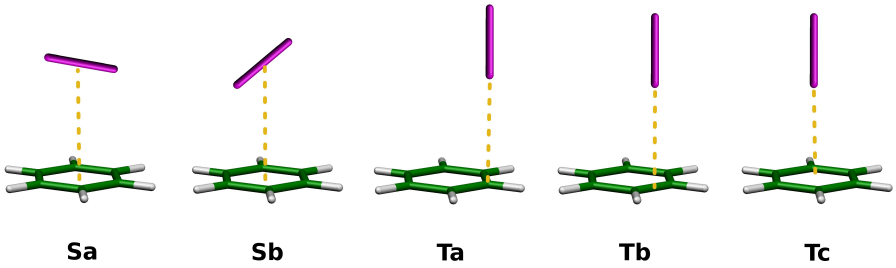

Figure 3.6: Schematic depiction of five different conformers considered for the benzene$\cdots X_2$ complexes.

The MP2 and CCSD(T) interaction energies together with CT data are shown in Table 3.5, whereas Table 3.6 (shown below) contains the DFT-SAPT energies. From Table 3.5, it is apparent that the T-shaped structures Ta and Tb are more stable than the symmetrical Tc structure. The energy difference between these structures is largest for dibromine (about 0.5

kcal.mol^{-1}) and smallest for difluorine (less than 0.1 kcal.mol^{-1}). The total stabilization energies of the complexes investigated are rather large, which is especially true for the dibromine-containing complexes. The stabilization energies of benzene$\cdots$Br$_2$, mesitylene$\cdots$Br$_2$, and hexamethylbenzene$\cdots$Br$_2$ (for the substituted benzene, only the Tc structure was considered) amount to 3.2, 4.2, and 5.7 kcal.mol^{-1}, respectively. By investigating the components of the CCSD(T)/CBS stabilization energy, we found that the $\Delta\Delta$CCSD(T) correction term was systematically repulsive and surprisingly large. For the Ta structure, it reduced the MP2/CBS stabilization energy of benzene$\cdots$ dihalogen complexes by 1.77, 1.32, and 0.46 kcal.mol^{-1} (33, 32, and 28%), respectively. The situation with benzene$\cdots$dinitrogen was different because here the sandwich structures was more stable than the T-shaped one. The role of the CCSD(T) correction term was similar here and the MP2/CBS stabilization energies were again strongly overestimated (by more than 30%).

The interpretation of the different structural preferences of the benzene$\cdots$dihalogen and benzene$\cdots$dinitrogen complexes in terms of the monomer quadrupole moments is straightforward. The signs and orientations of the quadrupole moments of benzene and dihalogens lead to a preference for T-shaped structures, whereas in the case of the benzene$\cdots$dinitrogen complex stacked structures are preferred. The electrostatic Q-Q interaction thus explains the structure of the complexes investigated (*cf.* Figure 3.7 and 3.8). Figure 3.7 shows the different character of the Q-Q interaction energy when moving to a benzene$\cdots$dinitrogen complex. The Q-Q interaction for the Sa structure is attractive, whereas in the case of the Tc structure it is repulsive. This is in full accordance with our previous conclusion based only on the signs and orientations of the quadrupole moments of benzene and dinitrogen. The Q-Q interaction energies for the Tc structures of the $C_6H_6\cdots Br_2$, $C_6H_6\cdots Cl_2$, and $C_6H_6\cdots F_2$ complexes for distances from 5.5 to 12 Å are depicted in Figure 3.8. The largest stabilization energies belong to the complexes of benzene with dibromine, followed by dichlorine, and difluorine. This order correlates with the magnitude of the σ-hole and naturally also with the ESP values along the main rotational axis of the dihalogens (*cf.* Figure 3.5).

complex	MP2/CBS	CCSD(T)/CBS	CT	E_2^{CT}
Ben$\cdots Br_2$(Sa)	-2.89	-1.98	0.0028	-
Ben$\cdots Br_2$(Sb)	-2.89	-1.98	0.0029	-
Ben$\cdots Br_2$(Ta)	-5.43	-3.66	-0.0167	-4.52
Ben$\cdots Br_2$(Tb)	-5.45	-3.66	-0.0169	-5.53
Ben$\cdots Br_2$(Tc)	-4.39	-3.17	-0.0024	-1.44
Ben$\cdots Cl_2$(Sa)	-2.47	-1.70	0.0035	-
Ben$\cdots Cl_2$(Sb)	-2.47	-1.70	0.0035	-
Ben$\cdots Cl_2$(Ta)	-4.16	-2.84	-0.0078	-2.61
Ben$\cdots Cl_2$(Tb)	-4.19	-2.86	-0.0077	-3.06
Ben$\cdots Cl_2$(Tc)	-3.52	-2.53	-0.0001	-0.90
Ben$\cdots F_2$(Sa)	-0.87	-0.73	0.0016	-
Ben$\cdots F_2$(Sb)	-0.87	-0.73	0.0016	-
Ben$\cdots F_2$(Ta)	-1.65	-1.19	-0.0006	-0.59
Ben$\cdots F_2$(Tb)	-1.62	-1.17	-0.0005	-0.75
Ben$\cdots F_2$(Tc)	-1.44	-1.10	0.0008	-
Ben$\cdots N_2$(Sa)	-2.28	-1.50	-0.0018	-0.40
Ben$\cdots N_2$(Sb)	-2.28	-1.50	-0.0018	-0.40
Ben$\cdots N_2$(Tc)	-1.26	-0.81	-0.0028	-0.30
Mesi$\cdots Br_2$(Sb)	-4.43	-3.13	0.0051	-
Mesi$\cdots Br_2$(Tc)	-5.90	-4.23	0.0023	-2.19
Hxme$\cdots Br_2$(Sb)	-5.95	-4.12	0.0007	-
Hxme$\cdots Br_2$(Tc)	-7.98	-5.66	0.0033	-

Table 3.5: The interaction energies (in kcal.mol^{-1}) calculated at the MP2-/CBS and CCSD(T)/CBS levels of theory; CT (in au), negative value indicates CT from benzene (and substituted benzene) to the diatomic molecule; E_2^{CT} - CT energies from bonding $\pi \rightarrow$ antibonding σ^* orbitals (in kcal.mol^{-1}).

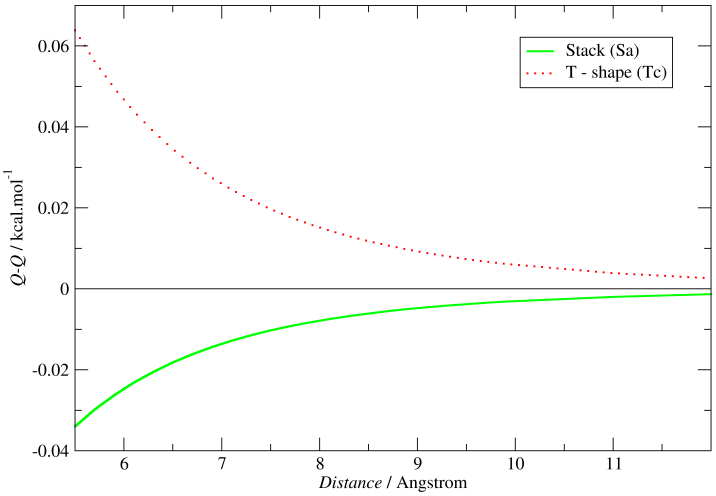

Figure 3.7: Quadrupole-Quadrupole interaction for the benzene$\cdots N_2$ complex in the Sa and Tc conformations.

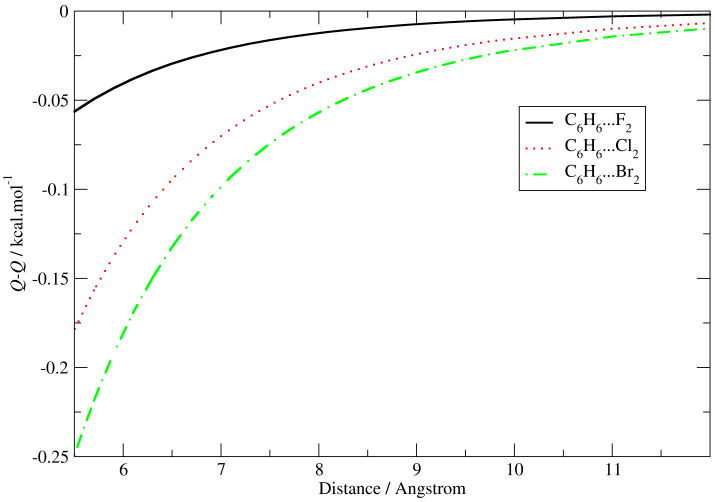

Figure 3.8: Q-Q interaction for the benzene$\cdots$dihalogen complexes in Tc conformation.

Table 3.6 shows that the dominant attractive energy contribution for all of the structures of all complexes originates in the dispersion energy. Only in the case of the T-shaped structures of benzene (and substituted-benzene) $\cdots$ dibromine is the first-order electrostatic E_1^{Pol} energy comparable to the dispersion energy; in all of the other complexes, the electrostatic energy is systematically smaller.

complex	E_1^{Pol}	E_1^{Ex}	$E^{"CT"}$	E_2^{Disp}	$E_{tot/CBS}$
Ben$\cdots$Br$_2$(Sa)	-0.89	3.29	-0.24	-4.00	-2.07
Ben$\cdots$Br$_2$(Sb)	-0.89	3.30	-0.24	-4.01	-2.07
Ben$\cdots$Br$_2$(Ta)	-4.88	9.60	-2.78	-5.81	-4.22
Ben$\cdots$Br$_2$(Tb)	-5.01	9.83	-2.87	-5.85	-4.23
Ben$\cdots$Br$_2$(Tc)	-2.60	4.82	-0.94	-4.48	-3.44
Ben$\cdots$Cl$_2$(Sa)	-0.69	2.71	-0.20	-3.44	-1.79
Ben$\cdots$Cl$_2$(Sb)	-0.71	2.75	-0.20	-3.47	-1.80
Ben$\cdots$Cl$_2$(Ta)	-3.19	6.55	-1.74	-4.52	-3.13
Ben$\cdots$Cl$_2$(Tb)	-3.17	6.52	-1.73	-4.53	-3.14
Ben$\cdots$Cl$_2$(Tc)	-1.87	3.69	-0.65	-3.69	-2.70
Ben$\cdots$F$_2$(Sa)	-0.08	0.55	-0.04	-1.11	-0.76
Ben$\cdots$F$_2$(Sb)	-0.08	0.55	-0.04	-1.12	-0.76
Ben$\cdots$F$_2$(Ta)	-0.97	2.14	-0.51	-1.81	-1.23
Ben$\cdots$F$_2$(Tb)	-0.96	2.08	-0.49	-1.77	-1.22
Ben$\cdots$F$_2$(Tc)	-0.55	1.20	-0.17	-1.52	-1.11
Ben$\cdots$N$_2$(Sa)	-1.09	2.78	-0.37	-2.79	-1.59
Ben$\cdots$N$_2$(Sb)	-1.09	2.77	-0.37	-2.79	-1.59
Ben$\cdots$N$_2$(Tc)	-0.28	1.88	-0.26	-2.12	-0.84
Mesi$\cdots$Br$_2$(Sb)	-1.63	5.22	-0.43	-6.07	-3.22
Mesi$\cdots$Br$_2$(Tc)	-4.04	7.31	-1.50	-6.34	-4.89
Hxme$\cdots$Br$_2$(Sb)	-2.52	7.25	-0.62	-7.96	-4.23
Hxme$\cdots$Br$_2$(Tc)	-5.43	9.60	-2.07	-7.79	-6.06

Table 3.6: DFT-SAPT interaction energies (in kcal.mol^{-1}) of investigated complexes; $E^{"CT"}$ - sum of sencond-order induction, exchange induction and δHF term.

The explanation for the attractive interaction between halogens (with the ex-

ception of fluorine) covalently bonded to carbon and an electron donor such as the carbonyl oxygen (halogen bond) was based on the existence of a σ-hole. The σ-hole is also present in dihalogens, even in difluorine, and surprisingly, it is very positive. Consequently, a halogen bond should also be formed between dihalogens and electron donors such as benzene or substituted benzene. The ESP of isolated benzene has its most negative part located above (and below) the skeleton of carbon atoms, whereas the potential becomes more positive when moving to the center of the aromatic ring. Consequently, the halogen bond between benzene and the dihalogens will be preferentially formed above the carbon skeleton and not above the center of the aromatic ring. The DFT-SAPT electrostatic energies fully confirm this assumption and the electrostatic energy for Ta and Tb structures of all of the benzene$\cdots$dihalogen complexes is considerably more attractive than that for the symmetrical Tc structure. We can thus conclude that the rather large stabilization energies of benzene$\cdots$dihalogen complexes, as well as the fact that the Ta and Tb structures are more favorable than Tc, can be explained by the existence of a halogen bond between the dihalogens and benzene.

The stabilization energy of the benzene$\cdots$dinitrogen complex is considerably smaller than that of the benzene$\cdots X_2$ (X = Br, Cl) complexes and is roughly comparable to that of benzene$\cdots$difluorine. Investigating the energy terms, we found that the larger stabilization of dibromine and dichlorine complexes aroses from the larger dispersion, induction, and mainly electrostatic terms. This is in accordance with previous conclusions, showing that the greater stabilization of benzene$\cdots X_2$ (X = Br, Cl) complexes arises from the existence of a halogen bond.

3.3.3.2.2 Charge-Transfer

Table 3.5 shows the amount of "electrons transferred" between the subsystems. According to expectations, CT is negligible for benzene$\cdots$dinitrogen.

Out of the benzene$\cdots$dihalogen complexes, the greatest electron transfer is found for the dibromine complex, followed by those of the dichlorine and difluorine complexes. An overlap between the benzene HOMO and dihalogen LUMO is significant only for the T-shaped structures. Consequently, structures Ta and Tb exhibited much greater electron transfer than the sandwich structure. In all cases, electron transfer is, however, rather small and does not exceed 0.02 e. From these values, we can deduce that electron-transfer energy will be also moderate. CT energy contributions can be deduced from DFT-SAPT analysis, for which the upper bound to CT energy is defined as the sum of the second-order induction and δHF terms. Sum of these three terms is labeled as E$^{"CT"}$. Table 3.6 shows that for all of the complexes and all of the structures, the E$^{"CT"}$ energy represents the smallest attractive energy term. This is true even for the hexamethylbenzene$\cdots Br_2$ complex, for which we expected the CT energy to be much more important if not dominant. It is non-negligible only for the Ta and Tb structures of the benzene$\cdots$dichlorine and benzene (and substituted-benzene)$\cdots$dibromine complexes. We can thus conclude that in the present complexes the CT energy contribution does not represent the dominant stabilization energy term. Table 3.5 lists the CT energy approximated by the second-order perturbation theory analysis of the Fock matrix,[352] see eqn. 3.1. It should be again mentioned that the E_2^{CT} values are overestimated (*cf.* Table 3.5 and 3.6) and that, owing to their different origin with respect to CT from SAPT theory, they cannot be directly compared. However, the E_2^{CT} energies can be compared for different structures of the complex or for different complexes. When analyzing these energies, we found that they reached the highest values for the Tb structures; for the Ta and Tc structures, they are significantly smaller. According to our expectations, E_2^{CT} energy was largest for benzene$\cdots Br_2$ (5.44 kcal.mol^{-1}), followed by the benzene$\cdots Cl_2$, and benzene$\cdots F_2$. For the sake of comparison, we list the E_2^{CT} values for some typical CT complexes: $NH_3\cdots Cl_2$ (14.0 kcal.mol^{-1}), $NH_3\cdots SO_2$ (7.4 kcal.mol^{-1}) and $NMe_3\cdots SO_2$ (58.7 kcal.mol^{-1}) (see Attachment **B**). It is thus clear that the CT E_2^{CT} energy within the benzene$\cdots$dibromine complex is not exceptional, but is in fact moderate.

3.4 Why Is the L-shaped Structure of $X_2 \cdots X_2$ (X = F, Cl, Br, I) Complexes More Stable Than Other Structures?

3.4.1 Introduction

The first nonzero electric multipole moment of dihalogens (X_2, X = F, Cl, Br, I) is the quadrupole moment (Q_{zz}), which is positive for all of these systems. The quadrupole moment of dinitrogen is negative. The explanation of the different signs of quadrupole moments of the dihalogens and the dinitrogen is not easy because both halogens and nitrogen bear lone electron pairs and pair, respectively. This is in accord with the sign of the quadrupole moment of dinitrogen but not with that of dihalogens. To our knowledge, no easy explanation of this fact was available until the recent introduction of the σ-hole, which has been used to explain the origin of the attraction in the halogen bond.[111] The existence of halogen bonding, described as an attractive interaction between a bound halogen and an electronegative atom, seems counterintuitive, given that an attractive noncovalent interaction is not expected to exist between two atoms that have high electronegativity, that is, possess a partial negative charge. The reason for the attractive noncovalent interaction that occurs in halogen bonds is the presence of a region of a positive electrostatic potential (σ-hole) along the extension of the C-X bond (X is most typically bound to carbon), which interacts electrostatically with an electron donor. The existence of the positive σ-hole explains positive quadrupole moments in dihalogens; similarly, the nonexistence of the σ-hole in dinitrogen explains its negative quadrupole moment. Are the two concepts fully equivalent? In other words, do they both lead to the same structure prediction? The aim of the present paper is to investigate different binding motifs of the (X_2, X = F, Cl, Br, I, N) complexes and elucidate how the presence of σ-hole affects the relative stability of investigated binding motifs.

3.4.2 Strategy of Calculation

3.4.2.1 Quadrupole Moments and Electrostatic Potentials

The electrostatic potential (ESP)[386–389] was computed for all subsystems at the B97/def2-QZVP level[255,256]and subsequently mapped on the 0.001 au isodensity surface.[124] The relativistic effects were included by considering pseudopotentials (PPs).[390,391] The point on the 0.001 au isodensity surface, which lies on the main rotational axis of the X_2 molecule, is referred to here as $V_{S,max}$. In the case of dihalogens, this point possesses the most positive value of the ESP. The angular dependence of the ESP was also investigated. The angle α, at which the ESP becomes negative when moving from the $V_{S,max}$ point on the 0.001 isodensity surface, was evaluated.

3.4.2.2 Structures and Geometries

Five different structures of the $(X_2)_2$ dimers (L-shaped (LS), T-shaped (TS), parallel (P), parallel-displaced (PD), and linear (L)) were considered, and the respective energy minima were determined by an unrelaxed potential energy scan along the main intermolecular coordinates (see Figure 3.9). The scans were performed on the grid with a point-to-point distance of 0.1 Å. The geometries of the X_2 molecules considered for all calculations were calculated at the B97-D3/def2-QZVP level of theory.

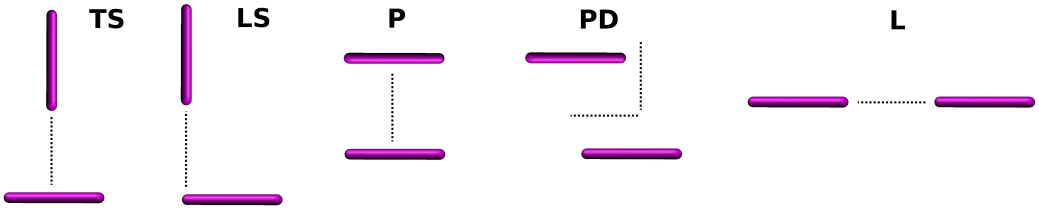

Figure 3.9: Structures of the L-shaped (LS), T-shaped (TS), parallel (P), parallel-displaced (PD), and linear (L) conformations.

3.4.2.3 Stabilization Energies

The benchmark interaction energies were evaluated at the CCSD(T) level and extrapolated to the complete basis set (CBS) limit. Specifically, the CCSD(T)-/CBS interaction energy was constructed as the sum of HF/CBS interaction energy and the correlation part of the MP2/CBS interaction energy. Both were determined by the two-point extrapolation scheme of Helgaker from aug-cc-pVTZ and aug-cc-pVQZ basis sets.[161–163] The CCSD(T) correction term ($\Delta\Delta$CCSD(T)) was evaluated utilizing the aug-cc-pVTZ basis set. The relativistic effects were included by considering the pseudopotentials (PPs).[390,391] Besides CCSD(T), two variants of the DFT method were applied. The M06-2X functional[246,247] was recommended for calculations of halogen-bonded complexes.[392] Therefore, this functional in combination with the aug-cc-pVTZ basis set was used in the present study along with the DFT-D (B97-D3/def2-QZVP) method.[236] The Grimme's empirical dispersion correction (D3) was calculated employing Becke-Johnson damping.[266] In the case of bromine and iodine at the M06-2X levels, PPs consistent with respective correlation consistent basis sets (aug-cc-pVXZ-PP) were applied. At the B97-D3/def2-QZVP level, the PPs were considered only for iodine.
The DFT-SAPT calculations were used for the decomposition of the total interaction energy. The subsystems were treated via the DFT approach, utilizing the asymptotically corrected LPBE0AC exchange-correlation functional[319] in the aug-cc-pVTZ basis set for nitrogen, fluorine, and chlorine. In the case of bromine and iodine, the aug-cc-pVTZ-PP basis set was used to account for relativistic effects. The exchange-induction and exchange-dispersion terms are merged into the respective induction and dispersion terms. Further, the δHF term, which represents higher than second-order terms covered by the Hartree-Fock approach, is also included in the induction energy. Hence, the induction energy in this study (E_2^{Ind}) represents the upper bound to the estimate of CT energy. (This term was in previous studies labeled as $E^{"CT"}$).

3.4.3 Results and Discussion

3.4.3.1 Isolated Systems

Table 3.7 summarizes the geometries, quadrupole moments, and size (α) and magnitude ($V_{S,max}$) of σ-holes calculated for all subsystems at the B97/def2-QZVP level of theory.

molecule	geometry	Q_{zz}	$V_{S,max}$
I_2	2.703	3.402	0.0474
Br_2	2.322	2.362	0.0443
Cl_2	2.013	1.574	0.0389
F_2	1.410	0.436	0.0243
N_2	1.116	-0.774	-0.0140

Table 3.7: The geometries (in Å), quadrupole moment (the Q_{zz} component in ea_0^2) and $V_{S,max}$ (in au) for dihalogens and dinitrogen.

The electrostatic potential for all subsystems is visualized in Figure 3.10. The quadrupole moments of all dihalogens are positive, while the quadrupole moment of dinitrogen is negative. The same is valid for the magnitudes of the σ-holes ($V_{S,max}$). The negative value of $V_{S,max}$ for dinitrogen indicates that the positive σ-hole is not present. The largest absolute value of both characteristics (quadrupole moments and $V_{S,max}$) was found for diiodine, and the smallest positive one was found for difluorine (the smallest $V_{S,max}$ was calculated for dinitrogen); quadrupole moments and $V_{S,max}$ correlate very well ($R^2 = 0.849$) for all systems. This finding is important because it shows that at least for the systems studied (and similar systems for which the first nonzero multipole moment is quadrupole) their nontrivial electronic structure is described already by the concept of electric quadrupoles and that the recently introduced concept of the σ-hole is not bringing any new information.

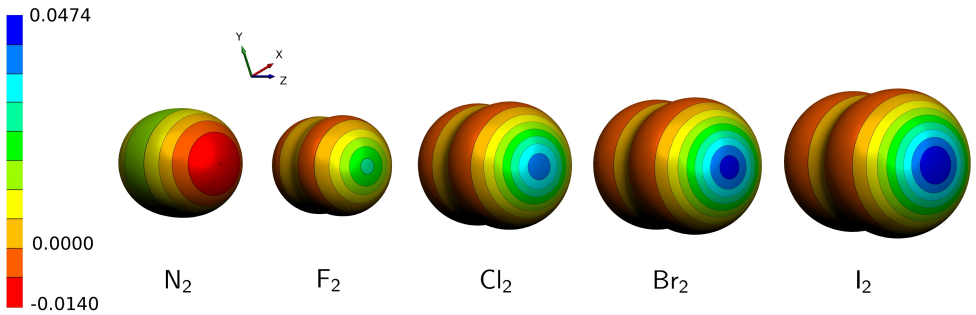

Figure 3.10: Electrostatic potential (in au) for all diatomics.

Table 3.8 presents the X-X bond orbital analysis together with the occupancy of the p-type valence natural atomic orbitals (NAOs) and $V_{S,max}$ values. It further lists the size (angle α) and magnitude ($V_{S,max}$) of the σ-hole.

	NHOs		NAOs				
X	s	p	p_x	p_y	p_z	$V_{S,max}$	α
F	4.9	94.9	1.999	1.999	1.038	0.0243	59
C	6.2	92.5	1.992	1.992	1.043	0.0389	62
B	4.3	94.8	1.994	1.994	1.029	0.0443	64
I	3.5	95.6	1.994	1.994	1.022	0.0474	65
N	36.6	62.9	0.996	0.996	1.325	-0.0140	-
	0.0	99.6					

Table 3.8: The $V_{S,max}$ (in au) and the size of the σ-hole (α in °). The hybridization state (in %) of the natural hybrid orbital (NHO) of atom X in the natural X-X bonding orbital. The occupancies of the p-type valence natural atomic orbitals (NAOs) of atom X; for more details see Table 2 in Attachment **D**.

When one moves from difluorine toward heavier dihalogens, the slight increase of α is in correlation with the significant increase of $V_{S,max}$. The small variation of the size of the σ-hole, which varies between 59° (F_2) and 65°(I_2), can be interpreted using the natural-bond orbital (NBO) analysis. First, the hybridization between s and p orbitals, in the case of the X-X natural-bond orbital, is negligible, not exceeding 6% (dichlorine). Second, the occu-

pancy of p-type valence NAOs does not differ much between different halogens. Specifically, the valence p_z occupancy varies between 1.022 (I_2) and 1.043 (Cl_2). Moreover, the occupancy of the p_x and p_z orbitals is essentially constant. Finally, it is clear that the hybridization state and the occupancy of NAO are more or less constant for all dihalogens. Consequently, we conclude that the angular redistribution of the valence electrons is very similar for all dihalogens. Hence, marginal differences in the size of the σ-hole are observed.

The character of the electron redistribution is entirely different for the dinitrogen molecule when compared with dihalogens. The relative occupancy of the p-type NAOs is reverse when compared with dihalogens.

3.4.3.2 Complexes

The total interaction energies of all dihalogen dimers determined by the B97-D3, M06-2X, CCSD(T), and DFT-SAPT techniques are given in Table 3.9. The numbers in parentheses listed with the B97-D3 energies correspond to the respective empirical dispersion energies.

In the case of DFT-SAPT calculations presented in Table 3.9, the potential-energy scans were not made, and the values correspond to the structures, which represent the CCSD(T)/CBS energy minima. Stabilization energies evaluated at different levels of theory correlate well with reference CCSD(T) stabilization energies. Specifically, coefficient of determination (R^2) amounts to 0.95, 0.92, and 0.99 for the B97-D3, M06-2X, and DFT-SAPT method. The root-mean-square deviation (RMSD) with respect to reference data, is following: 0.15, 0.34, and 0.44 kcal.mol^{-1} for DFT-SAPT, M06-2X and B97-D3, respectively.

Investigating the reference CCSD(T)/CBS stabilization energies of dihalogen complexes, we found that the LS structure corresponds with the most stable structure, followed by the TS, PD, P, and L structures. For heavier

	R	B97-D3	R	M06-2X	R	CCSD(T)/CBS	DFT-SAPT
LS							
I_2	5.0	-4.77 (-3.44)	5.1	-2.96	5.1	-3.16	-3.64
Br_2	4.6	-2.86 (-2.68)	4.6	-2.31	4.7	-2.38	-2.47
Cl_2	4.4	-1.58 (-1.94)	4.3	-1.18	4.4	-1.47	-1.20
F_2	4.0	-0.30 (-0.23)	3.7	-0.39	3.6	-0.41	-0.30
N_2	4.3	-0.47 (-0.46)	4.4	-0.21	4.2	-0.28	-0.26
TS							
I_2	5.2	-3.04 (-3.23)	5.1	-2.27	5.2	-2.45	-2.61
Br_2	4.9	-1.96 (-2.25)	4.7	-1.61	4.7	-1.89	-1.93
Cl_2	4.6	-1.25 (-1.69)	4.4	-0.87	4.5	-1.25	-1.05
F_2	3.8	-0.29 (-0.24)	3.6	-0.38	3.6	-0.39	-0.31
N_2	4.2	-0.48 (-0.45)	4.3	-0.21	4.1	-0.29	-0.27
P							
I_2	4.5	-2.22 (-3.36)	4.4	-0.53	4.5	-1.44	-1.46
Br_2	4.2	-1.50 (-2.39)	4.2	-0.37	4.2	-1.14	-1.08
Cl_2	4.1	-1.04 (-1.49)	4.0	-0.40	3.9	-0.91	-0.78
F_2	3.6	-0.26 (-0.22)	3.1	-0.36	3.2	-0.31	-0.25
N_2	3.8	-0.51 (-0.57)	4.0	-0.20	3.7	-0.23	-0.22
PD							
I_2	4.5	-2.82 (-3.85)	4.3	-1.73	4.4	-2.13	-2.26
Br_2	4.2	-1.83 (-2.64)	4.0	-1.19	4.0	-1.61	-1.58
Cl_2	3.9	-1.23 (-2.08)	3.7	-0.89	3.8	-1.24	-1.02
F_2	3.6	-0.26 (-0.25)	3.3	-0.34	3.2	-0.36	-0.28
N_2	3.9	-0.55 (-0.60)	4.4	-0.20	4.0	-0.31	-0.29
L							
I_2	6.9	-0.72 (-1.61)	6.5	-0.18	6.6	-0.57	-0.44
Br_2	6.4	-0.43 (-1.01)	5.8	0.17	6.0	-0.28	-0.15
Cl_2	5.8	-0.34 (-0.81)	6.1	0.02	5.5	-0.25	-0.11
F_2	4.7	-0.11 (-0.13)	4.4	-0.14	4.3	-0.12	-0.11
N_2	5.0	-0.13 (-0.20)	5.1	-0.09	4.8	-0.07	-0.07

Table 3.9: The interaction energies (in kcal.mol^{-1}) calculated at the B97-D3, M06-2X, CCSD(T)/CBS and DFT-SAPT levels of theory. The numbers in parentheses correspond to dispersion energy. R (in Å) corresponds to the center of mass distance.

dihalogens, the LS and TS structures are more stable than other structures; in the case of difluorine, all the structures, except linear (L), are comparably stable (*cf.* Figure 3.11). This finding is surprising because it was long believed that the TS structure of dihalogens, stabilized by quadrupole-quadrupole electrostatic interaction, corresponds to the global minimum.

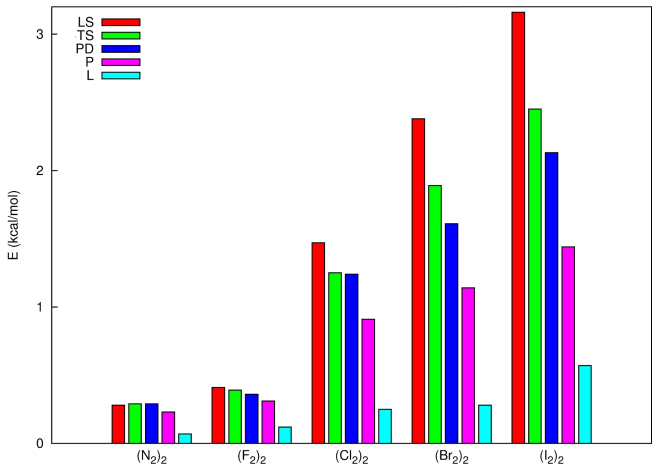

Figure 3.11: CCSD(T)/CBS stabilization energies (in kcal.mol^{-1}) for all X_2 complexes.

Analyzing the DFT-SAPT energy components (*cf.* Figure 3.12), we found that this is mainly caused by the Coulomb E_1^{Pol} energies. The other attractive energies (dispersion and induction) are also the largest for the LS structure, but the absolute difference with respect to other structures (TS, PD, P, and L) is much smaller. This can be interpreted as a consequence of the higher orientation dependence of the Coulomb interaction, in contrast with induction and dispersion. However, it should be stressed that the largest attractive contribution for all structures (including LS and TS) comes from dispersion energy, followed by Coulomb energy.

To exclude the overlap effects (see below), we evaluate the total interaction energies at the DFT-SAPT level for larger intermolecular distances. These scans are performed for the two most stable structures (LS and TS) of the diiodine dimer and will be discussed in the following paragraphs.

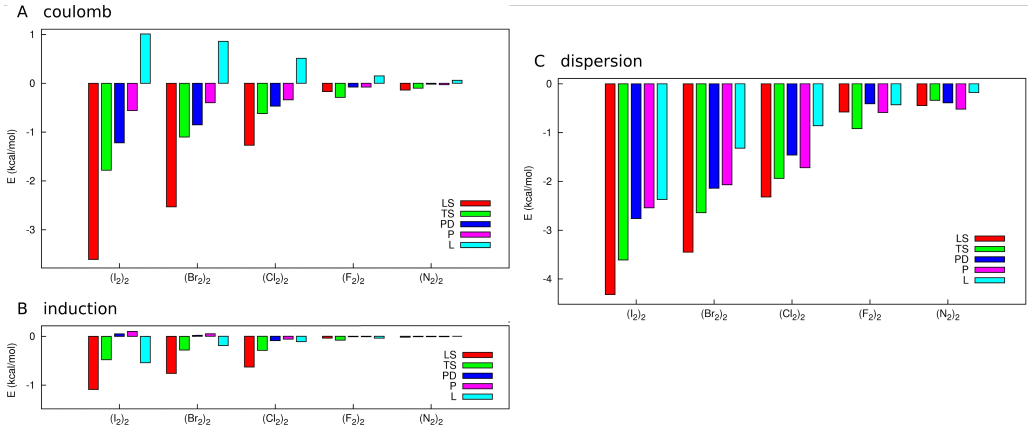

Figure 3.12: Coulomb E_1^{Pol} (A), induction E_2^{Ind} (B), and dispersion E_2^{Disp} (C) components of the DFT-SAPT interaction energy (in kcal.mol^{-1}), listed for all X_2 complexes.

Figure 3.13 shows the distance dependence of the total DFT-SAPT as well as Coulomb (E_1^{Pol}) energies for the LS and TS structures of the diiodine dimer. At short distances, the Coulomb energy is evidently more attractive for the LS structure, while the opposite is true at larger distances. On the other hand, the total DFT-SAPT energy is systematically larger for the LS structure. This finding can be easily explained on the basis of penetration energy (*cf.* section 2.2.3.1). The systematically attractive penetration energy, which is included in the DFT-SAPT Coulomb energy, is overlap-dependent. The overlap in the LS structure is clearly larger than that in the TS structure because of a closer X$\cdots$X contact in the former structure. Therefore, the Coulomb energy is larger for the LS structure at short distances. At large distances, the penetration of both molecules becomes negligible and the DFT-SAPT Coulomb energy is exclusively represented by long-range electrostatic energy. The long-range electrostatic energy covering all the interactions among all the permanent multipole moments is less attractive for the L-shaped structure than for the T-shaped structure (*cf.* Figure 3.13).

Regarding explanation of the larger total stabilization energy in the L-shaped structure. Figure 3.14 shows the distance dependence of the remaining at-

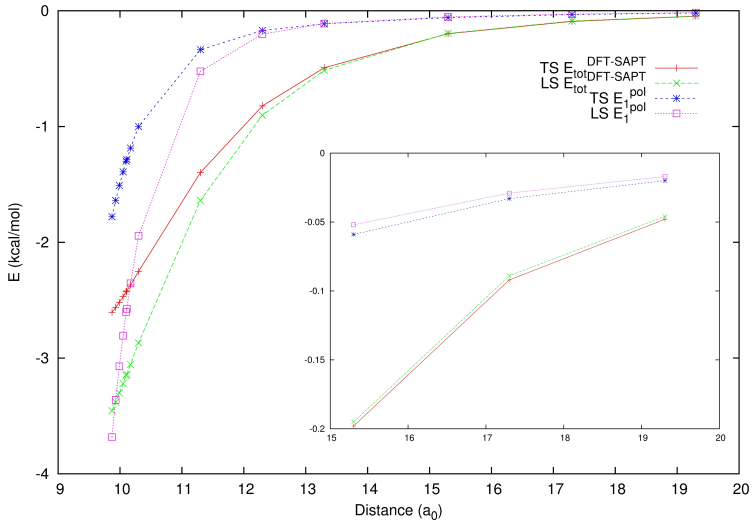

Figure 3.13: Distance dependence of the total interaction (E_{tot} DFT-SAPT) and the Coulomb (E_1^{Pol}) energies for the LS and TS structures of the diiodine dimer.

tractive energy terms, dispersion (E_2^{Disp}) and induction (E_2^{Ind}). It is evident that both energies are systematically more attractive for the LS structure. This was also observed for equilibrium geometries (see above). The reason for this is again the shorter interatomic distances for the LS structure. The preference of the LS structure in the entire distance region cannot be explained solely by either electrostatic or Coulomb energies. It is a result of the concert action of all three attractive energies, Coulomb, dispersion and induction.

Now we will discuss the question whether the interaction of two dihalogen molecules, each possessing the σ-hole, can be explained or at least schematically interpreted on the basis of the monomer electrostatic potential. Further, whether it would be possible to explain the preferential binding of the L-shaped structure within this concept. Figure 3.15 schematically shows the LS and TS structures together with the ESP of isolated molecules. Evidently, the first structure exhibits a strong attractive electrostatic interaction between the most positive σ-hole of the upper (vertical) dihalogen and the less positive belt of the lower (horizontal) one (the so-called dihalogen bond). Such an interaction is not expressed as strongly in the T-shaped structure.

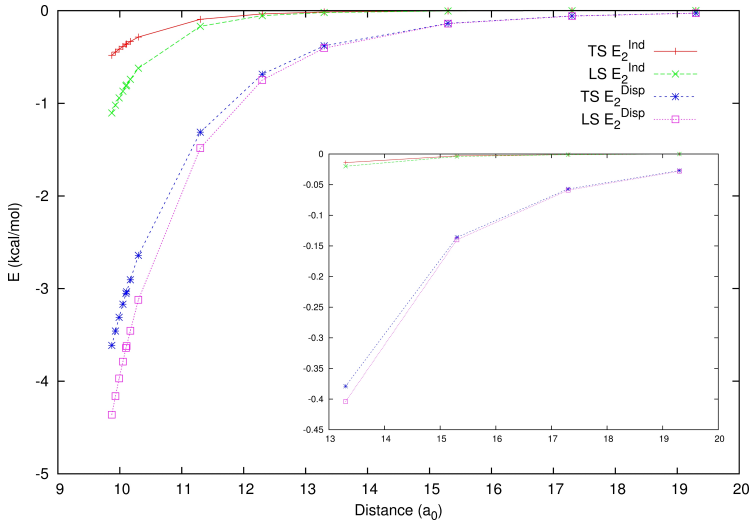

Figure 3.14: Distance dependence of the induction (E_2^{Ind}) and dispersion (E_2^{Disp}) energies for the LS and TS structures of the diiodine dimer.

Thus, it is possible to state that the preferential binding of the LS structure in all the dihalogen dimers investigated can be interpreted by the dihalogen bonding. The classical concept of electric multipoles leads to wrong conclusion, which is preferential binding of the TS structure.

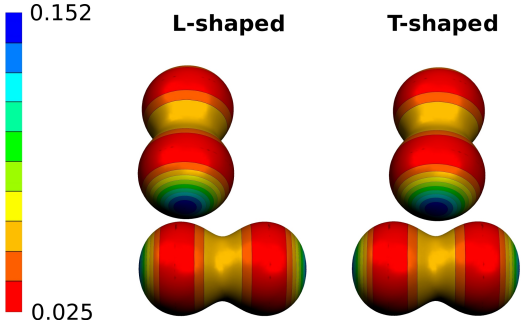

Figure 3.15: Electrostatic interaction based on the electrostatic potentials (ESP, in au) of isolated monomers for L-shaped (left) and T-shaped (right) structures of the X_2 dimer.

The situation with dinitrogen not possessing the σ-hole is different. Here, the total stabilization energy at the CCSD(T)/CBS level is comparable for the LS, TS, and PD structures. The L and P structures are less stable (*cf.*

Figure 3.11). Analyzing the DFT-SAPT energies, we found that dispersion energy is clearly a dominant stabilization term for all five conformers (*cf.* Figure 3.12), while Coulomb energy is marginal.

3.5 On the Origin of the Substantial Stabilization of the DTCA, DABCO$\cdots$I$_2$ complexes

3.5.1 Introduction

Complexes containing halogens participating in halogen bonding (X-bonding) are characterized by large stability, mostly comparable with the stabilization of similar H-bonded complexes. Indeed, stabilization energies of complexes in the X40 data set[176] and H-bonded complexes from the S66 data set[169,171] are well comparable. In both cases, energy decomposition reveals at least one similar feature: the electrostatic energy playing a dominant role. A counterintuitive electrostatic attraction in the case of a X-Y$\cdots$D halogen bond, where Y is Cl, Br or I, X is an electronegative atom (mostly carbon) and D is an electron donor like oxygen, nitrogen or sulfur, is explained by the existence of a positive σ-hole on top of the halogen atom. The electrostatic attraction thus occurs between the positive σ-hole and a negative electron donor. In the case of an X-H$\cdots$D hydrogen bond, the electrostatic attraction is caused by the interaction between a positively charged hydrogen and a negatively charged electron donor. Further comparison of other energy contributions reveals that dispersion energy is more negative in X-bonds than in H-bonds, which is explained by the fact that halogen and electron donors, both having large polarizability, are close to each other. The last attractive energy component, induction energy, is mostly smaller than dispersion energy; in X- and H-bonds, it is comparable. Induction energy also contains not only classical multiple-induced multiple induction energy but also CT energy. The CT energy becomes important only if an electron donor effectively

interacts with an electron acceptor. This means that besides the highest-occupied molecular orbital (HOMO) of the donor and the lowest-unoccupied molecular orbital (LUMO) of the acceptor, there must also be a favorable overlap between these orbitals (*cf.* eqn. 3.1).
In this study, we investigated the crystal structures containing iodine: 1,3-dithiole-2-thione-4-carboxyclic acid (DTCA)$\cdots$$I_2$ and 1,4-diazabicyclo[2.2.2]-octane (DABCO)$\cdots$$I_2$, for which there is an evidence of surprisingly large stabilization energies.[393,394] Since both calculations found in literature were made at a lower theoretical level, it is not clear whether these surprising numbers are correct. In case they are, where does this large stabilization come from? Is it only caused by halogen bonding with heavy iodine or does CT play an important role here? The aim of the present study is to investigate in detail the nature of the interactions in the above-mentioned complexes. Further, we will also study complexes where I_2 is replaced by lighter halogens (Br_2, Cl_2), hetero dihalogen (IF) as well as other systems (ICH_3, N_2).

3.5.2 Strategy of Calculation

3.5.2.1 Electrostatic Potentials and Quadrupole Moments

The electrostatic potentials were computed on molecular surfaces, with a surface being defined as the 0.001 au outer contour of the electron density. The most positive value of the potentials at the halogen (the local maximum) is referred to as $V_{S,max}$. Here, the electrostatic potentials as well as the geometries of electron acceptors and their electric quadrupole moments were calculated at the B97-D3/def2-QZVP level.[236]

3.5.2.2 Structures and Geometries

The coordinates of heavy atoms for both the I_2 complexes were taken from X-ray structures.[393,394] Afterwards, hydrogen atoms were manually added and subsequently optimized at the B97-D3/def2-QZVP level of theory, while keeping the coordinates of the heavy atoms frozen.

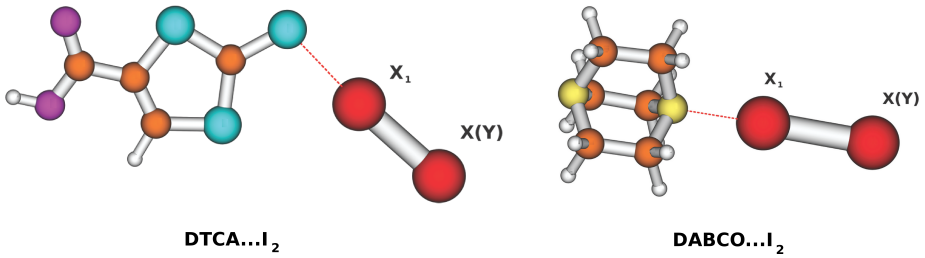

Figure 3.16: Structures of the DTCA$\cdots$I$_2$ and DABCO$\cdots$I$_2$ complexes.

When constructing the geometries of other binary complexes, the following procedure was utilized. Firstly, the coordinates of the DTCA and DABCO molecules were taken from structures of the respective I_2 complexes. Secondly, when the I_2 molecule was replaced from the corresponding I_2 complex structure by X_2 (X = Br, Cl, N) or XY (Y = F, CH_3) systems, the closer halogen atom X1 coincides with the closer iodine atom (*cf.* Figure 3.16). Finally, the rest of the electron acceptor molecule was constructed using the optimized geometry of the isolated acceptor, which was calculated at the B97-D3/def2-QZVP level.

3.5.2.3 Stabilization Energies

The benchmark stabilization energies were evaluated using the CCSD(T)/CBS method. Specifically, these stabilization energies were constructed as the sum of HF/CBS and MP2/CBS interaction energies. Both CBS energies were obtained via two point Helgaker extrapolation from aug-cc-pVDZ and aug-cc-pVTZ basis sets. The CCSD(T) correction term ($\Delta\Delta$CCSD(T)) was evaluated using aug-cc-pVDZ basis set.
The M06-2X functional was recommended for calculations of halogen-bonded complexes and it was also used in the present study.[392] Besides DFT-D (B97-D3/def2-QZVP), M06-2X/def2-QZVP calculations were also performed. All interaction energy calculations were corrected for the basis set superposition error (BSSE) utilizing counterpoise correction.[199]
Energy decomposition of the stabilization energies of all complexes was obtained by using the DFT-SAPT method. The subsystems were treated

using an asymptotically corrected PBE0AC exchange-correlation functional in combination with an aug-cc-pVDZ basis set.

3.5.3 Results and Discussion

3.5.3.1 Isolated systems

The LUMO energies of the acceptors are summarized in Table 3.10, which also contains the $V_{S,max}$ values and the quadrupole moment of all electron acceptors.

molecule	Q_{zz}	LUMO	$V_{S,max}$
I_2	3.402	-0.168	0.0474
Br_2	2.362	-0.173	0.0443
Cl_2	1.574	-0.162	0.0389
N_2	-0.774	-0.076	-0.0140
IF	0.273	-0.183	0.0896
ICH_3	3.217	-0.076	0.0214

Table 3.10: The Q_{zz} component of the quadrupole moment, LUMO and $V_{S,max}$ (all in au) for investigated monomers.

The electrostatic potentials of selected monomers are visualized in Figure 3.17.

Investigating the LUMO values, we find that IF, I_2 and Br_2 are the best acceptors. The Cl_2 systems are slightly worse and the N_2 molecule has the LUMO at higher energies, which agrees with the fact that N_2 is not an electron acceptor. The same is true for ICH_3 systems. The HOMO value for the electron donors DTCA and DABCO is -0.199 and -0.144 au, respectively; making DABCO the better electron donor. As expected, the magnitude of the σ-hole (see the $V_{S,max}$ value) for I_2 is larger than that of Br_2 and Cl_2. When the iodine was replaced by the more electronegative fluorine, the $V_{S,max}$ value increased considerably. The $V_{S,max}$ for N_2 is negative, which

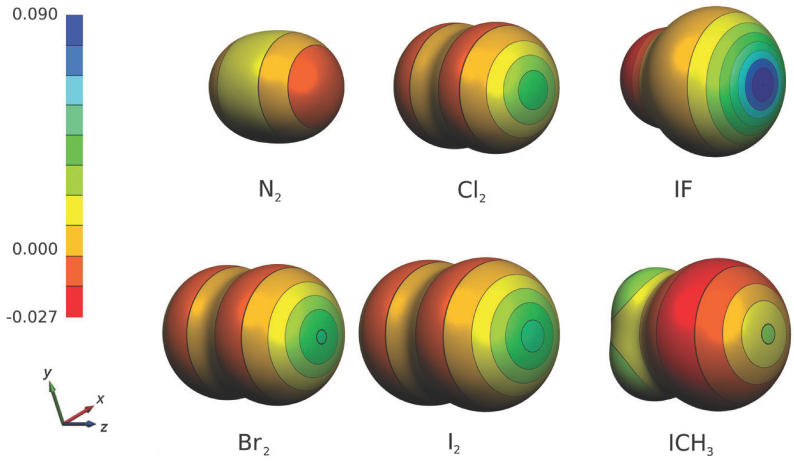

Figure 3.17: Electrostatic potential (in au) for all the monomers: I_2, Br_2, Cl_2, N_2, IF and ICH_3.

provides evidence that the positive σ-hole does not exist here. Comparing the quadrupoles of X_2 molecules, we find that they have different signs for halogens (I_2, Br_2, Cl_2) and nitrogen. The correlation between the $V_{S,max}$ and the quadrupole moment for the X_2 systems is very high, it amounts to R^2 = 0.902. This finding is surprising and helpful, because it tells us that for the explanation of the different binding of the halogens (Cl_2, Br_2, I_2) and the nitrogen to electron donors like O or N, it is not necessary to introduce a concept of the σ-hole, but it is enough to consider classical quadrupole moments. The electron donors with halogens exhibit attraction while the electron donors with nitrogen repulsion. This can be easily explained by the values of $V_{S,max}$ but comparably easily by quadrupole moments.

3.5.3.2 Complexes

Table 3.11 contains interaction energies determined for all complexes investigated using various computational techniques.

The B97-D3 stabilization energies for the complexes with halogens are very large; for DABCO complexes, they are even 40-70% larger. The larger stabilization of the DABCO complexes can be easily explained by the fact that DABCO is a better electron acceptor (see above). In both sets of complexes,

complex/term	B97-D3/ def2-QZVP	M06-2X/ def2-QZVP	MP2/ CBS	DFT-SAPT/ aug-cc-pVDZ	CCSD(T)/ CBS
DTCA$\cdots$ I_2	-13.80 (-5.38)	-8.81	-12.36	-5.98	-8.20
DTCA$\cdots$ Br_2	-11.25 (-4.45)	-6.97	-10.27	-5.42	-7.21
DTCA$\cdots$ Cl_2	-8.34 (-3.82)	-3.40	-6.01	-1.36	-3.75
DTCA$\cdots$ N_2	10.11 (-2.61)	10.03	9.11	9.57	9.93
DTCA$\cdots$ IF	-29.27 (-5.01)	-25.10	-27.21	-20.81	-23.77
DTCA$\cdots$ ICH_3	4.33 (-5.20)	8.12	3.98	10.43	7.08
DABCO$\cdots$ I_2	-18.97 (-8.33)	-17.18	-20.31	-24.19	-15.01
DABCO$\cdots$ Br_2	-16.72 (-6.58)	-13.04	-17.21	-14.68	-13.31
DABCO$\cdots$ Cl_2	-13.79 (-5.39)	-8.47	-11.72	-5.68	-8.98
DABCO$\cdots$ N_2	14.55 (-3.44)	14.40	14.02	13.92	14.40
DABCO$\cdots$ IF	-28.22 (-8.03)	-26.94	-30.29	-31.98	-26.49
DABCO$\cdots$ ICH_3	-2.69 (-8.19)	-4.90	-0.89	-3.01	-4.21

Table 3.11: The B97-D3, M06-2X, MP2, DFT-SAPT and CCSD(T) interaction energies (in kcal.mol^{-1}) for DTCA and DABCO complexes.

dispersion energy (shown in Table 3.11, 1^{st} column in parentheses) is an important stabilization component, but it is not dominant. For further energy decomposition, see the DFT-SAPT calculations. As mentioned in the Introduction, the DFT stabilization energies for the CT complexes could be overestimated due to an improper description of the virtual space. The benchmark CCSD(T)/CBS stabilization energies are considerably smaller than the DFT ones (*cf.* Table 3.11). Considering all the complexes with attractive interaction, we found that the CCSD(T)/CBS stabilization energy forms on average 62% of the DFT-D stabilization energy for the DTCA complexes and 79% for the DABCO complexes. The CCSD(T)/CBS stabilization energies of the DABCO complexes are larger than those of the DTCA complexes (by 11-139%). Surprisingly accurate numbers are obtained using the M06-2X functional. R^2 between the M06-2X and the CCSD(T)/CBS energies for both sets of complexes DTCA and DABCO amount to 0.983 and 0.994. The MP2/CBS stabilization energies are systematically overestimated with respect to CCSD(T)/CBS values. The average relative overestimation for the DABCO and the DTCA complexes evaluates to 32% and 37%, respectively. The DFT-SAPT calculations provide stabilization energies smaller than the benchmark CCSD(T)/CBS values, but the correlation between both energies is quite close ($R^2 = 0.990$ and $R^2 = 0.923$ for DTCA and DABCO, respectively). The underestimation of the DFT-SAPT energies arises from the use of a small aug-cc-pVDZ basis set. Among various energies, dispersion energy is the most underestimated.[319] We, however, use the DFT-SAPT not for generating accurate total stabilization energies but for a mere decomposition of the total stabilization energies. Passing from iodine to chlorine, the stabilization energies of both complexes decrease, the drop between iodine and bromine is moderate, but it becomes larger between bromine and chlorine. The stabilization energies of chlorine complexes are considerably smaller than those of iodine complexes, but they are still substantial. A dramatic increase of the stabilization energies of both complexes occurs when I_2 is replaced (at the same geometry) by IF. The electronegative fluorine withdraws electrons from iodine, which results in a much larger magnitude of the σ-hole (see Table 3.10 and Figure 3.17). Consequently, the total stabilization ener-

gies also increase. On the other hand, when one iodine atom in the iodine molecule is replaced by an electron-donating CH_3 group, the $V_{S,max}$ decreases and the total stabilization energy decreases dramatically and even becomes repulsive. Very large stabilization energies of complexes with IF are also caused due to the fact that this molecule is the best electron acceptor among all the systems investigated (see Table 3.10). The replacement of an iodine molecule with a nitrogen molecule also results in larger repulsive interaction energy. Here again, a certain role is played by both effects (nitrogen is not a good electron acceptor and does not contain a positive σ-hole). These findings indicate that electrostatic and CT energies play an important role in the complexes investigated.

complex/term	E_{tot}	E_1^{Pol}	E_1^{Ex}	E_2^{Ind}	δHF	$E^{"CT"}$	E_2^{Disp}
DTCA$\cdots I_2$	-5.98	-42.06	59.00	-67.07	55.25	-11.82	-11.11
DTCA$\cdots Br_2$	-5.42	-25.82	36.97	-22.09	14.05	-8.03	-8.54
DTCA$\cdots Cl_2$	-1.36	-17.68	33.34	-4.37	-5.86	-10.23	-6.79
DTCA$\cdots N_2$	9.57	-7.28	23.01	-0.33	-1.45	-1.78	-4.39
DTCA$\cdots$IF	-20.81	-40.19	49.97	-66.41	45.86	-20.55	-10.05
DTCA$\cdots ICH_3$	10.43	-42.01	63.60	-66.01	66.10	0.09	-11.25
DABCO$\cdots I_2$	-24.19	-65.36	83.02	-93.51	66.57	-26.94	-14.91
DABCO$\cdots Br_2$	-14.68	-40.43	52.52	-31.82	16.73	-15.09	-11.29
DABCO$\cdots Cl_2$	-5.68	-28.34	46.47	-7.20	-7.91	-15.11	-8.95
DABCO$\cdots N_2$	13.92	-9.65	31.82	-0.71	-1.82	-2.54	-5.71
DABCO$\cdots$IF	-31.98	-63.24	71.35	-90.82	64.28	-26.54	-13.55
DABCO$\cdots ICH_3$	-3.01	-64.32	89.58	-92.07	78.97	-13.10	-15.17

Table 3.12: The DFT-SAPT/aug-cc-pVDZ interaction energies (in kcal.-mol^{-1}) for the DTCA and the DABCO complexes.

The individual energy terms from the DFT-SAPT calculations are shown in Table 3.12. The largest term (in the absolute value) is exchange-repulsion energy E_1^{Ex}, which indicates short intermolecular distances. Among attrac-

tive terms, the largest energy is electrostatic energy E_1^{Pol}. In both complexes, dispersion energy (E_2^{Disp}) is large but induction energy (E_2^{Ind}) is comparable or in some cases even larger. This is clearly a new phenomenon: in all of our previous studies on non-covalent complexes including X-bonded complexes, the induction energy was systematically the smallest attractive term.
In this paragraph we will discuss the magnitude of the dispersion interaction, electrostatic and induction will follow in the next. Comparing the value of E_2^{Disp} from DFT-SAPT and D3 from B97-D3 (*cf.* Table 3.12 and 3.11) it is obvious that E_2^{Disp} term is systematically more negative. The E_2^{Disp} term is larger (in absolute value) on average by 94% and 73% for DTCA and DABCO complexes, respectively. This is in contrast to magnitudes of whole stabilization energies (as discussed above). This counterintuitive result can be understood as a consequence of the vagueness, when defining the dispersion interaction within the framework of the DFT. Grimme's empirical correction to dispersion interaction (D3) tries to remove one of the most important drawbacks of the exchange-correlation functional in DFT, which is its inability to reproduce the dispersion interaction, but not only at the asymptotic region ($1/R^6$ dependence) but in the whole range of distances. However, we should keep in mind that the "local" or "semi-local" functional, such as B97, can cover some part of the dispersion interaction. When describing the medium-range attractive non-covalent interaction at the van der Waals distances, the intermolecular overlap is not negligible. Hence, we stress that Grimme's D3 correction represents only a part of the dispersion. On the other hand, the E_2^{Disp} term from DFT-SAPT, which is based on second order perturbation theory, represents a better approximation to exact dispersion. That is why, the E_2^{Disp} term covers a bigger portion of dispersion (i.e. more negative) than the empirical D3 correction. Finally, we would like to point out that the presented difference between the E_2^{Disp} term and the D3 dispersion correction is underestimated. This follows from the fact that the aug-cc-pVDZ basis set does not provide a sufficiently converged value of the E_2^{Disp} term. This term is underestimated roughly by 10-20% at this level.[319]

Regarding the magnitude of electrostatic and induction energies. First, we

will investigate the correlation between electrostatic energy and the values of $V_{S,max}$ on the one hand and the quadrupole moment of X_2 molecules on the other. Evidently, both correlations (R^2 values for DTCA: 0.635, 0.894; DABCO 0.673, 0.917) are high, showing again that the σ-hole as well as the quadrupole moment explain the significant electrostatic stabilization in X-bonded complexes. When passing from X_2 molecules to other electron acceptors (IF, ICH_3) for which the first non-zero multipole moment is the dipole moment, the correlation between investigated entities is expected to deteriorates. Specifically, the correlation between electrostatic energy and $V_{S,max}$ for all six electron acceptors and both electron donors is not very high (DTCA $R^2 = 0.404$; DABCO $R^2 = 0.441$).

As mentioned above, induction energy (E_2^{Ind}) includes the CT which depends on the overlap and ability of an electron donor to donate electrons and an electron acceptor to accept electrons. Hence, the extremely large values of the induction energy ($E_2{}^{Ind}$) can be seen as a consequence of the short intermolecular distance as well as good abilities of partner molecules for CT interaction. The correlations between the sum of second-order induction, exchange-induction and δHF terms, labeled here as $E^{"CT"}$ (*cf.* section 2.2.3), and the LUMO energy of the electron acceptors for DTCA and DABCO are comparable (DTCA $R^2 = 0.756$ and DABCO $R^2 = 0.593$). This tells us how important is the CT energy within the $E^{"CT"}$ energy term. We have seen above that the $V_{S,max}$ value does not correlate tightly with electrostatic energy. However, the correlation between the SAPT interaction energy and the $V_{S,max}$ value (DTCA $R^2 = 0.873$ and DABCO $R^2 = 0.910$) as well as between the CCSD(T)/CBS interaction energy and the $V_{S,max}$ value (DTCA $R^2 = 0.932$ and DABCO $R^2 = 0.950$) is high. Little worse correlation has been found between the SAPT interaction energy and the LUMO energy of the acceptor (DTCA 0.787 and DABCO 0.688) and the CCSD(T)/CBS energy and the LUMO energy of the acceptor (DTCA $R^2 = 0.782$ and DABCO $R^2 = 0.873$). Thus, it is possible to conclude that the stabilization of the investigated complexes can be quit accurately predicted based on the $V_{S,max}$ values as well as the LUMO energies of the electron acceptor. The total DFT-SAPT energy correlates best with the $E^{"CT"}$ energy: for the DTCA and DABCO

complexes $R^2 = 0.953$ and 0.935, respectively. Further, the correlation with electrostatic and dispersion energies is considerably worse. R^2 values for the electrostatic energy (E_1^{Pol}) equals to 0.168 and 0.5874 for DTCA and DABCO, respectively. In the case of dispersion energy (E_2^{Disp}) corresponding values are 0.119 and 0.501. Among the molecular characteristics of the acceptor, the $V_{S,max}$ value as well as the LUMO energy correlate best with the total interaction energies. Putting together this and previous conclusions, we can state that within the complexes investigated, the stabilization is determined mostly by the electrostatic and the CT interaction.

3.6 Differences in the Sublimation Energies of Benzene and Hexahalogenbenzenes

3.6.1 Introduction

Two dimer structures are supposed to coexist at the respective potential energy surface of the benzene dimer, the T-shaped structure and the parallel-displaced (PD) structure. The parallel C_{2h} structure, which was expected to be the global minimum (because of the maximal overlap, dispersion contribution), is actually penalized by the quadrupole-quadrupole (Q-Q) electrostatic interaction, which is repulsive here.[395] The Q-Q interaction becomes less repulsive or attractive in the case of PD and T-shaped structures, respectively. Evidently, the electrostatic energy plays an important role in the interaction of benzene molecules not only in the benzene dimer but also in crystalline and plastic-crystalline phases.[396,397] It is thus not surprising that there have been attempts to interpret the sublimation energy of the benzene crystal only in terms of electrostatic quadrupole energy.[398] The resulting sublimation energy of 10.7 kcal.mol^{-1} agreed exactly with the respective experimental value.[398]

When passing to hexahalogenbenzenes, the quadrupole moment remains the first nonzero multipole moment, and it is hence possible to expect that

the sublimation energies of hexahalogenbenzenes could be interpreted solely by the electrostatic Q-Q interaction. Table 3.13 shows the z-components (perpendicular to the carbon ring) of quadrupole moments Q, polarizabilities α, and sublimation energies E_{sub} of benzene and hexahalogenbenzenes. In the case of benzene, the x and y components are positive, while in the cases of hexafluoro (C_6F_6) and hexachlorobenzenes (C_6Cl_6), these quadrupole components are negative. Hexabromobenzene has again the components with the same sign as benzene. A quick inspection of the quadrupole moments and the respective sublimation energies in Table 3.13 reveals no correlation between them. The quadrupole moments of hexafluorobenzene and benzene have the opposite sign, but their absolute values are similar (the former is slightly larger). With respect to this fact, we could expect the sublimation energy of the C_6F_6 to be slightly larger than that of the benzene, which actually holds true (*cf.* Table 3.13). When passing from hexafluorobenzene to hexachlorobenzene, the situation is dramatically changed, and the quadrupole moment of the latter molecule is more than an order of magnitude smaller. The sublimation energy of C_6Cl_6, however, has increased. Evidently, the assumption that the sublimation energy of hexahalogenbenzenes is predominantly determined by the electrostatic quadrupole energy is not fulfilled, and other energy terms may also have their contribution. From Table 3.13 we can see a close correlation between the polarizabilities and the sublimation energies, which tells us that the dispersion energy plays an important role in the interaction between hexahalogenbenzenes because there is a direct connection between the molecular polarizability and dispersion forces.

In the case of benzene dimer (or the crystal), both the electrostatic and dispersion energies are important attractive energy terms. These two terms are thus responsible for the structure determination. The situation is exactly the same in the case of hexahalogenbenzenes. However, with hexachloro- and hexabromobenzenes, a new interaction motif appears. Specifically, the dihalogen bond is formed between two molecules of hexachlorobenzenes or hexabromobenzenes, namely, between a halogen, X_1 (Cl, Br, I), which is covalently bound to a less electronegative atom (e.g. carbon), and another

molecule	Q_{zz}	α	E_{sub}
C_6H_6	-6.59	56.23	10.7
C_6F_6	7.89	57.24	11.8
C_6Cl_6	0.25	120.63	23.8
C_6Br_6	-4.72	152.93	-

Table 3.13: The z component of the quadrupole moments (Q_{zz}, in au), polarizabilities (α, in $\mathring{A}^3$) and sublimation energies (E_{sub}, in kcal.mol^{-1}) of the C_6X_6 (X = H, F, Cl, Br).

halogen, X_2 (C-$X_1\cdots X_2$).[115] This counterintuitive interaction is explained by the fact that a halogen atom is not isotropically negatively charged but it has a region with a positive electrostatic potential located on its top. This region is usually called a σ-hole;[111] it is depicted in Figure 3.18 as the blue disk on the halogens in C_6Cl_6 and C_6Br_6.

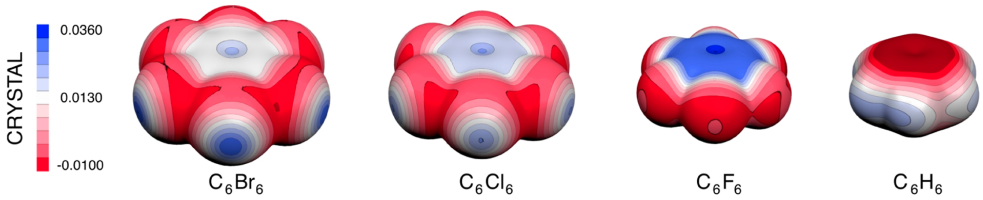

Figure 3.18: ESP of the C_6X_6 (X = H, F, Cl, Br) molecules.

The strength of the dihalogen bond is expected to increase with the atomic number of the halogens; in other words, the C-Cl$\cdots$Cl dihalogen bond is weaker than the C-Br$\cdots$Br or C-I$\cdots$I bonds. The σ-hole also exists at fluorine covalently bound to carbon, but this is typical only for small inorganic compounds such as NCF and not for aromatic species.[130] Consequently, the C-F$\cdots$F dihalogen bonds between two C_6F_6 are mostly impossible to form. It has to be added that in the case of the dihalogen bond, the dominant energy term is dispersion energy followed by electrostatic energy.[143]

The aim of the present study is to examine the nature of noncovalent binding within the crystals of C_6X_6 benzenes (X = H, F, Cl, Br). Specifically, we focused on identification of the binding motifs in various dimer structures appearing in these crystals. An attempt is made to correlate the experimental sublimation energy with the total interaction energies calculated for the crystal structures.

3.6.2 Strategy of Calculation

3.6.2.1 Structures

The X-ray structures of the hexahalogenbenzene crystals were obtained from the Cambridge Structural Database.[399,400] The X-ray structure of the benzene crystal[401] was obtained from the Crystallography Open Database (2100-0348.cif).[402] Within each crystal, the pairwise interactions were identified in the following manner: a reference molecule was chosen arbitrarily, and 20 pairs were created. Each pair contains the reference molecule and one of the 20 nearest neighbors.

3.6.2.2 Interaction Energies

The interaction energies for various dimers and for a large cluster, consisting of 21 molecules, were evaluated at the DFT (B3-LYP-D3) level using the TZVPP basis set and the empirical pairwise dispersion correction.[236] No deformation energy nor counterpoise correction was included. The interaction energy (ΔE) for a pair was determined via eqn. 2.6. Further, the energy of the central reference molecule E(1) and the energy of the cluster containing all but the central molecule E(20) were subtracted from the energy of the entire cluster E(21), providing the total interaction energy ΔE_{tot} (eqn. 3.2)

$$\Delta E_{tot} = E(21) - E(1) - E(20). \tag{3.2}$$

Finally, the average interaction energy (ΔE_{aver}) was evaluated according to eqn. 3.3

$$\Delta E_{aver} = [E(21) - 21.E(1)]/21, \tag{3.3}$$

where E(21) stands for the energy of the entire cluster and E(1) is the energy of the central reference molecule. The energy decomposition for all of the dimers was found by the DFT-SAPT method using the aug-cc-pVDZ basis set. It is a known fact that using the DFT-SAPT decomposition with an aug-cc-pVDZ basis set provides an unconverged dispersion (Disp) contribution, while the other contributions are converged, indeed, when compared with the complete basis set limit values.[319] Hence, the dispersion contribution was scaled by a factor that was calculated as follows. For the most stable dimers, we performed calculations with the aug-cc-pVTZ and aug-cc-pVDZ basis sets, and the scaling coefficients were obtained as the ratio between the dispersion term with the aug-cc-pVTZ and aug-cc-pVDZ basis sets. The coefficients are 1.42, 1.09, 1.09, and 1.12 for benzene, C_6F_6, C_6Cl_6, and C_6Br_6. The induction energy within DFT-SAPT approach, labeled as Ind (*cf.* Table 3.15), represents the sum of the induction, exchange-induction and δHF terms.

3.6.3 Results and Discussion

3.6.3.1 Interaction Energies

The Table 3.14 presents the total DFT-D3 interaction energies of the central reference molecule with the 20 neighboring molecules evaluated for four molecular crystals. Besides the total interaction energies, likewise, their DFT and dispersion components are listed. Table 3.14 also shows the average interaction energies, and also here, their DFT and dispersion components are presented.

The total interaction energies of benzene and C_6F_6 are almost equal, and also, the DFT and dispersion components are roughly comparable. These results

molecule	Total $E_{DFT+Disp}$	E_{DFT}	E_{Disp}	Average $E_{DFT+Disp}$	E_{DFT}	E_{Disp}
C_6H_6	-27.6	4.6	-32.2	-6.0	1.0	-7.0
C_6F_6	-27.9	1.0	-28.8	-10.5	-2.5	-8.0
C_6Cl_6	-45.6	19.3	-64.9	-13.5	5.5	-19.0
C_6Br_6	-61.6	24.3	-85.9	-17.6	7.3	-24.9

Table 3.14: The interaction energies (in kcal.mol^{-1}) of the central molecule with the 20 neighboring molecules (the Total columns) and the average interaction energies (the Average columns) for the clusters.

are not surprising regarding the molecular properties (*cf.* Table 3.13). However, the relatively large difference between the average interaction energies of C_6H_6 and C_6F_6 is surprising. This difference may arise from the symmetry of particular crystal structures. This issue will be discussed in more details below. When passing from C_6F_6 to C_6Cl_6 and C_6Br_6, a significant increase of the total stabilization energy and roughly the same increase of the average stabilization energy were found. In both cases, the dispersion contribution is much larger than that in the previous two crystals, and it is responsible for the total stabilization energy increase.

Table 3.15 shows the interaction energies for various pairs of benzene and hexahalogenbenzenes. The interaction energy is determined at the DFT-SAPT levels, and various pairs are ordered along decreasing stabilization energy; only the pairs with the stabilization energy higher than 1.0 kcal.mol^{-1} are presented.

Firstly, all dimers are mainly stabilized by dispersion and electrostatic interactions. Secondly, by comparing the pair interaction energies of C_6H_6 and C_6F_6 with those of C_6Cl_6 and C_6Br_6, we found an important difference. The stabilization energies for the most and least attractive pairs differ for the former two systems only marginally (by less than 1.2 kcal.mol^{-1}), while

molecule	bind. motif	deg.	DFT-D3 -ΔE	DFT-SAPT -Coul	-Ind	-Disp	-E_{int}
C_6H_6	T-shape	3	2.8 (3.3)	1.2	0.1	3.3	2.2
	distorted T-shape	3	2.0 (2.4)	1.0	0.0	2.4	1.5
	L-shape	3	1.6 (2.0)	0.4	0.0	1.9	1.2
C_6F_6	PD	1	3.3 (2.9)	1.4	0.1	3.7	2.4
	distant PD	0	3.3 (2.7)	1.7	0.1	3.7	2.4
	distorted T-shape	0	3.0 (2.8)	1.0	0.1	3.6	2.3
	distorted T-shape	0	2.7 (2.9)	1.0	0.1	3.7	1.9
	distorted T-shape	0	2.5 (2.4)	1.1	0.1	3.2	1.8
	distorted T-shape	0	2.4 (2.2)	1.1	0.1	2.9	1.7
	distorted T-shape	1	2.0 (1.7)	0.5	0.0	2.0	1.6
	distorted T-shape	1	1.9 (2.1)	0.7	0.0	2.8	1.2
C_6Cl_6	PD	1	11.5 (16.6)	6.0	0.4	19.6	9.7
	dihalogen bonded	1	2.0 (2.6)	1.2	0.1	3.8	2.1
	distorted T-shape	3	1.9 (2.4)	1.2	0.1	3.4	1.9
	distant PD	1	1.9 (2.9)	1.1	0.0	3.7	1.7
	distorted T-shape	3	1.6 (2.3)	1.3	0.1	3.4	1.6
C_6Br_6	PD	1	14.1 (20.9)	8.3	0.5	22.8	11.5
	dihalogen bonded	1	3.0 (3.8)	2.3	0.3	5.3	2.9
	distorted T-shape	3	2.7 (3.4)	2.3	0.3	4.8	2.7
	distorted T-shape	3	2.1 (3.1)	2.4	0.3	4.8	2.2
	distant PD	1	2.6 (4.1)	1.8	0.1	4.6	2.1

Table 3.15: The DFT-D3 and DFT-SAPT pair stabilization energies (-ΔE and -E_{int}) for the energetically most favorable pairs. The numbers in parentheses refer to the absolute value of the dispersion component of the DFT-D3 energy; for more details see Attachment **F**.

this difference is much more pronounced for the latter two systems (8.1 and 9.4 kcal.mol^{-1}, respectively). This difference can be documented also with the corresponding relative numbers. The relative increase of interaction from the weakest to the strongest dimer is 83, 100, 506, and 448% for C_6H_6 , C_6F_6, C_6Cl_6, and C_6Br_6, respectively.

The ratios of the dispersion and interaction energies (Disp/E_{int}) as well as of the Coulomb and interaction energies (Coul/E_{int}) provide a picture on the balance between the two most important attractive forces. The Coul/-E_{int} ratios averaged over the pairs with stabilization higher than 1 kcal.mol^{-1} are 0.51, 0.54, 0.68, and 0.89 for C_6H_6, C_6F_6, C_6Cl_6, and C_6Br_6, respectively. Clearly, the relative importance of the electrostatic contribution increases with the atomic number of the halogen. However, the value of neither the quadrupole (*cf.* Table 3.13) nor the Q-Q electrostatic interaction can interpret these ratios. An important increase of this ratio when passing from C_6H_6 and C_6F_6 to C_6Cl_6 and C_6Br_6 could be connected with the fact that a new binding motif is created in the latter group of crystals. Selected dimers of C_6Cl_6 and C_6Br_6 are stabilized by dihalogen bonds that do not exist in the former two crystals. The value of the Coul/E_{int} ratio for the dihalogen-bonded dimers of the C_6Cl_6 and C_6Br_6 molecules is even more pronounced. The values of 0.70 and 0.94 support our previous statement. Hence, the mere formation of dihalogen bonds in selected dimers of C_6Cl_6 and C_6Br_6 can explain the increase of the Coul/E_{int} ratios for the C_6Cl_6 and C_6Br_6 dimers. The different electrostatic potentials of C_6Cl_6 and C_6Br_6 with respect to the other two molecules, which is the reason for the formation of dihalogen-bond structures, may potentially be responsible for the increased value of the Coul/E_{int} ratio (*cf.* Figure 3.18). A more detailed view of the electrostatic potentials of all four molecules will be presented below. In Table 3.15, other relatively interesting features can be observed. The PD structure is either the most stable or one of the most stable dimer structures. When investigating the Coul DFT-SAPT energies for this structure, we found its dramatic increase for hexachloro- and hexabromobenzenes, which contradicts the decrease of the quadrupole moment when passing from C_6H_6 and

C_6F_6 to C_6Cl_6 and C_6Br_6. Visualizing the PD structures of all crystals (see Figure 3.19), we found that monomers in C_6Cl_6 and C_6Br_6 PD dimers are much closer to each other than those in the C_6F_6 dimer; the distance between the centers of mass of the C_6F_6, C_6Cl_6, and C_6Br_6 crystals amounts to 5.76, 3.76, and 3.95 Å, respectively.

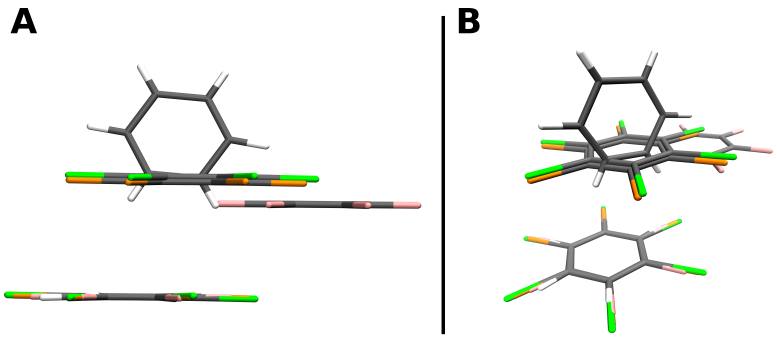

Figure 3.19: The most stable pair structures for benzene, hexafluorobenzene, hexachlorobenzene, and hexabromobenzene; silver = C, white = H, pink = F, orange = Cl, and green = Br; (A) side view; (B) perspective view.

A closer contact in the C_6Cl_6 and C_6Br_6 PD structures, which contradicts the larger vdW radii of Cl and Br than that of F, is clearly due to a very large dispersion energy (*cf.* Table 3.15). The penetration energy, defined as a difference between the SAPT electrostatic energy and multipole-multipole electrostatic energy, is negligible at the distances larger than equilibrium and becomes important (attractive) at shorter distances. Large SAPT electrostatic energies for the C_6Cl_6 and C_6Br_6 dimers are thus due to attractive penetration energies and have no connection with the Q-Q electrostatic energy. The Disp/E_{int} ratios averaged over the pairs with stabilization higher than 1 kcal.mol^{-1} are 1.56, 1.69, 2.00, and 1.99 for C_6H_6, C_6F_6, C_6Cl_6, and C_6Br_6, respectively. Unsurprisingly, the relative importance of the dispersion contribution is the lowest for C_6H_6 and the highest for C_6Cl_6 and C_6Br_6.

3.6.3.2 Structural Analysis

The differences in the binding motifs themselves, along with the different energetic degeneracy for all four molecular crystals, reveal that the relative

arrangement of the molecules in the cluster models is different (*cf.* Table 3.15). The highest degree of the energetic as well as binding motif degeneracy is exhibited by the benzene crystal. The 12 closest molecules that surround the central molecule are grouped into three structural motifs, each including four dimers (*cf.* the first part of Table 3.15). Several structural motifs can be recognized, T-shape, distorted T-shape, and L-shape (*cf.* Figure 3.20).

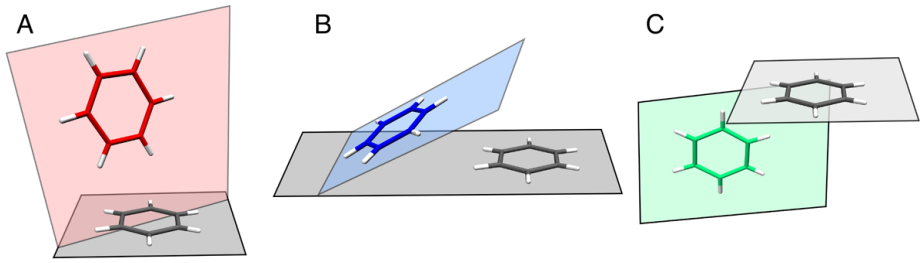

Figure 3.20: Structural motifs of benzene pairs found in the crystal. (A) T-shape, (B) distorted T-shape, and (C) L-shape.

The crystal of C_6F_6 possesses the lowest degree of structural motif and energetic degeneracy. The 11 neighboring molecules are divided into 8 groups (*cf.* the second part of Table 3.15). The three most stable dimers correspond to the PD structures. The structures of the remaining eight dimers can be classified as T-shape or distorted T-shape structures.

The crystals of C_6Cl_6 and C_6Br_6 are almost identical, hence possessing similar energetic and structural characteristics. The 14 neighboring molecules are divided into 5 groups. The most stable are two PD structures followed by two planar structures with two dihalogen bonds. As already mentioned above, dimers with dihalogen bonds are considerably less stable than the PD structures. Another two dimers represent a distant PD structure. The eight least stable dimers were included in the category of distorted halogen-bonded structures. However, they represent two distinct stabilization levels (*cf.* the third and fourth parts of Table 3.15).

One could expect that the similarity or the dissimilarity in the mutual arrangement of the neighboring molecules in the molecular crystals can be predicted for different chemical species based on the values of molecular properties, such as permanent multipole moments, polarizabilities, and so forth. However, the crystal structure analysis showed that such an assumption would lead to wrong interpretations. The structural differences between the crystals of C_6H_6 and C_6F_6 are remarkable, while the opposite is true when the crystals of C_6Cl_6 and C_6Br_6 are compared. Nevertheless, in the first example, the values of molecular properties are very similar, whereas in the second, there are significant differences (*cf.* Table 3.13). This leads us to the statement that more sophisticated approaches are necessary for the interpretation of the structural motif among noncovalently bound clusters.

In the next paragraphs, the geometrical parameters of individual dimers will be discussed. The most attractive pair of C_6H_6 is represented by the T-shape structure (*cf.* Figure 3.20 A), while the distorted T-shape and L-shaped structures (*cf.* Figure 3.20 B and C) are considerably less stable (by 29 and 43%, respectively). The situation with the remaining three hexahalogenbenzenes is different, and here, the most attractive pairs correspond to the PD structures. However, while the stabilization of the PD structure of C_6F_6 is comparable to that of the remaining structures, because of relatively large distance between the centers of mass (5.8 Å, *cf.* Figure 3.19). In the case of the chloro- and bromoderivates, the equivalent distance ranges between 3.8 and 4.0 Å, respectively, what is significantly smaller than equivalent distances for other conformers. Hence, for C_6Cl_6 and C_6Br_6 the PD structure is significantly more stable than the other structures. A further comparison of the most attractive PD structure for the three studied halogenbenzenes leads to the electrostatic term being larger for C_6Cl_6 and C_6Br_6 (than that for C_6F_6) by 4.6 and 6.9 kcal.mol^{-1}, respectively. This difference is, however, significantly larger (by 14.6 and 17 kcal.mol^{-1}) for the dispersion contribution. Consequently, it is mostly the dispersion energy for the PD structures that makes the total stabilization energy of C_6Cl_6 and C_6Br_6 much larger

than that of C_6F_6 (*cf.* polarizabilities of hexahalogenbenzenes presented in the Table 3.13).

Investigating other less stable pairs, we again found more pronounced differences between C_6H_6, C_6F_6, C_6Cl_6, and C_6Br_6. The three most stable structures of the second crystal possess a PD structure, while all of the others have a T-shaped structure. The crystals of hexachloro- and hexabromobenzenes differ from the crystals of benzene and hexafluorobenzene by the presence of structures possessing dihalogen bonds (*cf.* Figure 3.21).

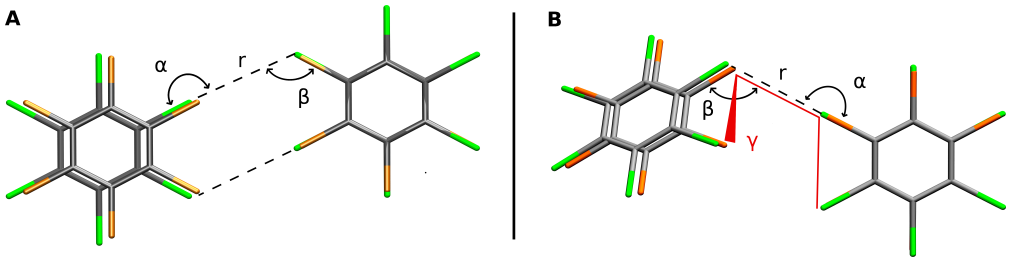

Figure 3.21: (A) Structure of the planar dihalogen-bonded dimer of hexachloro- and hexabromobenzene, with two ("cyclic") dihalogen bonds, (B) Structure of the distorted dihalogen-bonded dimer of hexachloro- and hexabromobenzene, with one dihalogen bond.

There are two structures with two ("cyclic") dihalogen bonds for each crystal with stabilization energies of 2.1 and 2.9 kcal.mol^{-1} for C_6Cl_6 and C_6Br_6, respectively. The C–X$\cdots$X angle (α) in these structures is, as expected, almost linear (171 and 173° for C_6Cl_6 and C_6Br_6, respectively), and the X$\cdots$X distance is 3.7 and 3.8 Å. The X$\cdots$X–C angle (β) is 123° for C_6Cl_6 and C_6Br_6 (*cf.* Figure 3.21 part A). Other dimer structures, named "distorted" dihalogen bonds, are not planar. One molecule is distorted from the imaginary plane (*cf.* Figure 3.21 part B); the dihedral angle $\gamma \neq 0°$. Hence, this structure contains only one dihalogen bond. We expected that due to rather short distance between heavy halogens, the stabilization energy of the structures with dihalogen bonds will be significantly higher. From the Table 3.15 it is, however, evident that these stabilization energies are only slightly larger than the stabilization energies of other structures. Investigating dif-

ferent structures of hexafluorobenzene, whose stabilization energy exceeds 1 kcal.mol^{-1}, we found neither planar nor distorted structures with a difluoro noncovalent bond. This is caused by the fact that fluorine covalently bound to an aromatic ring usually does not exhibit a σ-hole, which is a prerequisite for the existence of halogen bonding (*cf.* Figure 3.18). This significant difference between the electrostatic potential of C_6F_6 and C_6Cl_6 (together with C_6Br_6) crystals can be seen as the reason for the significant differences in the crystal structures. The region of the positive electrostatic potential (σ-hole), present at the top of each chlorine and bromine atom in a hexahalogenbenzene molecule, is the moiety via which the intermolecular interaction is realized (*cf.* Figure 3.21). Nevertheless, the stabilization energies of various hexafluorobenzene structures mostly having the T-shaped structure without a direct X$\cdots$X interaction are comparable to the stabilization energies of the structures possessing dihalogen bonds. It must be emphasized that no σ-hole$\cdots\pi$ interactions were found in the crystals. The positive σ-hole could be attracted by the negative π-electrons of the aromatic rings, but this is not the case of the C_6X_6 crystals, which are composed of one type of monomer only. The π-system of C_6Cl_6 and C_6Br_6 is a poor σ-hole acceptor. For mixed crystals, however, the σ-hole$\cdots\pi$ interactions could play a role.

Similar total interaction energies (1 + 20) of benzene and hexafluorobenzene agree with similar sublimation energies of these two crystals, and the much larger total interaction energy of hexachlorobenzene again agrees with its much larger sublimation energy. The relatively large difference in the average interaction energy of C_6H_6 and C_6F_6 (of as much as 75%) can be interpreted as a consequence of a different spatial arrangement of the pairs within the clusters considered. Cluster model of the C_6H_6 molecule is spherically less symmetric than in the case of C_6F_6 molecule, which means that the molecules around the central one are ordered less compactly. Hence, the average stabilization energy of the C_6H_6 molecule is substantially smaller.

The significant differences between the binding motif of the most stable dimer

of C_6H_6 and C_6F_6 crystals (T-shape and PD structure) can be seen as a consequence of a subtle difference in the electrostatic potential (*cf.* Figure 3.18). In the case of the C_6H_6 molecule, where the hydrogen atom regions are represented by a continuously increased positive potential (σ-hole), the T-shape conformer is energetically more preferred. The T-shape structure represent in fact the σ-hole$\cdots\pi$ interaction. On the other hand, the electrostatic potential of C_6F_6 in the regions of fluorine atoms does not show the same properties. Even though the fluorine atoms are surrounded by a negative region of the potential, an increase of the potential on top of each fluorine can be observed. This is a consequence of a mutual electron repulsion; hence, the T-shape structure is not as preferred as the PD structure.

3.7 Interactions of Boranes and Carboranes with Aromatic Systems

3.7.1 Introduction

Heteroboranes (substituted boron hydrides) can form two types of noncovalent interactions. The first are the thoroughly studied dihydrogen bonds (DHB, see section 1.1.3) of the B-H$\cdots$H-X type, which underlie their binding to biomolecules[92] or stabilize noncovalent complexes of substituted boranes in crystals.[61,78,403–405] The second type of interaction is stacking, that is, boranes and carboranes are located above aromatic rings while pointing their positively charged hydrogen to the center of the ring, thus forming a B-H$\cdots\pi$/C-H$\cdots\pi$ type of interaction. Such a structural motif has been observed in a few dicarba-borane nanostructures[406,407] and in a recent crystal structure of the n-$B_{18}H_{22}\cdots$benzene complex.[408] This "stacking" arrangement was subsequently studied computationally in a model diborane$\cdots$benzene complex (Figure 3.22 A).[409–411] The authors asserted that this $B_2H\cdots\pi$ weak hydrogen bond was of a dispersive nature and evaluated the stabilization energy using different optimization protocols at the CCSD(T) level at 4.3 kcal.mol^{-1} (ref. 409) or 2.45 kcal.mol^{-1} (ref. 410). Furthermore, a study em-

ploying a faster DFT approach (the M05-2X functional with the 6-311++G** basis set) calculated the stabilization energy of the diborane$\cdots$benzene complex as 3.43 kcal.mol^{-1} (ref. 411). This span in the energy values made us revisit this model system. The CCSD(T) method is known to provide accurate stabilization energies, but it is necessary to combine it with an extended basis set or, preferentially, to extrapolate these calculations to the complete basis set (CBS) limit. The CCSD(T) procedure yields accurate total interaction energies but does not provide insight into the nature of stabilization. The symmetry-adapted perturbation theory (SAPT) calculations, on the other hand, yield single-energy components, which are physically clearly defined and thus provide information on the interplay of the various terms. Moreover, the use of an extended basis set brings the SAPT interaction energies into a very close agreement with those from CCSD(T). An analysis of the interaction energy components was carried out in both previous studies to provide insight into the relative importance of various energy components. Li *et al.*[409] used a simple decomposition scheme based on a comparison of the CCSD(T) and HF energies and minima, whereas Tian *et al.*[410] employed the HF-based symmetry-adapted perturbation theory (SAPT). Both approaches showed a major contribution of dispersion energy (86%[409] or 170%[410] of the total interaction) and differed in the magnitude of other terms as well (for example, the contribution of electrostatics was assessed at $\sim$0 (ref. 409) or 2.68 (ref. 410) kcal.mol^{-1}). This span of the values available in the literature motivated us to determine a correct partitioning of the physically well-defined interaction-energy components using the novel density functional theory/symmetry adapted perturbation theory (DFT-SAPT) method.

The first aim of this work is an in-depth exploration of the diborane$\cdots$benzene interaction. The second is to investigate the effect of modifying either of the interacting partners. A similar task, which was however aimed at investigating the substituted diborane$\cdots$borazine complexes, has been carried out recently.[411] Here, we modified (i) benzene by heteroatom- and exosubstitutions or replaced it by cyclic aliphatic systems, and (ii) diborane, which was re-

placed by larger cage closo-borane and carba-closo-borane anions.

Cage carboranes have not only been found to form host-guest complexes stabilized by a C-H$\cdots\pi$ interaction[406,407] but also have been increasingly used as hydrophobic pharmacophores in molecular medicine.[412] As such, their ability to form noncovalent interactions with aromatic fragments of amino acids and nucleic acids has been studied here using benzene and its modifications as the model. The analysis of possible binding motifs and the DFT-SAPT interaction energy components broadens our understanding of the borane$\cdots$aromatic interaction.

3.7.2 Strategy of Calculation

3.7.2.1 Systems Studied

Noncovalent complexes of boranes or carboranes with aromatic and cyclic aliphatic systems have been studied. The starting model system was the diborane$\cdots$benzene complex (see Figure 3.22 A). The benzene moiety was replaced by both heteroatom- and exosubstituted aromatic compounds as well as cyclic aliphatic ones. Most of the molecules represent aminoacid side chains (phenol for tyrosine, imidazole for histidine, pyrrolidine for proline) or selected nucleic acid bases (cytosine). Furthermore, we have explored the effect of aromaticity/planarity by utilizing pairs of aromatic/planar and cyclic aliphatic systems: benzene-cyclohexane, 1,3-cyclopentadiene-cyclopentane, pyrrole-pyrrolidine. The diborane$\cdots$cyclopentane complex is shown in Figure 3.22 B. In other complexes with benzene, diborane was replaced by two icosahedral cages: the monoanionic $CB_{11}H_{12}^-$ carborane (C_{5v} symmetry) (*cf.* Figure 3.22, C, D) and the dianionic $B_{12}H_{12}^{2-}$ borane (I_h symmetry). Note that the $B_{12}H_{12}^{2-}$ dianion exhibits zero dipole moment because of its symmetry. The dipole moment of the monocarbaborane coincides with the space diagonal of this deformed icosahedron.

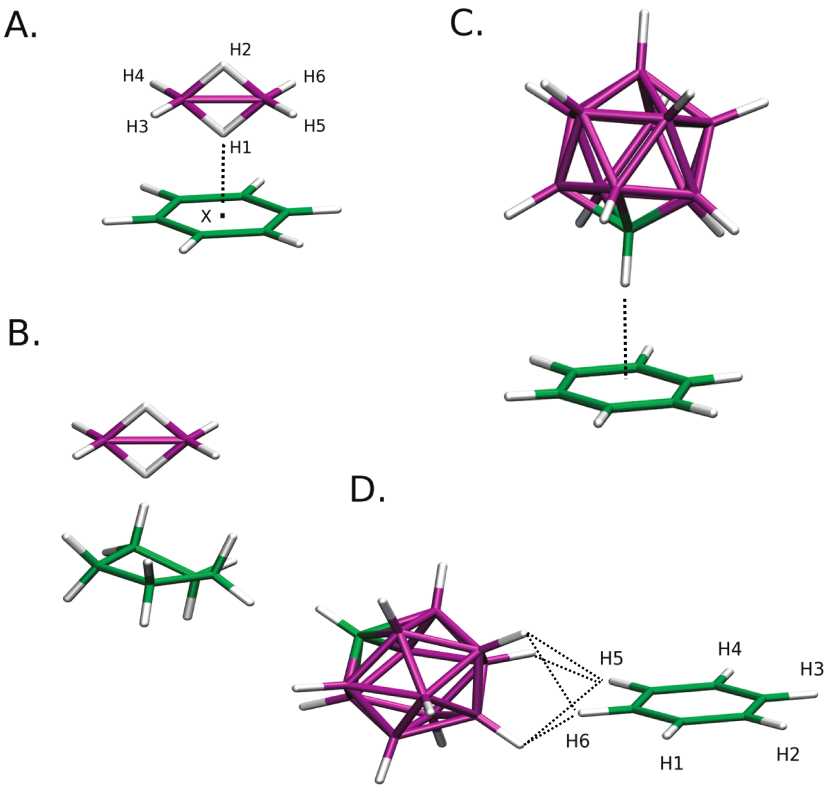

Figure 3.22: The optimized structures of the model complexes: (A) $C_6H_6\cdots B_2H_6$, (B) $C_5H_{10}\cdots B_2H_6$, (C) stacked $C_6H_6\cdots CB_{11}H_{12}^-$, and (D) planar $C_6H_6\cdots CB_{11}H_{12}^-$.

3.7.2.2 Structures

The initial structures of the investigated noncovalent complexes were built in a "stacked" motif (i.e., borane above the face of the ring). In the monocarbaborane cage complex, the C-H bond was positioned so as to point toward or away from the centroid of the benzene ring. Two important geometrical characteristics of the stacked structures are the distances between the centroid of the aromatic ring (X) and the hydrogen atom (H) pointing toward it or the non-hydrogen atom (Y) bound to it. We thus characterize the geometries of the complexes with the XH or XY distances (d_{XH}, d_{XY}). The models of the anion$\cdots$benzene interactions were investigated in a series of monoanions (F^-, Cl^-, and Br^-) and dianion SO_4^{2-}. All the complexes were built in a "planar" arrangement, the halide$\cdots$benzene geometries were taken from ref. 413.

3.7.2.3 Computational Details

Geometry optimizations of the complexes were carried out at the MP2/cc-pVTZ level of theory. Frozen core approximation was systematically used. Vibrational frequencies were calculated numerically at the above-mentioned level to confirm that the complexes represented true minima on the respective potential energy surfaces. The stabilization energies of the investigated complexes were determined at different levels of theory: MP2/aug-cc-pVD(T)Z, MP2/CBS, MP2.5/CBS,[232] DFT-SAPT/aug-cc-pVDZ,[319] DFT-SAPT/CBS, and CCSD(T)/CBS. The MP2 interaction energies were calculated with a medium-size aug-cc-pVDZ basis set, as the overestimation of stabilization energies caused by MP2 and underestimation owing to this basis set should approximately compensate for one another and give quite accurate results.[414] The MP2/aug-cc-pVDZ interaction energies were corrected for BSSE.[199] The DFT part, within DFT-SAPT calculations, was treated using the PBE0AC exchange-correlation functional with density fitting and the aug-cc-pVDZ basis set.[327,328] This combination of functional and basis set has been shown to provide a reasonably good description of electrostat-

ics and induction, with the dispersion component being underestimated by approximately 10-20%.[319] The use of a larger basis set for covering the dispersion term more appropriately is computationally too demanding for larger systems, so we performed this type of calculation with aug-cc-pVTZ only for the smallest model diborane$\cdots$benzene complex. This more precise calculation enabled an extrapolation toward the CBS limit (DFT-SAPT/CBS) for all the systems to be carried out. The contributions of individual DFT-SAPT terms toward the interaction energy are calculated as a percentage of the term in question from the total DFT-SAPT/CBS stabilization energy. The extrapolation to CBS in the DFT-SAPT approach (DFT-SAPT/CBS) was performed as follows. From the DFT-SAPT/aug-cc-pVDZ and DFT-SAPT/aug-cc-pVTZ results for the model diborane$\cdots$benzene complex, the Helgaker[161] two-point extrapolation of the dispersion energy (E_2^D + E_2^{Ex-D} terms) was made. Subsequently, this extrapolated dispersion energy (E_2^D/CBS) was combined with all the other terms from the DFT-SAPT/aug-cc-pVDZ calculation. The benchmark calculations were conducted at the CCSD(T)/CBS level using the extrapolation scheme defined by eqn. 2.8.

The NPA analysis[352,353] were carried out at the MP2/cc-pVTZ level of theory. The 1H NMR chemical shifts were calculated at the optimized geometry employing the gauge-including atomic orbitals (GIAO)[415] method at the B3-LYP level of theory with a Huzinaga's TZP basis set.[416,417] Electron densities were calculated at the B3-LYP/6-31G* level.

3.7.3 Results and Discussion

3.7.3.1 Diborane$\cdots$Benzene Complex

The optimized structure of the diborane$\cdots$benzene complex is depicted in the Figure 3.22 A. The novel type of a weak hydrogen-bond interaction (denoted here as $B_2H\cdots\pi$) take place in this complex. It is formed between the π-electron density of benzene and the bridging hydrogen (H1) of diborane. The H1 atom bears a slightly positive charge, unlike the terminal hydrogen

atoms H3-H6, which are slightly negatively charged.[92] The two subsystems are in a "stacking" arrangement (*cf.* Figure 3.22 A) with a characteristic H1$\cdots$X distance (where X is the centroid of the benzene ring), which amounts to 2.33 Å.

3.7.3.1.1 Benchmark Value of Interaction Energy

The benchmark CCSD(T)/CBS stabilization energy for diborane$\cdots$ benzene complex is -4.0 kcal.mol^{-1}. The DFT-SAPT/aug-cc-pVDZ approach underestimates the interaction by 1.2 kcal. mol^{-1}. However using the extrapolation scheme to pass from the aug-cc-pVDZ basis set to the CBS limit only for dispersion component of the energy yield stabilization of 3.6 kcal.mol^{-1}, which is in close agreement with the benchmark value. This result shows that the crucial role is played by the dispersion term (see also below). The MP2/CBS level overestimates the stabilization by 0.9 kcal.mol^{-1}. When passing to the MP2.5 method, the overestimation of MP2 is reduced and the agreement improves to a 0.3 kcal.mol^{-1} difference from the benchmark energy. This is a remarkable result given the feasibility of such a calculation in comparison with the CCSD(T)/CBS approach.

The capability of the DFT-SAPT/CBS method to provide reliable values of stabilization energies for more extended complexes was tested toward two larger complexes: $B_{12}H_{12}^{2-}\cdots$benzene and planar $CB_{11}H_{12}^{-}\cdots$benzene. The DFT-SAPT/CBS calculations reproduce the benchmark CCSD(T)/CBS values to within 10% for both complexes. However, the DFT-SAPT/aug-cc-pVDZ method underestimates the benchmark data by 10-30% (for more details see Table 2 in Attachment **G**). Hence, we can see that the correction of dispersion for the CBS limit in the DFT-SAPT approach is crucial for obtaining reliable results for the stabilization energy for this type of complexes.

We have calculated the benchmark CCSD(T)/CBS value of stabilization en-

ergy of 4.0 kcal.mol^{-1}. One of the previous studies was in close agreement (4.3 kcal.mol^{-1}),[409] whereas the other substantially underestimated the stabilization energy (2.45 kcal.mol^{-1}).[410] In another study, a faster DFT method (M05-2X/6-311++G**) yielded an interaction energy of 3.43 kcal.mol^{-1},[411] in fair agreement with the CCSD(T)/CBS value.

3.7.3.1.2 Interaction Energy Components

The DFT-SAPT decomposition of the stabilization energy for the diborane$\cdots$ benzene complex is presented in Table 3.16. It is clear that the complex is stabilized mainly by dispersion with a large contribution of 7.6 kcal.mol^{-1}. The second important stabilizing term is electrostatic, which yields 4.5 kcal.mol^{-1}. Another major term is exchange-repulsion, which opposes the interaction by 10.2 kcal.mol^{-1}. The last two terms are small; induction contributes 0.5 kcal.mol^{-1}, and δHF is 1.2 kcal.mol^{-1}. These rather small values of the induction and δHF energy terms indicate that the CT between the subsystems is relatively small. This supposition was independently confirmed by evaluating CT from the NPA analysis, which revealed a small transfer of 0.016 e from the electron-rich benzene to the electron-poor diborane.

The DFT-SAPT/CBS approach (presented in this work) slightly underestimated the benchmark value (a stabilization energy of 3.6 kcal.mol^{-1}). With such close agreement, we are confident about the interaction-energy components obtained using this method. The previous finding that the major stabilizing force is dispersion has thus been confirmed.[409–411] Quantitatively, the relative importance of this term was underestimated in the first study, where it accounted only for 86% of the total stabilization energy.[409] In the last two works[410,411] as well as in our current study, the E_2^D term was roughly twice as large as the total stabilization energy. Besides dispersion, the other stabilizing term according to the DFT-SAPT/CBS approach was electrostatic. Again, its relative importance was neglected in the former work ($\sim$0 kcal.mol^{-1}),[409] whereas in the later studies (this work and in refs.

405 and 406) its contribution has been calculated to be approximately 120% of the total stabilization. In summary, the presented DFT-SAPT/CBS calculations on the model diborane$\cdots$benzene complex agree quantitatively with the highly accurate CCSD(T)/CBS benchmark interaction energies. The energy components show a major stabilization caused by dispersion, followed by electrostatics.

3.7.3.2 Complexes of Diborane with Aromatic Systems

In this part we will focus our attention to study the effect of replacing benzene with other aromatic interaction partners of diborane. We chose several systems with either a heteroatom and/or exocyclic substitution that at the same time cover the basic building blocks of biomolecules (see Methods). The diborane$\cdots$borazine complex, which has recently been investigated computationally, is included for comparison. Table 3.16 lists the DFT-SAPT/CBS interaction energies for the benzene, borazine, pyrimidine, phenole, cytosine, and pyrrole complexes with diborane (B_2H_6).

complex	E_1^{Pol}	E_1^{Ex}	$E^{"CT"}$	E_2^{Disp}	$E_{tot/CBS}$
$B_2H_6\cdots$benzene	-4.5	10.2	-1.7	-7.6	-3.6
$B_2H_6\cdots$borazine	-2.3	5.6	-0.7	-5.3	-2.7
$B_2H_6\cdots$pyrimidine	-1.9	6.0	-0.9	-5.7	-2.5
$B_2H_6\cdots$phenol	-4.8	11.3	-2.0	-8.3	-3.7
$B_2H_6\cdots$cytosine	-3.3	9.6	-1.6	-7.7	-2.9
$B_2H_6\cdots$pyrrole	-4.9	10.6	-2.1	-7.2	-3.6

Table 3.16: DFT-SAPT/CBS interaction energies (in kcal.mol^{-1}) for the complexes of diborane with aromatic molecules.

The total values of the stabilization energies range from 2.5 to 3.7 kcal.mol^{-1}. It should be noted that the $d_{X\cdots H}$ distances range from 2.31 Å for phenol and pyrrole to 2.46 Å for pyrimidine. The diborane complexes with benzene, phenole, and pyrrole (which have smaller intersystem separation) exhibit

stronger binding (-3.6 to -3.7 kcal.mol^{-1}), whereas borazine, pyrimidine, and cytosine interact less strongly (-2.5 to -2.9 kcal.mol^{-1}). Like in the case of benzene, the main stabilization energy comes from the dispersion term followed by electrostatics. The drop in these two terms is mostly responsible for the differences in the total stabilization. Given the slight variations in the size of the systems, the effect of a heteroatom or exocyclic substitution of the aromatic system interacting with diborane is rather small.

3.7.3.3 Complexes of Diborane with Cyclic Aliphatic Systems

To investigate the effect of aromaticity/planarity of the subsystems treated in the previous section on the interaction with diborane, we have systematically examined their cyclic aliphatic counterparts, yielding the following pairs: benzene - cyclohexane, 1,3-cyclopentadiene - cyclopentane, and pyrrole - pyrrolidine. The DFT-SAPT/CBS interaction energy components for these complexes are shown in Table 3.17. Overall, the leading stabilizing term is dispersion followed by electrostatics. A comparison of the binding of the aromatic compounds on the one hand and the aliphatic ones on the other reveals that the stabilization energies as well as all of their energy components of the former group are systematically larger than the latter (*cf.* Table 3.17).

complex	E_1^{Pol}	E_1^{Ex}	$E^{"CT"}$	E_2^{Disp}	$E_{tot/CBS}$
$B_2H_6\cdots$benzene	-4.5	10.2	-1.7	-7.6	-3.6
$B_2H_6\cdots$cyclohexane	-1.0	3.6	-0.5	-4.0	-1.9
$B_2H_6\cdots$1,3-cyclopentadiene	-4.5	9.6	-2.0	-7.1	-3.9
$B_2H_6\cdots$cyclopentane	-1.4	4.7	-0.5	-5.0	-2.3
$B_2H_6\cdots$pyrrole	-5.6	11.1	-2.5	-7.6	-4.5
$B_2H_6\cdots$pyrrolidine	-6.0	10.1	-1.8	-6.2	-3.9

Table 3.17: DFT-SAPT/CBS interaction energies (in kcal.mol^{-1}) for the complexes of diborane with cyclic aliphatic molecules; a - stands for aromatic (planar) molecule; b - stands for aliphatic molecule.

The main difference between these two groups (with the exception of the last, nitrogen-containing molecular pair) is the large decrease in the E_1^{Pol} term for the aliphatic compounds. Thus, the interaction with the aromatic compounds is stabilized by both electrostatics and dispersion, whereas the interaction with the aliphatic molecules is governed mainly by dispersion. The energy differences are also connected with a different type of binding. Although the aromatic compounds interact with diborane in a stacking arrangement via their π-electron density involving a weak hydrogen bond ($B_2H\cdots\pi$ interaction, *cf.* Figure 3.22 A), their aliphatic counterparts bind to diborane via a weak van der Waals interaction (*cf.* Figure 3.22 B). The former interaction type is characterized by the position of the B–H1–B bridge over the center of the aromatic ring with the X–H1 distances ranging from 2.31 Å for phenol and pyrrole to 2.46 Å for pyrimidine (*cf.* Figure 3.22 A, with 2.33 Å for benzene optimized at the BSSE corrected MP2 level).[410] The latter van der Waals interaction is characterized by maximizing the number of $H\cdots H$ contacts between the subsystems where the $H\cdots H$ distances are equal or slightly larger than the sum of their atomic van der Waals radii (2.4 Å for hydrogens plus a cutoff of 0.3 Å).[418] Indeed, in the complexes studied here, the $H\cdots H$ distances range from 2.37 to 2.60 Å. Overall, the above-mentioned type of van der Waals interaction contacts is characterized by a large dispersion contribution to the total stability.[419–421] In the case of pyrrolidine, we additionally observe a short distance of 2.46 Å between the bridging hydrogen (H1) of diborane and the nitrogen atom of pyrrolidine. The geometry suggests an electrostatic interaction between the slightly positively charged H1 atom with the nitrogen lone-pair. The actual increase of the E_1^{Pol} term (in its absolute value) in comparison with pyrrole (Table 3.17) supports this idea.

In summary, we find that the aromatic compounds investigated in this study interact up to twice as strongly with diborane than their aliphatic counterparts. From the geometries and DFT-SAPT results, it is evident that there are two types of interactions of diborane stabilized by a different interplay of energy terms: (i) "stacking" interaction with aromatic compounds via a weak $B_2H\cdots\pi$ hydrogen bond, characterized by stabilizing dispersion, and

electrostatic terms; and (ii) van der Waals complexes with aliphatic compounds, where dispersion is the major stabilizing term.

3.7.3.4 Complexes of Cage Borane and Carborane Anions with Benzene

3.7.3.4.1 Structures

To further investigate the borane$\cdots$aromatic type of interaction, we have exchanged diborane for the larger carba-closo-borane $CB_{11}H_{12}^-$ and closo-dodecaborane $B_{12}H_{12}^{2-}$ cages. It should be stressed that both these cages are anionic, which influences their binding motifs. The geometries of their optimized complexes with benzene are shown in Figure 3.22 C and D. The $CB_{11}H_{12}^-\cdots$benzene complex exhibits two minima at the MP2/cc-pVTZ level: (i) a stacking arrangement with a C-H$\cdots\pi$ hydrogen bond (the X$\cdots$H distance is 2.26 Å, *cf.* Figure 3.22 C) and (ii) a planar binding motif with five B-H$\cdots$H-C dihydrogen bonds spanning an interval of 2.50-2.61 Å (*cf.* Figure 3.22 D). The dianionic borane $B_{12}H_{12}^{2-}$ has shifted during the optimization process from a starting stacked to a planar structure with five C-H$\cdots$H-B dihydrogen bonds (analogous to the planar geometry of the carborane$\cdots$benzene complex) ranging from 2.28 to 2.45 Å. A less stable minimum (see bellow) of the $CB_{11}H_{12}^-\cdots$benzene complex exhibits a stacking binding motif with a C-H$\cdots\pi$ hydrogen bond (*cf.* Figure 3.22 C) reminiscent of typical C-H$\cdots\pi$ interactions between aliphatic and aromatic hydrocarbons. Although the cage is on the whole an anion, the charge is delocalized on its surface with a partial positive charge on the carbon-attached hydrogen of the carborane,[92] positioned so that it can interact with the π-electron density.

3.7.3.4.2 Interaction Energy Decomposition

Table 3.18 shows the components of the total stabilization energy of the benzene complexes with the cage carborane $CB_{11}H_{12}^-$ (in both, the stacked and planar geometries), the borane $B_{12}H_{12}^{2-}$ and diborane for comparison.

complex	E_1^{Pol}	E_1^{Ex}	$E^{"CT"}$	E_2^{Disp}	$E_{tot/CBS}$
$C_6H_6\cdots B_2H_6$[a]	-4.5	10.2	-1.7	-7.6	-3.6
$C_6H_6\cdots CB_{11}H_{12}^-$[a]	-3.2	10.5	-1.5	-9.7	-3.9
$C_6H_6\cdots CB_{11}H_{12}^-$[b]	-3.3	6.4	-2.5	-6.0	-5.3
$C_6H_6\cdots B_{12}H_{12}^{2-}$[b]	-6.5	11.9	-7.6	-7.9	-10.1

Table 3.18: DFT-SAPT/CBS interaction energies (in kcal.mol^{-1}) for the $C_6H_6\cdots B_2H_6$, $CB_{11}H_{12}^-$ (stacked and planar binding motifs) and $B_{12}H_{12}^{2-}$ complexes; a - stands for stacked binding motif; b - stands for planar binding motif.

Larger stabilization energies are expected for the benzene complexes with closo-(car)boranes as compared to the diborane owing to the size of the cages. The main source of stabilization in all the complexes is dispersion energy, followed by electrostatics. These terms have been acknowledged to stabilize C-H$\cdots\pi$ hydrogen bonding,[422,423] dihydrogen bonding,[76,92,405] and anion$\cdots$arene interactions.[424,425] The exchange repulsion term opposes binding, the least in the planar $CB_{11}H_{12}^-\cdots$benzene complex, due to its largest intersystem separation (center-of-mass$\cdots$center-of-mass distance of 6.2 Å as compared to 6.1 and 4.8 Å for the $B_{12}H_{12}^{2-}$ and the stacked $CB_{11}H_{12}^-$ complexes, respectively. The induction term contributes a little to the total stabilization except for the $B_{12}H_{12}^{2-}$ complex in which it is responsible for a significant portion of the stabilization (61%). The role of induction in noncovalent complexes, especially of anion$\cdots$benzene type, will be discussed in detail below. In all, carborane and borane cages adopt either stacking or planar binding modes upon interaction with benzene, and the complexes are stabilized by dispersion, electrostatics, and in one case also by induction energy terms.

3.7.3.4.3 Cage Carborane, Borane$\cdots$Benzene Planar Interactions

The more stable minimum of the $CB_{11}H_{12}^{-}\cdots$benzene complex as well as the only minimum found for the $B_{12}H_{12}^{2-}$ complex displays a "planar" binding motif with five B-H$\cdots$H-C dihydrogen bonds. Such structures are analogous to small anion (e.g., halides, nitrate, perchlorate)$\cdots$ benzene planar complexes that have been thoroughly studied.[424,425] The partially negatively charged boron-bound hydrogens of the carborane (borane) cage (*cf.* Figure 3.22 D) present interaction partners for the C-H groups of benzene, similarly as the halides or oxygens of the nitrate, perchlorate or sulfate anions do. We note that the cage carborane offers those hydrogens for the interaction that are the most distant from the cage C-H group, that is, the most negatively charged (*cf.* Figure 3.22 D).[92] The calculated H$\cdots$H distances in the carborane/borane$\cdots$benzene complexes of 2.3-2.6 Å compare well with the X(O)$\cdots$H distances in small anion complexes of 2.1-3.1 Å in which the larger atomic radius of X(O) must be taken into account.[424,425] In terms of the DFT-SAPT interaction energy components, the dispersion and electrostatic terms constitute the main attraction terms in the planar carborane$\cdots$benzene complex, whereas in the $B_{12}H_{12}^{2-}$ complex induction is just as important (*cf.* Table 3.18). The dispersion term is the leading attractive component, which makes up 79-114% of the total stabilization in the carborane/borane$\cdots$benzene complexes. This is due to the high polarizability of both interaction partners and their close proximity. The attractive electrostatic term (62-64% of the total stabilization) is due to the interactions between the negative charge of the cages and the quadrupole of benzene (*cf.* Figure 3.23 A) and also due to five B-H$\cdots$H-C dihydrogen bonds (*cf.* Figure 3.22 D). The strength of these interactions is enhanced in the $B_{12}H_{12}^{2-}$ complex by the electron density redistribution due to induction. The induction interaction in the $B_{12}H_{12}^{2-}\cdots$benzene complex (61% of the total stabilization) is exerted by the permanent monopole of $B_{12}H_{12}^{2-}$ on the polarizable benzene ring in the complex geometry. The electron density redistribution is apparent from the calculated 1H NMR chemical shifts (the H5 and H6 atoms are deshielded the most, by 3 ppm) and also from the visualization of difference electron-density maps (*cf.* Figure 3.23 B).

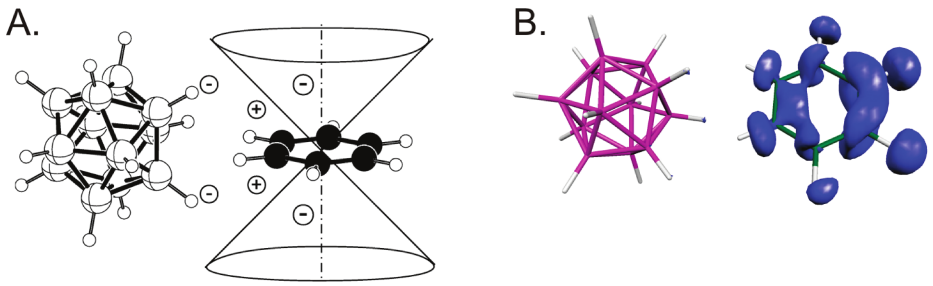

Figure 3.23: The $B_{12}H_{12}^{2-}\cdots$benzene interactions: (A) The electrostatic interactions; (B) The difference electron density of the complex describes the induction.

3.7.3.5 Induction in Planar Anion$\cdots$Benzene Complexes

Induction is not as usual source of stabilization in noncovalent complexes as electrostatics or dispersion, which drive, for example, the formation of hydrogen bonds or π-interactions, respectively.[419–421] Two examples of an interaction type where induction has gradually been acknowledged to play an important role are cation$\cdots\pi$[426–428] and anion$\cdots$arene[424] binding motifs. The contributions of the three major attractive terms (dispersion, electrostatics and induction) toward the stabilization of planar anion$\cdots$benzene complexes are plotted for the $CB_{11}H_{12}^{-}$ and $B_{12}H_{12}^{2-}$ cages and compared with halides and sulfate (*cf.* Figure 3.24).

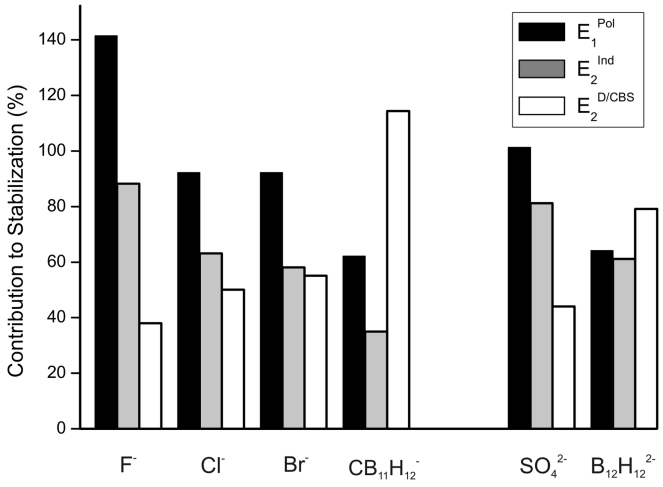

Figure 3.24: The relative importance of the individual attractive components in the total stabilization energy for the anion$\cdots$benzene complexes.

It can be seen that the percentages of the E_1^{Pol} and E_2^{Ind} terms decrease with the increasing size of the anions due to increasing intersystem separation and decreasing surface charge density, whereas the contribution of the E_2^D/CBS term increases. Altogether, the $B_{12}H_{12}^{2-}\cdots$benzene planar complex is stabilized by three major energy components, dispersion and electrostatics and moreover induction, the contributions of which are uniquely balanced (E_2^D/-CBS) = 79%, E_1^{Pol} = 64%, E_2^{Ind} = 61%).

Chapter 4

Conclusions

4.1 L7 Data Set

- The set of seven extended molecular complexes (called the L7 set), stabilized mostly by dispersion interaction is presented. We believe that considered complexes are reasonable representative of dispersion-dominated supramolecular associations in biological systems, although a few systems are partially stabilized by hydrogen bonding. The performance of the methods is evaluated by comparing them to high-level QCISD(T) (and CCSD(T)) results. These benchmark results, along with the geometries of the L7 complexes, are available on line in the BEGDB database (www.begdb.com).

- The best method in this study in absolute performance is MP2.5, delivering binding energies with a rRMSD of only 4%.

- The performance of the MP2.X method, parametrized toward noncovalent interactions, is surprisingly slightly worse.

- MP2C provides fairly accurate results, with an average relative error of 8%. It has a clear computational advantage over the MP2.5 by being an order of magnitude faster.

- The SCS-MP2 and SCS-(MI)-MP2 methods are rather disappointing, although they perform better for this test set than for the S66 data set of medium-sized noncovalent complexes. SCS-MP2 does not resolve fully the overbinding of MP2 for π stacking, while it underestimates dispersion between σ systems.

- Among density functional based methods, DFT-D3 is clearly superior to other approaches tested in this work. It represents the best trade-off between the accuracy and computational cost and, with rRMSD ranging from 9-17%, it outperforms some more sophisticated methods, such as SCS(MI)-MP2 or M06-2X.

- The accuracy of the semiempirical quantum mechanical methods, with empirical corrections for dispersion or hydrogen bonding is less satisfactory. The best are SCC-DF-TB-D and PM6-D3H4, and their relative standard deviation exceeds 25%. Nevertheless, the performance of these methods is not much worse than that of dramatically more expensive ones, such as M06-2X, MP2, or SCS(MI)-MP2, and thus their price/performance ratio is excellent.

4.2 Charge-Transfer Complexes

- The CT energy contribution is covered already at the HF level and the HF/MP2 ratio increases for strong CT complexes.

- The DFT-SAPT stabilization energy for all of the CT complexes studied is close to the CCSD(T)/CBS values and only exception represents the $NH_3 \cdots F_2$ complex. Agreement between DFT-SAPT and CCSD(T) stabilization energies even for complexes with substantial CT is surprisingly good and gives evidence that the former method can be safely used even for complexes with dative bond.

- The perturbation E_2^{CT} energies yield the same order of stabilization between studied electron donors and acceptors as the CCSD(T)/CBS

calculations, which supports the idea that the complexes studied are stabilized by the "CT" term.

- The DFT-SAPT analysis shown the dominant position of the first-order polarization (Coulomb) term, which is systematically the most important one. Dispersion energy is, for the complexes with relatively small CT (less than 0.05 e), the second most important term. For the complexes with the CT exceeding the value 0.05 e is the sum of induction and δHF terms covering the CT energy more attractive than dispersion term. The importance of the induction and the δHF terms covering the CT energy is systematically larger than in the H-bonded and stacked complexes of the S22 set.

- The MP2/CBS stabilization energy is mostly larger than the CCSD(T)/-CBS one, but the difference is rather small (less than 10%). This means that the computationally favorable MP2/CBS calculations can be safely used for this type of CT complexes, especially for strong type of CT complexes (those with dative bond).

4.3 On the Nature of the Stabilization of Benzene$\cdots$Dihalogen and Benzene$\cdots$Dinitrogen Complexes

- The quadrupole moments of dihalogens and dinitrogen have the opposite sign, which can be easily understood on the basis of the ESP of these systems. All of the dihalogens, including difluorine, possess a strong positive σ-hole.

- Benzene$\cdots X_2$ (X = Br, Cl) complexes are stronger than benzene$\cdots X_2$ (X = F, N) complexes. Analyzing the DFT-SAPT electrostatic, dispersion, induction, and δHF energies, we conclude that the former complexes are stabilized mainly by dispersion energy, followed by the

electrostatic term, whereas the latter are stabilized mainly by the dispersion interaction.

- The CT energy of the benzene$\cdots$dihalogen complexes is only moderate. This is true even for methylated benzene$\cdots$dihalogen complexes, for which a considerably larger role of the CT energy was expected.

- Benzene$\cdots$dichlorine and especially benzene (and methylated benzene) $\cdots$dibromine are stabilized mainly by dispersion and electrostatic energies; the role of CT energy is smaller. These complexes could thus be characterized as halogen-bonded ones rather than CT complexes.

4.4 Why Is the L-shaped Structure of $X_2\cdots X_2$ (X = F, Cl, Br, I) Complexes More Stable Than Other Structures?

- The most stable structure among the complexes of homodiatomics possessing the σ-hole is the L-shaped structure. For dinitrogen dimer (a system without the σ-hole), it is T-shaped structure.

- The stabilization in the case of homodihalogen complexes results from the existence of a dihalogen bond having comparable Coulomb and dispersion energies. The T-shaped structure of dinitrogen dimer is stabilized dominantly by the dispersion energy.

4.5 On the Origin of the Substantial Stabilization of the DTCA, DABCO$\cdots I_2$ complexes

- The CCSD(T)/CBS stabilization energies of the DTCA$\cdots I_2$ and DABCO$\cdots I_2$ CT complexes are very large, exceeding 8 and 15

kcal.mol^{-1}, respectively. The B97-D3/def2-QZVP stabilization energies of these complexes are strongly overestimated while the M06-2X/def2-QZVP energies agree with the benchmark values very well. DFT-SAPT stabilization energies are smaller than the benchmark values, which arise from the use of the aug-cc-pVDZ basis set which underestimates the dispersion interaction.

- The stabilization energies of both complexes decrease when passing from iodine to chlorine and dramatically increase when iodine is replaced by IF. When replacing halogen electron acceptors with ICH_3 or nitrogen, the stabilization energy strongly decreases and becomes repulsive.

- The total stabilization energies correlate well with the induction energy including the CT energy as well as with the $V_{S,max}$ value and the LUMO energy.

4.6 Differences in the Sublimation Energies of Benzene and Hexahalogenbenzenes

- Both the total and the average interaction energies increase when passing from benzene through hexafluorobenzene over hexachlorobenzene, and this increase is proportional to the increase of sublimation energy.

- The most stable dimer structure of benzene corresponds to the T-shaped structure, while that for hexahalogenbenzenes corresponds systematically to the PD structure. Because of the much higher polarizability of the hexachloro- and hexabromobenzene, the dispersion energy in this structure is also much higher than that in the hexafluorobenzene. The significant increase in the total interaction energy as well as that in the experimental sublimation energy when passing from hexafluorobenzene to hexachloro- and hexabromobenzene is thus mainly caused by the increase in dispersion energy. Indeed, the DFT-SAPT

decomposition shows that the dominant part of the interaction energy originates in the dispersion energy. Nevertheless, the relative importance of the electrostatic contribution increases when passing to heavier halogens.

- The new structural type, found in the crystals of hexachloro and hexabromobenzenes, is stabilized by dihalogen bonds. However, the stabilization energies of these structures do not differ much (except the PD structures) from the stabilization energies of other, mainly T-shaped structures of hexahalogenbenzene. The existence of the structures with dihalogen bonds thus cannot be responsible for the higher total interaction and sublimation energies of hexachloro- and hexabromobenzene. However, the presence of dihalogen bonds in hexachloro- and hexabromobenzenes is crucial for the determination of geometries of their crystals.

4.7 Interactions of Boranes and Carboranes with Aromatic Systems

- The benchmark interaction energy at the CCSD(T)/CBS level in the diborane$\cdots$benzene complex is -4.0 kcal.mol^{-1}.
- The DFT-SAPT/CBS approach yield a stabilization energy for the diborane$\cdots$benzene complex of 3.6 kcal.mol^{-1}, which is in close agreement with the benchmark energy. The major stabilization comes from dispersion (7.6 kcal.mol^{-1}), followed by electrostatics (4.5 kcal.mol^{-1}). The induction and δHF terms are negligible.
- Diborane interactions with aromatic systems are virtually unaffected by either the heteroatom or exosubstitutions of the benzene ring.
- Replacing aromatic systems with cyclic aliphatic ones results in a change of the binding motif and a roughly 50% drop in the stabilization energy. Specifically, the stacking $B_2H\cdots\pi$ interaction is replaced by van

der Waals contacts of the interacting hydrogen atoms. The source of stabilization comes in the former case from both dispersion and electrostatic terms, whereas the latter interactions are governed by dispersion only.

- Larger icosahedral cage carborane, monoanionic $CB_{11}H_{12}^-$, interacts with benzene employing two binding motifs: (i) The less stable minimum displays a stacking arrangement with a C-H$\cdots\pi$ hydrogen bond. The interaction is slightly stronger (DFT-SAPT/CBS value of -3.9 kcal.mol^{-1}) than for the diborane. Like the $B_2H\cdots\pi$ bonding of diborane, the leading stabilizing terms are dispersion, followed by electrostatics. (ii) The more stable minimum has a planar arrangement, reminiscent of small anion$\cdots$benzene complexes, with five bifurcated dihydrogen bonds of the B-H$\cdots$H-C type. Its interaction energy amounts to -5.3 kcal.mol^{-1} and assembly mainly of attractive dispersion and electrostatic terms.

- The dianionic $B_{12}H_{12}^{2-}$ molecule interacts with benzene only in the planar arrangement (and is repelled from the stacking one) with five bifurcated dihydrogen bonds of the B-H$\cdots$H-C type. The interaction energy amounts to -10.1 kcal.mol^{-1}, which is 2-3 times stronger than for the diborane$\cdots$benzene complex. The $B_{12}H_{12}^{2-}\cdots$benzene complex is stabilized by dispersion (79%), electrostatics (64%) and in addition by induction (61%). The balance of these three attractive components in stabilizing the $B_{12}H_{12}^{2-}\cdots$benzene complex is rather unique.

Bibliography

[1] Wilcken R.; Zimmermann M. O.; Lange A.; Joerger A. C.; Boeckler F. M. *J. Med. Chem.* **2013**, 56, 1363.

[2] Bissantz C.; Kuhn B.; Stahl M. *J. Med. Chem.* **2010**, 53, 5061.

[3] Gonnade R. G.; Shashidhar M. S.; Bhadbhade M. M. *J. Indian I. Sci. A-Eng.* **2007**, 87, 149.

[4] Burley S. K.; Petsko G. A. *Science* **1985**, 229, 23.

[5] Hunter C. A.; Singh J.; Thornton J. M. *J. Mol. Biol.* **1991**, 218, 837.

[6] Riley K. E.; Hobza P. *Acc. Chem. Res.* **2013**, 46, 927.

[7] Desiraju G. R.; Steiner T. *The Weak Hydrogen Bonds: In Structural Chemistry and Biology* Oxford University Press: Oxford New York, **1999**.

[8] Munshi P.; Guru Row T. N. *Crystallogr. Rev.* **2005**, 11, 199.

[9] Arunan E.; Desiraju G. R.; Klein R. A.; Sadlej J.; Scheiner S.; Alkorta I.; Clary D. C.; Crabtree R. H.; Dannenberg J. J.; Hobza P.; Kjaergaard H. G.; Legon A. C.; Mennucci B.; Nesbitt D. J. *Pure Appl. Chem.* **2011**, 83, 1637.

[10] Havlas Z.; Hobza P. *Chem. Rev.* **2000**, 100, 4253.

[11] Havlas Z.; Hobza P. *Theor. Chem. Acc.* **2002**, 108, 325.

[12] Dixon D. A.; Dobbs K. D.; Valentini J. J. *J. Phys. Chem.* **1994**, 98, 13435.

[13] Kim K. S.; Tarakeshwar P.; Lee J. Y. *Chem. Rev.* **2000**, 100, 4145.

[14] Taylor R.; Kennard O. *J. Am. Chem. Soc.* **1982**, 104, 5063.

[15] Shimoni L.; Carrell H. L.; Glusker J. P.; Coombs M. M. *J. Am. Chem. Soc.* **1994**, 116, 816.

[16] Choudhury A. R.; Guru Row T. N. *Cryst. Growth Des.* **2004**, 4, 47.

[17] Jeffrey G. A. *Crystallogr. Rev.* **2003**, 9, 135.

[18] Steiner T. *Crystallogr. Rev.* **2003**, 9, 177.

[19] Auffinger P.; Westhof E. *Biophys. J.* **1996**, 71, 940.

[20] Berger I.; Egli M.; Rich A. *Proc. Nat. Acad. Sci. U.S.A.* **1996**, 93, 12116.

[21] Bella J.; Berman H. M. *J. Mol. Biol.* **1996**, 264, 734.

[22] Musah R. A.; Jensen G. M.; Rosenfeld R. J.; McRee D. E.; Goodin D. B.; Bunte S. W. *J. Am. Chem. Soc.* **1997**, 119, 9083.

[23] Wahl M. C.; Sundaralingam M. *Trends Biochem Sci.* **1997**, 22, 97.

[24] Ninshio M *Cryst. Eng. Comm.* **2004**, 6, 130.

[25] Tatko C. D.; Waters M. L. *J. Am. Chem. Soc.* **2004**, 126, 2028.

[26] Brandl M.; Weiss M. S.; Jabs A.; Sühnel J.; Hilgenfeld R. *J. Mol. Biol.* **2001**, 307, 357.

[27] Umezawa Y.; Tsuboyama S.; Takahashi H.; Uzawa J.; Nishio M. *Bioorg. Med. Chem.* **1999**, 7, 2021.

[28] Yamakawa M.; Yamada I.; Noyori R. *Angew. Chem., Int. Ed.* **2001**, 40, 2818.

[29] Steiner T.; Koellner G. *J. Mol. Biol.* **2001**, 305, 535.

[30] Toth G.; Watts C. R.; Murphy R. F.; Lovas S. *Proteins Struct. Funct. Genet.* **2001**, 43, 373.

[31] Burley S. K.; Petsko G. A. *Adv. Protein Chem.* **1988**, 39, 125.

[32] Bhattacharyya R.; Samanta U.; Chakrabarti P. *Protein Eng.* **2002**, 15, 91.

[33] Kannan N.; Vishveshwara S. *Protein Eng.* **2000**, 13, 753.

[34] McGaughey G. B.; Gagne M.; Rappe A. K. *J. Biol. Chem.* **1998**, 273, 15458.

[35] Gilman A. G.; Rall T. W.; Mies A. S.; Taylor P. *The Pharmaceutical Basis of Therapeutics*, 8th Ed., McGraw Hill, Inc., New York, **1993**.

[36] Ishida T.; Doi M.; Ueda H.; Inoue M.; Scheldrick B. M. *J. Am. Chem. Soc.* **1988**, 110, 2286.

[37] Kamiichi K.; Danshita M.; Minamino N.; Doi M.; Ishidia T.; Inoue M. *FEBS Lett.* **1986**, 195, 57.

[38] Lu C. H.; Yang H. H.; Zhu C. L.; Chen X.; Chen G. N. *Angew. Chem., Int. Ed.* **2009**, 48, 4785.

[39] Shan C. S.; Yang H. F.; Song J. F.; Han D. X.; Ivaska A.; Niu L. *Anal. Chem.* **2009**, 81, 2378.

[40] Baby T. T.; Aravind S. S. J.; Arockiadoss T.; Rakhi R. B.; Ramaphabhu S. *Sens. Actuators. B* **2010**, 145, 71.

[41] Grimme S.; Antony J.; Schwabe T.; Muck-Lichtenfeld C. *Org. Biomol. Chem.* **2007**, 5, 741.

[42] Hunter A. C. *Angew. Chem.* **1993**, 105, 1653.

[43] Gung W. B.; Amicangelo C. J. *J. Org. Chem.* **2006**, 71, 9261.

[44] Tsuzuki S.; Honda K.; Uchimaru T.; Mikami M. *J. Chem. Phys.* **2006**, 124, 114304.

[45] Ahlrichs R.; Bar M.; Haser M.; Horn H.; Kolmel C. *Chem. Phys. Lett.* **1989**, 162, 165.

[46] Hobza P.; Selzle H. L.; Schlag E. W. *Chem. Rev.* **1994**, 94, 1767.

[47] Hobza P.; Selzle H. L.; Schlag E. W. *J. Phys. Chem.* **1993**, 97, 3937.

[48] Hobza P.; Selzle H. L.; Schlag E. W. *J. Am. Chem. Soc.* **1994**, 116, 3500.

[49] Watson D. M.; Fechtenkotter A.; Mullen K. *Chem. Rev.* **2001**, 101, 1267.

[50] Grimme S. *Angew. Chem. Int. Ed.* **2008**, 47, 3430.

[51] Kristyan S.; Pulay P. *Chem. Phys. Lett.* **1994**, 229, 175.

[52] Jaffe R. L.; Smith G. D. *J. Chem. Phys.* **1996**, 105, 2780.

[53] Sinnokrot M. O.; Valeev E. F.; Sherrill C. D. *J. Am. Chem. Soc.* **2002**, 124, 10887.

[54] Cybulski S. M.; Lytle M. L. *J. Chem. Phys.* **2007**, 127, 141102.

[55] Hesselmann A. *J. Chem. Phys.* **2008**, 128, 9.

[56] Zachariansen W. H.; Mooney R. C. L. *J. Chem. Phys.* **1934**, 2, 34.

[57] Burg A. B. *Inorg. Chem.* **1964**, 3, 1325.

[58] Brown M. P.; Heseltine R. W. *Chem. Commun.* **1968**, 1551.

[59] Brown M. P.; Heseltine R. W.; Smith P. A.; Walker P. J. *J. Chem. Soc. A* **1970**, 410.

[60] Brown M. P.; Walker P. J. *Spectrochim. Acta* **1974**, 30A, 1125.

[61] Richardson T. B.; de Gala S.; Crabtree R. H.; Siegbhan P. E. M. *J. Am. Chem. Soc.* **1995**, 117, 12875.

[62] Atwood J. L.; Koutsantonis G. A.; Lee F.-C.; Raston C. L. *Chem. Commun.* **1994**, 91.

[63] Campbell J. P.; Hwang J. W.; Young V. G.; Von Dreele R. B.; Cramer C. J.; Gladfelter W. L. *J. Am. Chem. Soc.* **1998**, 120, 521.

[64] Liu Q.; Hoffmann R. *J. Am. Chem. Soc.* **1995**, 117, 10108.

[65] Alkorta I.; Elguero J.; Foces-Foces C. *Chem. Commun.* **1996**, 1633.

[66] Remko M. *Mol. Phys.* **1998**, 94, 839.

[67] Kulkarni S. A. *J. Phys. Chem. A* **1998**, 102, 7704.

[68] Kulkarni S. A.; Srivastava A. K. *J. Phys. Chem. A* **1999**, 103, 2836.

[69] Lundell J.; Petterson M. *Phys. Chem. Chem. Phys.* **1999**, 1, 1601.

[70] Lee J. C.; Rheingold A. L.; Muller B.; Pregosin P. S.; Crabtree R. H. *Chem. Commun.* **1994**, 1021.

[71] Wessel J.; Lee J. C.; Peris E.; Yap G. P. A.; Fortin J. B.; Ricci J. S.; Sini G.; Albinati A.; Koetzle T. F.; Eisenstein O.; Rheingold A. L.; Crabtree R. H. *Angew. Chem., Int. Ed. Engl.* **1995**, 34, 2507.

[72] Patel B. P.; Wessel J.; Yao W; Lee J. C. Jr.; Peris E.; Koetzle T. F.; Yap G. P. A.; Fortin J. B.; Ricci J. S.; Sini G.; Albinati A.; Eisenstein O.; Rheingold A. L.; Crabtree R. H. *New J. Chem.* **1997**, 21, 413.

[73] Sini G.; Eisenstein O.; Yao W.; Crabtree R. H. *Inorg. Chim. Acta* **1998**, 280, 26.

[74] Belkova N. V.; Shubina E. S.; Ionidis A. V.; Epstein L. M.; Jacobsen H.; Messmer A.; Becke H. *Inorg. Chem.* **1997**, 36, 1522.

[75] Messmer A.; Jacobsen H.; Becke H. *Chem. Eur. J.* **1999**, 5, 3341.

[76] Custelcean R.; Jackson J. E. *Chem. Rev.* **2001**, 101, 1963.

[77] Cramer C. J.; Gladfelter W. L. *Inorg. Chem.* **1997**, 36, 5358.

[78] Klooster W. T.; Koetzle T. F.; Siegbahn P. E. M.; Richardson T. B.; Crabtree R. H. *J. Am. Chem. Soc.* **1999**, 121, 6337.

[79] Popelier P. L. A. *J. Phys. Chem. A* **1998**, 102, 1873.

[80] Nöth H.; Thomas S. *Eur. J. Inorg. Chem.* **1999**, 1373.

[81] Zottola M.; Pedersen L. G.; Singh P.; Shaw B. R. *"Borohydrides as Novel Hydrogen-Bond Acceptors" in Modeling the Hydrogen Bond, Proceedings from American Chemical Society Symposia Series, Computers in Chemistry, ed. D. A. Smith* **1994**, 569, 277.

[82] Singh P.; Zottola M.; Huang S.; Shaw B. R.; Pedersen L. G. *Acta. Crystallogr.* **1996**, C52, 693.

[83] Padilla-Martínez I. R.; Rosalez-Hoz M. de J.; Tlahuext H.; Camacho-Camacho C.; Ariza-Castolo A.; Contreras R. *Chem. Ber.* **1996**, 129, 441.

[84] Flores-Parra A.; Sánchez-Ruiz S. A.; Guadarrama C.; Nöth H.; Contreras R. *Eur. J. Inorg. Chem.* **1999**, 2069.

[85] Gusev D. G.; Lough A. J.; Morris R. H. *J. Am. Chem. Soc.* **1998**, 120, 13138.

[86] Abdur-Rashid K.; Gusev D. G.; Lough A. J.; Morris R. H. *Organometallics* **2000**, 19, 834.

[87] Lee J. C.; Peris E.; Rheingold A. L.; Crabtree R. H. *J. Am. Chem. Soc.* **1994**, 116, 11014.

[88] Ayllón J. A.; Gervaux C.; Sabo-Etienne S.; Chaudret B. *Organometallics* **1997**, 16, 2000.

[89] Epstein L. M.; Shubina E. S. *Ber Bunsen-Ges. Phys. Chem.* **1998**, 102, 359.

[90] Custelcean R.; Jackson J. E. *J. Am. Chem. Soc.* **1998**, 120, 12935.

[91] Hwang J. W.; Campbell J. P.; Kozubowski J.; Hanson S. A.; Evans J. F.; Gladfelter W. L. *Chem. Matter.* **1995**, 7, 517.

[92] Fanfrlík J.; Lepšík M.; Horinek D.; Havlas Z.; Hobza P. *ChemPhysChem* **2006**, 7, 1100.

[93] Cigler P.; Kozisek M.; Rezacova P.; Brynda J.; Otwinowski Z.; Pokorna J.; Plesek J.; Gruner B.; Doleckova-Maresova L.; Masa M.; Sedlacek J.; Bodem J.; Krausslich H. G.; Kral V.; Konvalinka J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 15394.

[94] Pecina A.; Lepsik M.; Rezac J.; Brynda J.; Mader P.; Rezacova P.; Hobza P.; Fanfrlik J. *J. Phys. Chem. B* **2013**, 117, 16096.

[95] Guthrie F. *J. Chem. Soc.* **1863**, 16, 239.

[96] Remsen I.; Norris J. F. *Am. Chem. J.* **1896**, 18, 90.

[97] Hassel O.; Rømming C. *Q. Rev. Chem. Soc.* **1962**, 16, 1.

[98] Bent H. A. *Chem. Rev.* **1968**, 68, 587.

[99] Murray-Rust P.; Motherwell W. D. S. *J. Am. Chem. Soc.* **1979**, 101, 4374.

[100] Dumas J.-M.; Kern M.; Janier-Dubry J. L. *Bull. Soc. Chim. Fr.* **1976**, 1785.

[101] Dumas J.-M.; Peurichard H.; Gomel M. J. *Chem. Res. (S)* **1978**, 54.

[102] Imakubo T.; Sawa H.; Kato R. *Synth. Met.* **1995**, 73, 117.

[103] Amico V.; Meille S. V.; Corradi E.; Messina M. T.; Resnati G. *J. Am. Chem. Soc.* **1998**, 120, 8261.

[104] Metrangolo P.; Neukirch H.; Pilati T.; Resnati G. *Acc. Chem. Res.* **2005**, 38, 386.

[105] Metrangolo P.; Meyer F.; Pilati T.; Resnati G.; Terraneo G. *Angew. Chem., Int. Ed.* **2008**, 47, 6114.

[106] Rissanen K. *J. CrystEngComm.* **2008**, 10, 1107.

[107] Legon A. C. *Phys. Chem. Chem. Phys.* **2010**, 12, 7736.

[108] Auffinger P.; Hayes F. A.; Westhof E.; Shing Ho P. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 16789.

[109] Murray J. S.; Riley K. E.; Politzer P.; Clark T. *Aust. J. Chem.* **2010**, 63, 1598.

[110] Politzer P.; Murray J. S. *ChemPhysChem* **2013**, 14, 278.

[111] Clark T.; Hennemann M.; Murray J. S.; Politzer P. *J. Mol. Model.* **2007**, 13, 291.

[112] Murray-Rust P.; Motherwell W. D. S. *J. Am. Chem. Soc.* **1979**, 101, 4374.

[113] Murray-Rust P.; Stallings W. C.; Monti C. T.; Preston R. K.; Glusker J. P. *J. Am. Chem. Soc.* **1983**, 105, 3206.

[114] Ramasubbu N.; Pathasarathy R.; Murray-Rust P. *J. Am. Chem. Soc.* **1986**, 108, 4308.

[115] Grabowski S. J. *J. Phys. Chem. A* **2012**, 116, 1838.

[116] Murray J. S.; Lane P.; Politzer P. *Int. J. Quantum Chem.* **2007**, 107, 2286.

[117] Murray J. S.; Lane P.; Clark T.; Riley K. E.; Politzer P. *J. Mol. Model.* **2012**, 18, 541.

[118] Tsirelson V. G.; Zou P. F.; Tang T.-H.; Bader R. F. W. *Acta. Crystallogr., Sect. A:Found. Crystallogr.* **1995**, 51, 143.

[119] Price S. L. *J. Chem. Soc. Faraday Trans.* **1996**, 92, 2997.

[120] Bilewicz E.; Rybarczyk-Pirek A. J.; Dubis A. T.; Grabowski S. J. *J. Mol. Struct.* **2007**, 829, 208.

[121] Delgado-Barrio G.; Prat R. F. *Phys. Rev. A: At., Mol., Opt. Phys.* **1975**, 12, 2288.

[122] Sen G. D.; Politzer P. *J. Chem. Phys.* **1989**, 90, 4370.

[123] Reed A. E.; Curtiss L. A.; Weinhold F. *Chem. Rev.* **1988**, 88, 899.

[124] Bader R. F. W.; Carroll M. T.; Cheeseman J. R.; Chang C. *J. Am. Chem. Soc.* **1987**, 109, 7968.

[125] Stewart R. F. *Chem. Phys. Lett.* **1979**, 65, 335.

[126] Politzer P.; Truhlar D. G. *Chemical Applications of Atomic and Molecular Electrostatic Potentials* Plenum Press, New York **1981**

[127] Politzer P.; Murray J. S.; Clark T. *Phys. Chem. Chem. Phys.* **2013**, 15, 11178.

[128] Politzer P.; Murray J. S.; Concha M. C. *J. Mol. Model.* **2007**, 13, 643.

[129] Chopra D.; Guru Row T. N. *CrystEngComm* **2011**, 13, 2175.

[130] Metrangolo P.; Murray J. S.; Pilati T.; Politzer P.; Resnati G.; Terraneo G. *Cryst. Growth Des.* **2011**, 11, 4238.

[131] Allen F. H. *Acta Crystallogr.* **2002**, B58, 380.

[132] Bernstein F. C.; Koetzle T. E.; Williams G. J. B.; Meyer J. E. F.; Brice M. D.; Rodgers J. R.; Kennard O.; Shimanouchi T.; Tasumi M. *J. Mol. Biol.* **1977**, 112, 535; http://www.rcsb.org/pdb/.

[133] Cody V.; Murray-Rust P. *J. Mol. Struct.* **1984**, 112, 189.

[134] Parisini E.; Metrangolo P.; Pilati T.; Resnati G.; Terraneo G. *Chem. Soc. Rev.* **2011**, 40, 2267.

[135] Lu Y.; Wang Y.; Zhu W. *Phys. Chem. Chem. Phys.* **2010**, 12, 4543.

[136] Xu Z.; Yang Z.; Liu Y.; Lu Y.; Chen K.; Zhu W. *J. Chem. Inf. Model.* **2014**, 54, 69.

[137] Cizek J. *J. Chem. Phys.* **1966**, 45, 4526.

[138] Paldus J.; Shavitt I.; Cizek J. *Phys. ReV. A* **1972**, 5, 50.

[139] Bartlett R. J. *Annu. Rev. Phys. Chem.* **1981**, 32, 359.

[140] Kucharski S. A. Noga J. Bartlett R. J., Watts J. D. *Chem. Phys. Lett.* **1990**, 165, 513.

[141] Bartlett R. J.; Musial M. *Rev. Mod. Phys.* **2007**, 79, 291.

[142] Møller Ch.; Plesset M. S. *Phys. Rev.* **1934**, 46, 618.

[143] Riley K. E.; Hobza P. *J. Chem. Theory Comput.* **2008**, 4, 232.

[144] Kolář M.; Hobza P. *J. Chem. Theory Comput.* **2012**, 8, 1325.

[145] Hardegger L. A.; Kuhn B.; Spinnler B.; Anselm L.; Ecabert R.; Stihle M.; Gsell B.; Thoma R.; Diez J.; Benz J.; Plancher J.-M.; Hartmann G.; Banner D. W.; Haap W.; Diederich F. *Angew. Chem., Int. Ed.* **2011**, 50, 314.

[146] Joergensen W. L.; Schyman P. *J. Chem. Theory Comput.* **2012**, 8, 3895.

[147] Wilcken R.; Zimmermann M.; Lange A.; Zahn S.; Boeckler F. *J. Comput-Aided. Mol. Des.* **2012**, 26, 935.

[148] Steiner T. *Angew. Chem., Int. Ed.* **2002**, 41, 49.

[149] Hardegger L. A.; Kuhn B.; Spinnler B.; Anselm L.; Ecabert R.; Stihle M.; Gsell B.; Thoma R.; Diez J.; Benz J.; Plancher J.-M.; Hartmann G.; Isshiki Y.; Morikami K.; Shimma N.; Haap W.; Banner D. W.; Diederich F. *ChemMedChem.* **2011**, 6, 2048.

[150] Wilcken R.; Liu X.; Zimmermann M. O.; Rutherford T. J.; Fersht A. R.; Joerger A. C.; Boeckler A. C. *J. Am. Chem. Soc.* **2012**, 134, 6810.

[151] Riley K. E.; Murray J. S.; Fanfrlik J.; Rezac J.; Sola R. J.; Concha M.; Ramos F.; Politzer P. *J. Mol. Model.* **2011**, 17, 3309.

[152] Metrangolo P.; Resnati G. *Chem.-Eur. J.* **2001**, 7, 2511.

[153] El Kerdawy A.; Wick C. R.; Hennemann M.; Clark T. *J. Chem. Inf. Model.* **2012**, 52, 1061.

[154] Cirtiss A. L.; Raghavachari K.; Trucks W. G.; Pople A. J. *J. Chem. Phys.* **1991**, 94, 7221.

[155] Goerigk L.; Grimme S. *J. Chem. Theory Comput.* **2010**, 6, 107.

[156] Zhao Y.; Truhlar D. G. *J. Chem. Theory Comput.* **2005**, 1, 415.

[157] Zhao Y.; Truhlar D. G. *J. Phys. Chem. A* **2005**, 109, 5656.

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

[159] Jurecka P.; Sponer J.; Cerny J.; Hobza P. *Phys. Chem. Chem. Phys.* **2006**, 8, 1985.

[160] Truhlar D. G. *Chem. Phys. Lett.* **1998**, 294, 45.

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

[162] Helgaker T.; Klopper W.; Koch H.; Noga J. *J. Chem. Phys.* **1997**, 106, 9639.

[163] Halkier A.; Helgaker T.; Jorgensen P.; Klopper W.; Olsen J. *Chem. Phys. Lett.* **1999**, 302, 437.

[164] Pitonak M.; Riley E. K.; Neograd y P.; Hobza P. *ChemPhysChem* **2008**, 9, 1636.

[165] Pitonak M.; Janowski T.; Neograd y P.; Pulay P.; Hobza P. *J. Chem. Theory Comput.* **2009**, 5, 1761.

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

[167] Takatani T.; Hohenstein E. G.; Malagoli M.; Marshall M. S.; Sherrill C. D. *J. Chem. Phys.* **2010**, 132, 144104.

[168] Riley E. K.; Hobza P. *J. Phys. Chem. A* **2007**, 111, 8257.

[169] Molnar L.; He X.; Wang B.; Merz M. K. *J. Chem. Phys.* **2009**, 131, 065102.

[170] Grafova L.; Pitonak M.; Rezac J.; Hobza P. *J. Chem. Theory Comput.* **2010**, 6, 2365.

[171] Rezac J.; Riley E. K.; Hobza P. *J. Chem. Theory Comput.* **2011**, 7, 2427.

[172] Grimme S. *Chemr Eur. J.* **2012**, 18, 9955.

[173] Sedlak R.; Janowski T.; Pitonak M.; Rezac J.; Pulay P.; Hobza P. *J. Chem. Theory Comput.* **2013**, 9, 3364.

[174] Valdes H.; Pluhackova K.; Pitonak M.; Rezac J.; Hobza P. *Phys. Chem. Chem. Phys.* **2008**, 10, 2747.

[175] Rezac J.; Hobza P. *J. Chem. Theory Comput.* **2013**, 9, 2151.

[176] Rezac J.; Riley E. K.; Hobza P. *J. Chem. Theory Comput.* **2012**, 8, 4285.

[177] Rezac J.; Jurecka P.; Riley E. K.; Cerny J.; Valdes H.; Pluhackova K.; Berka K.; Rezac T.; Pitonak M.; Vondrasek J.; Hobza P. *Collect. Czech. Chem. Commun.* **2008**, 73, 1261.

[178] Shin I.; Park M.; Min S. K.; Lee E. C.; Suh S. B.; Kim K. S. *J. Chem. Phys.* **2006**, 125, 234305.

[179] Lee E. C.; Kim D.; Jurecka P.; Tarakeshwar P.; Hobza P.; Kim K. S. *J. Phys. Chem. A* **2007**, 111, 3446.

[180] Min S. K.; Lee E. C.; Lee H. M.; Kim D. Y.; Kim D.; Kim K. S. *J. Comput. Chem.* **2008**, 29, 1208.

[181] Klopper W.; Samson C. C. M. *J. Chem. Phys.* **2002**, 116, 6397.

[182] Manby F. R. *J. Chem. Phys.* **2003**, 119, 4607.

[183] May A. J.; Valeev E.; Polly R.; Manby F. R. *Phys. Chem. Chem. Phys.* **2005**, 7, 2710.

[184] Noga J.; Kedzuch S.; Simunek J. *J. Chem. Phys.* **2007**, 127, 034106.

[185] Ten-No S. *Chem. Phys. Lett.* **2004**, 398, 56.

[186] Valeev E. F. *Chem. Phys. Lett.* **2004**, 395, 190.

[187] Ten-No S.; Manby F. R. *J. Chem. Phys.* **2003**, 119, 5358.

[188] Werner H. J.; Adler T. B.; Manby F. R. *J. Chem. Phys.* **2007**, 126, 164102.

[189] Fliegl H.; Hattig C.; Klopper W. *Int. J. Quantum Chem.* **2006**, 106, 2306.

[190] Fliegl H.; Klopper W.; Hattig C. *J. Chem. Phys.* **2005**, 122, 084107.

[191] Noga J.; Kedzuch S.; Simunek J.; Ten-No, S. *J. Chem. Phys.* **2008**, 128, 174103.

[192] Noga J.; Kutzelnigg W. *J. Chem. Phys.* **1994**, 101, 7738.

[193] Noga J.; Valiron P. *Chem. Phys. Lett.* **2000**, 324, 166.

[194] Shiozaki T.; Kamiya M.; Hirata S.; Valeev E. F. *Phys. Chem. Chem. Phys.* **2008**, 10, 3358.

[195] Shiozaki T.; Kamiya M.; Hirata S.; Valeev E. F. *J. Chem. Phys.* **2008**, 129, 071101.

[196] Tew D. P.; Klopper W.; Hattig C. *Chem. Phys. Lett.* **2008**, 452, 326.

[197] Tew D. P.; Klopper W.; Neiss C.; Hattig C. *Phys. Chem. Chem. Phys.* **2007**, 9, 1921.

[198] Marchetti O.; Werner H. *J. Phys. Chem. Chem. Phys.* **2008**, 10, 3400.

[199] Boys S. F.; Bernardi F. *Mol. Phys.* **1970**, 19, 553.

[200] Pople J. A.; Head-Gordon M.; Raghavachari K. *J. Chem. Phys.* **1987**, 87, 5968.

[201] Lee T. J.; Rendell A. P.; Taylor P. R. *J. Phys. Chem.* **1990**, 94, 5463.

[202] Janowski T.; Pulay P. *Chem. Phys. Lett.* **2007**, 447, 27.

[203] Scuseria G. E.; Schaefer H. F. *J. Chem. Phys.* **1989**, 90, 3700.

[204] Janowski T.; Ford A. R.; Pulay P. *Mol. Phys.* **2010**, 108, 249.

[205] Pitonak M.; Neogrady P.; Hobza P. *Phys. Chem. Chem. Phys.* **2010**, 12, 1369.

[206] Podeszwa R.; Szalewicz K. *J. Chem. Phys.* **2007**, 126, 194101.

[207] Rode M.; Sadlej J.; Moszynski R.; Wormer P. E. S.; van der Avoird A. *Chem. Phys. Lett.* **1999**, 314, 326.

[208] Rybak S.; Jeziorski B.; Szalewicz K. *J. Chem. Phys.* **1991**, 95, 6576.

[209] Boese A. D.; Martin J. M. L.; Klopper W. *J. Phys. Chem. A* **2007**, 111, 11122.

[210] Jurecka P.; Hobza P. *Chem. Phys. Lett.* **2002**, 365, 89.

[211] Min S. K.; Lee E. C.; Lee H. M.; Kim D. Y.; Kim D.; Kim K. S. *J. Comput. Chem.* **2008**, 29, 1208.

[212] Sherrill C. D.; Takatani T.; Hohenstein E. G. *J. Phys. Chem. A* **2009**, 113, 10146.

[213] Pitonak M.; Riley K. E.; Neograd y P.; Hobza P. *ChemPhysChem* **2008**, 9, 1636.

[214] Pitonak M.; Janowski T.; Neograd y P.; Pulay P.; Hobza P. *J. Chem. Theory Comput.* **2009**, 5, 1761

[215] Pittner J.; Hobza P. *Chem. Phys. Lett.* **2004**, 390, 496.

[216] Crittenden D. L. *J. Phys. Chem. A* **2009**, 113, 1663.

[217] Hopkins B. W.; Tschumper G. S. *J. Phys. Chem. A* **2004**, 108, 2941.

[218] Pitonak M.; Neograd y P.; Rezac J.; Jurecka P.; Urban M.; Hobza P. *J. Chem. Theory Comput.* **2008**, 4, 1829.

[219] Rezac J.; Simova L.; Hobza P. *J. Chem. Theory Comput.* **2013**, 9, 364.

[220] Grimme S. *J. Chem. Phys.* **2003**, 118, 9095.

[221] Antony J.; Grimme S. *J. Phys. Chem. A* **2007**, 111, 4862.

[222] Bachorz R. A.; Bischoff F. A.; Hofener S.; Klopper W.; Ottiger P.; Leist R.; Frey J. A.; Leutwyler S. *Phys. Chem. Chem. Phys.* **2008**, 10, 2758.

[223] Distasio R. A.; Head-Gordon M. *Mol. Phys.* **2007**, 105, 1073.

[224] King R. A. *Mol. Phys.* **2009**, 107, 789.

[225] Takatani T.; Sherrill C. D. *Phys. Chem. Chem. Phys.* **2007**, 9, 6106.

[226] Sherrill C. D.; Takatani T.; Hohenstein E. G. *J. Phys. Chem. A* **2009**, 113, 10146.

[227] Hill J. G.; Platts J. A. *J. Chem. Theory Comput.* **2007**, 3, 80.

[228] Takatani T.; Hohenstein E. G.; Sherrill C. D. *J. Chem. Phys.* **2008**, 128, 129901.

[229] Pitonak, M.; Rezac, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2010**, 12, 9611.

[230] Pitonak M.; Hesselmann A. *J. Chem. Theory Comput.* **2009**, 6, 168.

[231] Della Sala F.; Gørling A. *J. Chem. Phys.* **2001**, 115, 5718.

[232] Pitonak M.; Neogrady P.; Cerny J.; Grimme S.; Hobza P. *ChemPhysChem* **2009**, 10, 282.

[233] Riley E. K.; Rezac J.; Hobza P. *Phys. Chem. Chem. Phys.* **2011**, 13, 21121.

[234] Grimme S. *J. Comput. Chem.* **2004**, 25, 1463.

[235] Grimme S. *J. Comput. Chem.* **2006**, 27, 1787.

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

[237] Jurecka P.; Cerny J.; Hobza P.; Salahub D. R. *J. Comput. Chem.* **2007**, 28, 555.

[238] Dion M.; Rydberg H.; Scroder E.; Langreth D. C.; Lundqvist B. I. *Phys. ReV. Lett.* **2004**, 92, 246401.

[239] Sato T.; Tsuneda T.; Hirao K. *Mol. Phys.* **2005**, 103, 1151.

[240] Sato T.; Tsuneda T.; Hirao K. *J. Chem. Phys.* **2007**, 126, 234114.

[241] Vydrov O. A.; Van Voorhis T. *J. Chem. Phys.* **2009**, 130, 104105.

[242] Vydrov O. A.; Van Voorhis T. *Phys. Rev. Lett.* **2009**, 103, 063004.

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

[244] Zhao Y.; Schultz E. N.; Truhlar D. G. *J. Chem. Phys.* **2005**, 123, 161103.

[245] Zhao Y.; Schultz E. N.; Truhlar D. G. *J. Chem. Theory and Comput.* **2006**, 2, 364.

[246] Zhao Y.; Truhlar D. G. *Acc. Chem. Res.* **2008**, 41, 157.

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

[248] Schwabe T.; Grimme S. *Phys. Chem. Chem. Phys.* **2007**, 9, 3397.

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

[250] Chai J.-D.; Head-Gordon M. *J. Chem. Phys.* **2008**, 128, 084106.

[251] Yanai T.; Tew P. D.; Handy C. N. *Chem. Phys. Lett.* **2004**, 393, 51.

[252] Kaplan, I. *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials* John Wiley and Sons, Ltd., **2006**.

[253] Riley E. K.; Pitonak M.; Jurecka P.; Hobza P. *Chem. Rev.* **2010**, 110, 5023.

[254] Grimme S. *Chem-Eur. J.* **2004**, 10, 3423.

[255] Becke D. A. *J. Chem. Phys.* **1997**, 107, 8554.

[256] Schmider H. L.; Becke A. D. *J. Chem. Phys.* **1998**, 108, 9624.

[257] Axilrod M. B.; Teller E. *J. Chem. Phys.* **1943**, 11, 299.

[258] Muto J. *Proc. Phys. Math. Soc. Jpn.* **1943**, 17, 629.

[259] Tkatchenko A.; Alfé D.; Kim K. S. *J. Chem. Theory Comput.* **2012**, 8, 4317.

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

[261] Goerigk L.; Kruse H.; Grimme S. *ChemPhysChem.* **2011**, 12, 3421.

[262] Casimir H.; Polder D. *Phys. Rev.* **1948**, 73, 360.

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

[264] Becke A. D.; Johnson E. R. *J. Chem. Phys.* **2005**, 123, 154101.

[265] Johnson E. R.; Becke A. D. *J. Chem. Phys.* **2005**, 123, 024101.

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

[267] Johnson E. R.; Becke A. D. *J. Chem. Phys.* **2006**, 124, 174104.

[268] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, 102, 6.

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

[270] Sato, T.; Nakai, H. *J. Chem. Phys.* **2009**, 131, 224104.

[271] Iikura H.; Tsuneda T.; Yanai T.; Hirao K. *J. Chem. Phys.* **2001**, 115, 3540.

[272] Dewar M. J. S.; Zoebisch E. G.; Healy E. F.; Stewart J. J. P. *J. Am. Chem. Soc.* **1985**, 107, 3902.

[273] Stewart J. J. P. *J. Comput. Chem.* **1989**, 10, 221.

[274] Stewart J. J. P. *J. Comput. Chem.* **1989**, 10, 209.

[275] Dewar M. J. S.; Thiel W. *J. Am. Chem. Soc.* **1977**, 99, 4907.

[276] Dewar M. J. S.; Thiel W. *J. Am. Chem. Soc.* **1977**, 99, 4899.

[277] Martin B.; Clark T. *Int. J. Quantum Chem.* **2006**, 106, 1208.

[278] Thiel W.; Voityuk A. *Theor. Chim. Acta* **1992**, 81, 391.

[279] Thiel W.; Voityuk A. *Theor. Chim. Acta* **1996**, 93, 315.

[280] Thiel W.; Voityuk A. *J. Phys. Chem.* **1996**, 100, 616.

[281] Paton R. S.; Goodman J. *J. Chem. Inf. Model.* **2009**, 49, 944.

[282] Stewart J. J. P. *J. Mol. Model.* **2007**, 13, 1173.

[283] Stewart, J. J. P. *J. Mol. Model.* **2008**, 15, 765.

[284] Rezac J.; Fanfrlik J.; Salahub D. R.; Hobza P. *J. Chem. Theory Comput.* **2009**, 5, 1749.

[285] Korth M.; Pitonak M.; Rezac J.; Hobza P. *J. Chem. Theory Comput.* **2010**, 6, 344.

[286] Rezac J.; Hobza P. *J. Chem. Theory Comput.* **2012**, 8, 141.

[287] Stewart J. J. P. *J. Mol. Model.* **2013**, 19, 1.

[288] Korth M. *J. Chem. Theory Comput.* **2010**, 6, 3808.

[289] Rezac J.; Hobza P. *Chem. Phys. Lett.* **2011**, 506, 286.

[290] Hostas J.; Rezac J.; Hobza P. *Chem. Phys. Lett.* **2013**, 568-569, 161.

[291] London F. *Z. Phys. Chem. B* **1930**, 11, 222.

[292] London F. *Z. Phys.* **1930**, 63, 245.

[293] Eisenschitz R.; London F. *Z. Phys.* **1930**, 60, 491.

[294] London F. *Trans. Faraday Soc.* **1937**, 33, 8.

[295] Hirschfelder J. O.; Curtiss C. F.; Bird R. B. *Molecular Theory of Gases and Liquids* Wiley: New York, **1964**.

[296] Buckingham A. D. *Adv. Chem. Phys.* **1967**, 12, 107.

[297] Ahlrichs R. *Theor. Chim. Acta* **1976**, 41, 7.

[298] Morgan J. D., III.; Simon B. *Int. J. Quantum Chem.* **1980**, 17, 1143.

[299] Damburg R. J.; Propin R. Kh.; Graffi S.; Grecchi V.; Harrell E. M., II.; Cizek J.; Paldus J.; Silverstone H. *J. Phys. Rev. Lett.* **1984**, 52, 1112.

[300] Cizek J.; Damburg R. J.; Graffi S.; Grecchi V.; Hassels E. M.; Harris J. G.; Nakai S.; Paldus J.; Propin R. K.; Silverstone H. *J. Phys. Rev. A* **1986**, 33, 12.

[301] Hirschfelder J. O. *Chem. Phys. Lett.* **1967**, 1, 325.

[302] Holstein T. *J. Phys. Chem.* **1952**, 56, 832.

[303] Korona T.; Moszynski R.; Jeziorski B. *Adv. Quantum Chem.* **1997**, 28, 171.

[304] Claverie P. *Int. J. Quantum Chem.* **1971**, 5, 273.

[305] Adams H. W. *Int. J. Quantum Chem.* **1996**, 60, 273.

[306] Adams H. W. *Int. J. Quantum Chem.* **1999**, 72, 393.

[307] Patkowski K.; Korona T.; Jeziorski B. *J. Chem. Phys.* **2001**, 115, 1137.

[308] Jeziorski B.; Szalewicz K.; Chalasinski G. *Int. J. Quantum Chem.* **1978** 14, 271.

[309] Murrell N. J.; Shaw G. *J. Chem. Phys.* **1967**, 46, 1768.

[310] Musher N. J.; Amos T. A. *Phys. Rev.* **1967**, 164, 31.

[311] Hirschfelder O. J.; Silbey R. *J. Chem. Phys.* **1966**, 45, 2188.

[312] Patkowski K.; Jeziorski B.; Korona T.; Szalewicz K. *J. Chem. Phys.* **2002**, 117, 5124.

[313] Hirschfelder O. J. *Chem. Phys. Lett.* **1967**, 1, 363.

[314] van der Avoird A. *J. Chem. Phys.* **1967**, 47, 3649.

[315] Amos T. A.; Musher I. J. *Chem. Phys. Lett.* **1969**, 3, 721.

[316] Jeziorski B.; Szalewicz K.; Chalasinski G. *Int. J. Quantum Chem.* **1978**, 14, 271.

[317] Patkowski K.; Jeziorski B.; Szalewicz K. *J. Chem. Phys.* **2004**, 120, 6849.

[318] Jerziorski B.; Szalewicz K.; Mosyznski R. *Chem. Rev.* **1994**, 94, 1886.

[319] Hesselmann A.; Jansen G.; Schutz M. *J. Chem. Phys.* **2005**, 122, 014103.

[320] Jansen G.; Hesselmann A. *J. Phys. Chem. A* **2001**, 105, 646.

[321] Hesselmann A.; Jansen G. *Chem. Phys. Lett.* **2002**, 357, 464.

[322] Hesselmann A.; Jansen G. *Chem. Phys. Lett.* **2002**, 362, 319.

[323] Hesselmann A.; Jansen G. *Chem. Phys. Lett.* **2003**, 367, 778.

[324] Hesselmann A.; Jansen G. *Phys. Chem. Chem. Phys.* **2003**, 5, 5010.

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

[326] Perdew P. J.; Burke K.; Ernzerhof M. *Phys. Rev. Lett.* **1997**, 78, 1396.

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

[328] Gruning M.; Gritsenko V. O.; van Gisbergen A. J. S.; Baerends J. E. *J. Chem. Phys.* **2001**, 114, 652.

[329] Della Sala F.; Goorling A. *J. Chem. Phys.* **2001**, 115, 5718.

[330] Hesselmann A.; Manby F. *J. Chem. Phys.* **2005**, 123, 164116.

[331] Hesselmann A. *J. Phys. Chem. A* **2011**, 115, 11321.

[332] Bendova L.; Jurecka P.; Hobza P.; Vondrasek J. *J. Phys. Chem. B* **2007**, 111, 9975.

[333] Cerny J.; Hobza P. *Phys. Chem. Chem. Phys.* **2007**, 9, 5291.

[334] Pluhackova K.; Jurecka P.; Hobza P. *Phys. Chem. Chem. Phys.* **2007**, 9, 755.

[335] Wang W.; Hobza P. *ChemPhysChem* **2008**, 9, 1003.

[336] Sponer J.; Riley E. K.; Hobza P. *Phys. Chem. Chem. Phys.* **2008**, 10, 2595.

[337] Biedermannova L.; Riley E. K.; Berka K.; Hobza P.; Vondrasek J. *Phys. Chem. Chem. Phys.* **2008**, 10, 6350.

[338] Ran J.; Hobza P. *J. Chem. Theory Comput.* **2009**, 5, 1180.

[339] Berka K.; Hobza P.; Vondrasek J. *ChemPhysChem* **2009**, 10, 543.

[340] Ran J.; Hobza P. *J. Phys. Chem. B* **2009**, 113, 2933.

[341] Kolar M.; Berka K.; Jurecka P.; Hobza P. *ChemPhysChem* **2010**, 11, 2399.

[342] Morgado A. C.; Jurecka P.; Svozil D.; Hobza P.; Sponer J. *Phys. Chem. Chem. Phys.* **2010**, 12, 3522.

[343] Riley E. K.; Pitonak M.; Cerny J.; Hobza P. *J. Chem. Theory Comput.* **2010**, 6, 66.

[344] Zgarbova M.; Otyepka M.; Sponer J.; Hobza P.; Jurecka P. *Phys. Chem. Chem. Phys.* **2010**, 12, 10476.

[345] Rezac J.; Hobza P. *J. Chem. Theory Comput.* **2011**, 7, 685.

[346] Riley E. K.; Hobza P. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, 1, 3.

[347] Zierkiewicz W.; Wieczorek R.; Hobza P.; Michalska D. *Phys. Chem. Chem. Phys.* **2011**, 13, 5105.

[348] Adhikari U.; Scheiner S. *Chem. Phys. Lett.* **2012**, 532, 31.

[349] Mulliken S. R. *J. Chem. Phys.* **1955**, 23, 1833.

[350] Bickelhaupt M. F.; van eikema Hommes J. R. N.; Guerra F. C.; Baerends J. E. *Organometallics* **1996**, 15, 2923.

[351] Wiberg B. K.; Rablen R. P. *J. Comput. Chem.* **1993**, 14, 1504.

[352] Reed E. A.; Weinstock B. R.; Weinhold F. *J. Chem. Phys.* **1985**, 83, 735.

[353] Reed E. A.; Curtiss A. L.; Weinhold F. *Chem. Rev.* **1988**, 88, 899.

[354] Reed E. A.; Weinhold F. *J. Chem. Phys.* **1983**, 78, 4066.

[355] Bader W. F. R. *Atoms in Molecules - A quantum theory* **1990**, Oxford University Press, New York.

[356] Bayly I. C.; Cieplak P.; Cornell D. W.; Kollman A. P. *J. Phys. Chem.* **1993**, 97, 10269.

[357] Stone J. A. *Chem. Phys. Lett.* **1981**, 83, 233.

[358] Stone J. A.; Alderton M. *Mol. Phys.* **1985**, 56, 1047.

[359] Tough A. J. R.; Stone J. A. *J. Phys.* **1977**, A10, 1261.

[360] Stone J. A. *J. Chem. Theory Comput.* **2005**, 1, 1128.

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

[362] Elstner M.; Hobza P.; Frauenheim T.; Suhai S.; Kaxiras E. *J. Chem. Phys.* **2001**, 114, 5149.

[363] Janowski T.; Ford A. R.; Pulay P. *Mol. Phys.* **2010**, 108, 249.

[364] Hays F. A.; Teegarden A.; Jones Z. J.; Harms M.; Raup D.; Watson J.; Cavaliere E.; Ho P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 7157.

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

[366] Woon D. E.; Dunning T. H. Jr. *J. Chem. Phys.* **1993**, 98, 1358.

[367] Dunning T. H. Jr. *J. Chem. Phys.* **1989**, 90, 1007.

[368] Kroon-Batenburg L. M. J.; van Duijneveldt F. B. *J. Mol. Struct.* **1985**, 121, 185.

[369] van Lenthe J. H.; van Duijneveldt-van de Rijdt J. G. C. M.; van Duijneveldt F. B. *Adv. Chem. Phys.* **1987**, 69, 521.

[370] Hobza P.; Sponer. J. *Chem. Rev.* **1999**, 99, 3247.

[371] Svozil D.; Hobza P.; Sponer J. *J. Phys. Chem. B* **2010**, 114, 1191.

[372] Janowski T.; Pulay P. *J. Am. Chem. Soc.* **2012**, 134, 17520.

[373] Krapf S.; Koslowski T.; Steinbrecher T. *Phys. Chem. Chem. Phys.* **2010**, 12, 9516.

[374] Warshel A. *Acc. Chem. Res.* **2002**, 35, 385.

[375] Hervas M.; Navarro J.; Rosa M. D. L. *Acc. Chem. Res.* **2003**, 36, 798.

[376] Mallajosyula S.; Gupta A.; Pati S. *J. Phys. Chem. A* **2009**, 113, 3955.

[377] Halls J. J. M.; Walsh C. A.; Greenham N. C.; Marseglia E. A.; Friend r. H; Moratti S. C.; Holmes A. B. *Nature* **1995**, 376, 498.

[378] Yu G.; Gao J.; Hummelen J. C.; Wudi F.; Heeger A. J. *Science* **1995**, 270, 1789.

[379] Schmidt-Mende L.; Fechtenkotter A.; Mullen K.; Moons E.; Friend R. H.; Mackenzie J. D. *Science* **2001**, 293, 1119.

[380] Hoppe H.; Sariciftci N. S. *J. Mater. Res.* **2004**, 19, 1924.

[381] Stone A. J.; Misquita A. J. *Chem. Phys. Lett.* **2009**, 473, 201.

[382] Dabkowska I.; Jurecka P.; Hobza P. *J. Chem. Phys.* **2005**, 122, 204322.

[383] Jackel F.; Perera U. G. E.; Iancu V.; Braun K.-F.; Koch N.; Rabe J. P.; Hla S.-W. *Phys. Rev. Lett.* **2008**, 100, 126102.

[384] Chesnut D. B. *In Annual report on NMR spectroscopy; Webb G. A., Ed.; Academic Press: London* **1994**; 29, 89.

[385] Barbosa A. G. H.; Barcelos A. M. *Theor. Chem. Acc.* **2009**, 122, 51.

[386] Politzer P.; Laurence P. R.; Jayasuriya K. *Environ. Health Perspect.* **1985**, 61, 191.

[387] Murray J. S.; Politzer P. *J. Mol. Struct.: THEOCHEM* **1998**, 425, 107.

[388] Politzer P.; Murray J. S. *Trends Chem. Phys.* **1999**, 7, 157.

[389] Politzer P.; Murray J. S.; Peralta-Inga Z. *Int. J. Quantum Chem.* **2001**, 85, 676.

[390] Peterson K. A.; Figgen D.; Goll E.; Stoll H.; Dolg M. *J. Chem. Phys.* **2003**, 119, 11113.

[391] Peterson K. A. *J. Chem. Phys.* **2003**, 119, 11099.

[392] Kozuch S.; Martin J. M. L. *J. Chem. Theory Comput.* **2013**, 9, 1918.

[393] Deepa P.; Pandiyan B. V.; Kolandaivel P.; Hobza P. *Phys. Chem. Chem. Phys.* **2014**, 16, 2038.

[394] Peuronen A.; Valkonen A.; Kortelainen M.; Rissanen K.; Lahtinen M. *Cryst. Growth Des.* **2012**, 12, 4157.

[395] Battaglia M. R.; Buckingham A. D.; Williams J. H. *Chem. Phys. Lett.* **1981**, 78, 421.

[396] Shahin M.; Murthy S. S. N.; Singh L. P. *J. Phys. Chem. B* **2006**, 110, 18573.

[397] Cansell F.; Fabre D.; Petitet J. P. *J. Chem. Phys.* **1993**, 99, 7300.

[398] Sen P.; Basu S. *J. Chem. Phys.* **1968**, 48, 4075.

[399] Boden N.; Davis P. P.; Stam C. H.; Wesselink G. A. *Mol. Phys.* **1973**, 25, 81.

[400] Reddy C. M.; Kirchner M. T.; Gundakaram R. C.; Padmanabhan K. A.; Desiraju G. R. *Chem. Eur. J.* **2006**, 12, 2222.

[401] Budzianowski A.; Katrusiak A. *Acta Crystallogr., Sect. B* **2006**, 62, 94.

[402] *Crystallography Open Database* http://www.crystallography.net, **2013**.

[403] Planas J. G.; Vinas C.; Teixidor F.; Comas-Vives A.; Ujaque G.; Lledos A.; Light M. E.; Hursthouse M. B. *J. Am. Chem. Soc.* **2005**, 127, 15976.

[404] Glukhov I. V.; Lyssenko K. A.; Korlyukov A. A.; Antipin M. Y. *Russ. Chem. Bull.* **2005**, 54, 547.

[405] Belkova N. V.; Shubina E. S.; Epstein L. M. *Acc. Chem. Res.* **2005**, 38, 624.

[406] Blanch R. J.; Williams M.; Fallon G. D.; Gardiner M. G.; Kaddour R.; Raston C. L. *Angew. Chem., Int. Ed.* **1997**, 36, 504.

[407] Raston C. L.; Cave G. W. V. *Chem. Eur. J.* **2004**, 10, 279.

[408] Hamilton E. J. M.; Kultyshev R. G.; Du, B.; Meyers E. A.; Liu S. M.; Hadad C. M.; Shore S. G. *Chem. Eur. J.* **2006**, 12, 2571.

[409] Li H. Z.; Min D. H.; Shore S. G.; Lipscomb W. N.; Yang W. *Inorg. Chem.* **2007**, 46, 3956.

[410] Tian S. X.; Li H. B.; Bai Y. B.; Yang J. L. *J. Phys. Chem. A* **2008**, 112, 8121.

[411] Ravinder P.; Subramanian V. *J. Phys. Chem. A* **2010**, 114, 5565.

[412] Lesnikowski J. Z. *Collect. Czech. Chem. Commun.* **2007**, 72, 1646.

[413] Coletti C.; Re N. *J. Phys. Chem. A* **2009**, 113, 1578.

[414] Kolar M.; Hobza P. *J. Phys. Chem. A* **2007**, 111, 5851.

[415] Cheeseman J. R.; Trucks G. W.; Keith T. A.; Frisch M. J. *J. Chem. Phys.* **1996**, 104, 5497.

[416] Huzinaga S. *J. Chem. Phys.* **1965**, 42, 1293.

[417] Dunning T. H. *J. Chem. Phys.* **1970**, 53, 2823.

[418] Desiraju G. R. *Acc. Chem. Res.* **2002**, 35, 565.

[419] Hobza P.; Muller-Dethlefs K. *In Non-CoValent Interactions* The Royal Society of Chemistry: Cambridge, UK, **2010**.

[420] Muller-Dethlefs K.; Hobza P. *Chem. ReV.* **2000**, 100, 143.

[421] Hobza P.; Zahradnik R.; Muller-Dethlefs K. *Collect. Czech. Chem. Commun.* **2006**, 71, 443.

[422] Tsuzuki S.; Honda K.; Uchimaru T.; Mikami M.; Tanabe K. *J. Am. Chem. Soc.* **2000**, 122, 11450.

[423] Tsuzuki S.; Fujii A. *Phys. Chem. Chem. Phys.* **2008**, 10, 2584.

[424] Hay B. P.; Bryantsev V. S. *Chem. Commun.* **2008**, 21, 2417.

[425] Bryantsev V. S.; Hay B. P. *J. Am. Chem. Soc.* **2005**, 127, 8282.

[426] Cubero E.; Luque F. J.; Orozco M. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 5976.

[427] Tsuzuki S.; Yoshida M.; Uchimaru T.; Mikami M. *J. Phys. Chem. A* **2001**, 105, 769.

[428] Singh N. J.; Min S. K.; Kim D. Y.; Kim K. S. *J. Chem. Theory Comput.* **2009**, 5, 515.

List of Attached Publications

This thesis is based on following articles:

Attachment A:
Sedlak R.; Janowski T.; Pitonak M.; Rezac J.; Pulay P.; Hobza P. "Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes", *Journal of Chemical Theory and Computation* 2013, 9, 3364.

Attachment B:
Karthikeyan S.; **Sedlak R.**; Hobza P. "On the Nature of Stabilization in Weak, Medium and Strong Charge-Transfer Complexes: CCSD(T)/CBS and SAPT Calculations", *Journal of Physical Chemistry A* 2011, 115, 9422.

Attachment C:
Munusamy E.; **Sedlak R.**; Hobza P. "On the Nature of the Stabilization of Benzene (and Methylated-Benzene)$\cdots$Diatomic (Dihalogen and Dinitrogen) Complexes: CCSD(T)/CBS and DFT-SAPT calculations", *ChemPhysChem* 2011, 12, 3253.

Attachment D:
Sedlak R.; Deepa P.; Hobza P. "Why is the L-Shaped Structure of $X_2\cdots X_2$ (X = F, Cl, Br, I) Complexes More Stable Than Other Structures", *Journal of Physical Chemistry A* 2014, 118, 3846.

Attachment E:
Deepa P.; **Sedlak R.**; Hobza P. "On the Origin of the Substantial Stabilisation of the Electron-Donor 1,3-dithiole-2-thione-4-carboxyclic Acid$\cdots I_2$ and DABCO$\cdots I_2$ Complexes", *Physical Chemistry and Chemical Physics* 2014, 16, 6679.

Attachment F:
Trnka J.; **Sedlak R.**; Kolar M.; Hobza P. "Differences in the Sublimation Energy of Benzene and Hexahalogenbenzenes Are Caused by Dispersion Energy", *Journal of Physical Chemistry A* 2013, 117, 4331.

Attachment G:
Sedlak R.; Fanfrlik J.; Hnyk D.; Hobza P.; Lepsik M. "Interaction of Boranes with Aromatic Systems: CCSD(T) Complete Basis Set Calculations and DFT-SAPT Analysis of Energy Components", *Journal of Physical Chemistry A* 2010, 114, 11304.

Attached Publications