

Jan Gerit Brandenburg

**Development and Application of
Electronic Structure Methods for
Noncovalent Interactions in Organic
Solids**

–2015–

Theoretische Chemie

Development and Application of Electronic Structure Methods for Noncovalent Interactions in Organic Solids

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften im Fachbereich Chemie

der Mathematisch-Naturwissenschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jan Gerit Brandenburg

aus Aachen

–2015–

Dekan: Prof. Dr. Ulf-G. Meißner

Erster Gutachter: Prof. Dr. Stefan Grimme

Zweiter Gutachter: Prof. Dr. Thomas Bredow

Tag der Disputation: 03. Juli 2015

Gesamtnote: *summa cum laude*

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

To Jens

Statement of Authorship

I, Jan Gerit Brandenburg, hereby declare that I am the sole author of this thesis. The ideas and work of others, whether published or unpublished, have been fully acknowledged and referenced in my thesis.

Bonn, July 6, 2015

Publications

Parts of this thesis have already been published in international, peer-reviewed journals:

1. J. G. Brandenburg and S. Grimme, “Dispersion Corrected Hartree-Fock and Density Functional Theory for Organic Crystal Structure Prediction”, *Top Curr Chem*, **2014**, *345*, 1–23.
2. J. G. Brandenburg, T. Maas, and S. Grimme, “Benchmarking DFT and Semiempirical Methods on Structures and Lattice Energies for Ten Ice Polymorphs”, *J. Chem. Phys.*, **2015**, *tba*.
3. J. G. Brandenburg, M. Alessio, B. Civalleri, M. F. Peintinger, T. Bredow, and S. Grimme, “Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic Density Functional Theory Calculations”, *J. Phys. Chem. A*, **2013**, *117*, 9282–9292.
4. J. G. Brandenburg and S. Grimme, “Accurate Modeling of Organic Molecular Crystals by Dispersion-corrected Density Functional Tight-Binding (DFTB)”, *J. Phys. Chem. Lett.*, **2014**, *5*, 1785–1789.
5. J. G. Brandenburg, M. Hochheim, T. Bredow, and S. Grimme, “Low-Cost Quantum Chemical Methods for Non-Covalent Interactions”, *J. Phys. Chem. Lett.*, **2014**, *5*, 4275–4284.
6. J. G. Brandenburg and S. Grimme, “A Dispersion-Corrected Density Functional Theory Case Study on Ethyl Acetate Conformers, Dimer, and Molecular Crystal”, *Theor. Chem. Acc.*, **2013**, *132*, 1399.
7. J. G. Brandenburg, S. Grimme, P. G. Jones, G. Markopoulos, H. Hopf, M. Cyranski, and D. Kuck, “Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State: A Combined X-Ray and Theoretical Study”, *Chem. Eur. J.*, **2013**, *19*, 9930–9938.

8. J. G. Brandenburg, G. Bender, J. Ren, A. Hansen, S. Grimme, H. Eckert, C. G. Daniliuc, G. Kehr, and G. Erker, "Crystal Packing Induced Carbon-Carbon Double-Triple Bond Isomerization in a Zirconocene Complex", *Organometallics*, **2014**, *33*, 5358–5364.

Further publications:

1. B.-H. Xu, K. Bussmann, R. Fröhlich, C. G. Daniliuc, J. G. Brandenburg, S. Grimme, G. Kehr, and G. Erker, "An Enamine/HB(C₆F₅)₂ Adduct as a Dormant State in Frustrated Lewis Pair Chemistry", *Organometallics*, **2013**, *32*, 6745–6752.
2. F. Malberg, J. G. Brandenburg, W. Reckien, O. Hollóczki, S. Grimme, and B. Kirchner, "Substitution effect and effect of axle's flexibility at (pseudo-)rotaxanes", *Beilstein J. Org. Chem.*, **2014**, *10*, 1299–1307.
3. D. Schweinfurth, S. Demeshko, S. Hohloch, M. Steinmetz, J. G. Brandenburg, S. Dechert, F. Meyer, S. Grimme, and B. Sarkar, "Spin Crossover in Fe (II) and Co (II) Complexes with the Same Click-Derived Tripodal Ligand", *Inorg. Chem.*, **2014**, *53*, 8203–8212.

Presentations:

1. Poster: "A Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic DFT Calculations", *Symposium on Weak Molecular Interactions*, **2013**, Pécs, Hungary
2. Poster: "Dispersion Corrected DFT and HF for Organic Crystal Structure Prediction", *European Summer School in Quantum Chemistry*, **2013**, Palermo, Italy
3. Talk: "Electronic Structure Modeling of Organic Molecular Crystals", *International Conference on Computational and Mathematical Methods in Science and Engineering*, **2014**, Cádiz, Spain
4. Poster: "Semiempirical Methods for Organic Crystal Modeling", *Symposium on Theoretical Chemistry*, **2014**, Vienna, Austria
5. Talk: "Low-Cost Quantum Chemical Methods for Non-Covalent Interactions", *Theoretical Chemistry Colloquia*, **2014**, Bochum, Germany
6. Poster: "Quantum Chemical Methods for Organic Crystal Modeling", *Computational Molecular Science*, **2015**, Warwick, England

Abstract

This thesis reports on multilevel electronic structure approaches for the description of noncovalent interactions. They play an important role in various areas of chemistry and physics, ranging from bio-molecular applications to organic semi-conductors. The main focus lies on the cost-effective yet reasonably accurate description of noncovalently bound solids in the framework of organic crystal structure prediction (CSP). In principle, high-level quantum chemical wavefunction theory methods can seamlessly describe all of the local and nonlocal interactions but are computationally too demanding for large organic complexes, specifically for molecular crystals of larger molecules.

London dispersion inclusive density functional theory (DFT-D) is state-of-the-art in molecular gas phase applications. However, its absolute accuracy for lattice energies and crystal geometries was still uncertain. The good performance of DFT-D on standard and newly compiled benchmark sets is shown to be close (or within) the chemical accuracy of 1 kcal/mol. Exact exchange and three-body dispersion overall improve the performance, e.g., the mean absolute relative deviation of the hybrid functional PBE0-D3^{atm} from the reference lattice energies of the X23 and ICE10 sets is 6.6% and 6.1%, respectively. Because the references are typically experimental sublimation energies and X-ray geometries at finite temperature, a correct treatment of zero-point and thermodynamic effects is mandatory. When compared to the experimental unit-cells, which are corrected for zero-point and thermal effects, the DFT-D unit cell volumes are accurate within 1–3%. Thus, DFT-D in principle is applicable to CSP, but the computational demands to sample a huge number of polymorphs, are too high.

In the second part of this thesis, alternative low-cost methods are developed, extended to periodic boundary conditions, and evaluated on standard benchmarks. Two approaches, namely the London dispersion corrected density functional tight-binding (DFTB3-D3) and the corrected small basis set Hartree-Fock (HF-3c) are especially promising. The empiricism of HF-3c is comparable to modern density functionals (nine global parameters) while the tight binding Hamiltonian relies on element-specific parametrized pair potentials. Both schemes are shown to accurately model both solid- and gas phase inter- and intramolecular noncovalent interactions. The mean absolute deviation for interaction (lattice) energies are typically 1–3 kcal/mol (5–20%), that is, only about two times larger

Abstract

than those for DFT-D. At the same time, a speed-up of two to three orders of magnitude can be achieved. HF-3c yields very reasonable unit cell volumes (mass densities) within 3–5% error, while DFTB3-D3 yields larger errors up to 15%. However, the deviations of thermodynamic corrections to sublimation energies between the DFTB3-D3 and DFT-D level is below 0.5 kcal/mol and the tight binding model can be ideally used in a multilevel approach. One can, for instance, combine the thermal corrections of DFTB3-D3 with the electronic energy from DFT-D or use the computationally cheaper method to screen a huge number of possible conformations.

The presented methods can be routinely applied to molecular crystals as demonstrated in the last part of the thesis. The correct description of a variety of crystal packing effects is presented. Specifically, the change of the molecular conformer of ethyl acetate, the stacking of π -systems, the spin state of iron spin-crossover compounds, and the bond isomerization of certain zirconium complexes are computed in agreement with corresponding experiments.

Zusammenfassung

Die Dissertation beschäftigt sich mit Elektronenstrukturmethoden zur Beschreibung von nicht-kovalenten Wechselwirkungen in einem Multilevelansatz. Diese Wechselwirkungen spielen eine wesentliche Rolle in vielen Bereichen der Chemie und der Physik. Sie sind sowohl wichtig in biochemischen Anwendungen, als auch bei organischen Halbleitern und vielen anderen modernen Materialien. Der Hauptfokus dieser Arbeit liegt auf kostengünstigen und trotzdem hinreichend genauen theoretischen Methoden zur Beschreibung von nicht-kovalent gebundenen Festkörpern. Insbesondere die mögliche Anwendung zur organischen Kristallstrukturvorhersage (*crystal structure prediction*, CSP) wird analysiert. Im Prinzip lassen sich alle lokalen und nicht-lokalen Wechselwirkungen mit hochgenauen quantenchemischen Methoden beschreiben. Allerdings ist dies für große organische Komplexe und besonders Molekülkristalle größerer Moleküle rechnerisch zu aufwendig.

Die Dichtefunktionaltheorie mit Korrekturen für die London Dispersion (DFT-D) hat sich zur modernen Standardmethode für molekulare Gasphasensysteme entwickelt. Die Genauigkeit von berechneten Gitterenergien und Kristallgeometrien wurde jedoch bisher kaum untersucht. In dieser Dissertation wird die gute Anwendbarkeit der DFT-D Methoden anhand von gängigen sowie neu zusammengestellten Benchmarksätzen gezeigt. Die Genauigkeit für diese Testsätze ist nahe, bzw. innerhalb der chemischen Genauigkeit von 1 kcal/mol. Es wird gezeigt, dass exakter Fock-Austausch und Dreikörperdispersion die Ergebnisse verbessern. Die mittlere absolute prozentuale Abweichung der mittels des Hybridfunktionals PBE0-D3^{atm} berechneten Gitterenergien der X23 und ICE10 Systeme von den entsprechenden Referenzen ist 6.6% und 6.1%. Die Referenzen sind typischerweise experimentelle Sublimationsenergien und Röntgenstrukturen, die bei endlichen Temperaturen gemessen werden. Deshalb ist eine Berücksichtigung von Nullpunkts- und thermischen Effekten entscheidend. Die via DFT-D berechneten Kristallvolumen stimmen mit den zurückkorrigierten gemessenen Einheitszellen innerhalb von 1–3% überein. Damit ist DFT-D im Prinzip zur Kristallstrukturvorhersage geeignet, aber die rechnerischen Anforderungen um eine sehr große Anzahl an Polymorphen zu beschreiben sind zu hoch.

Im zweiten Teil der Dissertation werden kostengünstige Alternativen entwickelt, auf pe-

periodische Randbedingungen erweitert und an Benchmarksystemen getestet. Insbesondere zwei Methoden stellen sich als besonders vielversprechend heraus. Dies ist zum einen das dispersionskorrigierte Dichtefunktional-Tight-Binding Modell (DFTB3-D3) und zum anderen ein in einer Minimalbasis ausgewertetes Hartree-Fock mit Korrekturtermen (HF-3c). Der empirische Charakter von HF-3c ist vergleichbar mit dem üblicher Dichtefunktionale (neun globale Parameter), während der DFTB3-D3 Hamiltonian auf parametrisierten element-spezifischen Paarpotentialen beruht. Es wird gezeigt, dass beide Ansätze inter- sowie intramolekulare Wechselwirkungen in der Gasphase und der festen Phase gut modellieren können. Der mittlere absolute Fehler von Bindungsenergien ist typischerweise 1–3 kcal/mol (5–20%), d.h. lediglich um einen Faktor zwei schlechter als DFT-D. Gleichzeitig werden die entsprechenden Rechnungen um zwei bis drei Größenordnungen beschleunigt. HF-3c ergibt sinnvolle Massendichten (Kristallvolumen) mit Fehlern von 3–5%, während DFTB3-D3 Geometrien Fehler von bis zu 15% haben. Trotzdem sind die thermodynamischen Korrekturen von DFTB3-D3 recht genau und weichen nur 0.5 kcal/mol von den entsprechenden DFT-D Rechnungen ab. Damit ist das Tight-Binding Modell ideal in einem Multilevelansatz einsetzbar, etwa durch Kombination von thermischen Korrekturen mit der elektronischen Energie von DFT-D oder als günstige Screeningmethode mit der man tausende mögliche Konfigurationen testen kann.

Die vorgestellten Methoden wurden in die entsprechenden Programmpakete implementiert und lassen sich routinemäßig zur Beschreibung von Molekulkristallen einsetzen. Die richtige Beschreibung von verschiedenen Kristallpackungseffekten wird präsentiert. Wie im letzten Teil dieser Dissertation gezeigt wird, gehört dazu die Änderung des Molekülkonformers von Ethylacetat, die relative Orientierung von gestapelten π -Systemen, der Spinzustand eines Eisen-Spincrossover Systems und die Bindungsisomerie in einem Zirkoniumkomplex.

Contents

Statement of Authorship	vii
Publications	ix
I. Introduction	1
0.1. Summary of Methodologies	8
II. Benchmarking Density Functional Theory	11
1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction	15
1.1. Introduction	16
1.2. Dispersion Corrected Density Functional Theory	19
1.2.1. London Dispersion Correction	19
1.2.2. Evaluation of Dispersion Corrected DFT	22
1.3. Dispersion Corrected Hartree-Fock with Basis Set Error Corrections	28
1.3.1. Basis Set Error Corrections	28
1.3.2. Evaluation of Dispersion and Basis Set Corrected DFT and HF	30
1.4. Conclusions	32
2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs	35
2.1. Introduction	36
2.2. Benchmark Setup	38
2.2.1. Systems under Consideration	38
2.2.2. Computational Details	39
2.2.3. Correction for Zero-Point Energies	41

2.3. Results and Discussion	43
2.3.1. Potential Energy Surface of Ice VIII	43
2.3.2. Equilibrium Structures for ICE10	45
2.3.3. Electronic Lattice Energy	47
2.3.4. Comparison to Gas Phase Water Clusters	50
2.4. Conclusions	51
 III. Development of Low-Cost Methods	 53
 3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT	 57
3.1. Introduction	59
3.2. Theory	60
3.3. Computational Details	63
3.4. Comparison of Counterpoise Corrections	65
3.5. X23 Benchmark Set	67
3.6. Graphite: Interlayer Distance and Energy	72
3.7. The gCP Correction for the pob-TZVP Basis	74
3.8. Conclusions	76
 4. Modeling Organic Molecular Crystals by Dispersion Corrected DFTB	 79
4.1. Results and Discussion	80
4.2. Computational Methods	86
 5. Low-Cost Quantum Chemical Methods for Noncovalent Interactions	 87
5.1. Introduction	88
5.2. Covered Methods	91
5.2.1. HF-3c	91
5.2.2. Neglect of Differential Overlap Approximations	92
5.2.3. Density Functional Tight-Binding	93
5.2.4. London Dispersion Interaction	94
5.2.5. Hydrogen Bonding Correction	95
5.3. Technical Setup	96
5.4. Results and Discussion	96
5.4.1. Small Molecular Systems	97
5.4.2. Large Molecular Complexes and Organic Solids	100
5.4.3. Summary and Future Directions	103

IV. Application to Organic and Organometallic Crystals	107
6. DFT-D Study on Ethyl Acetate: Conformers, Dimer, and Molecular Crystal	111
6.1. Introduction	112
6.2. Computational Details	114
6.3. Results and Discussion	115
6.4. Summary	119
7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State	121
7.1. Introduction	123
7.2. Results and Discussion	125
7.2.1. Experimental Studies on the Solid-State Structures of Tribenzotriquinacenes 2 and 3	125
7.2.2. Theoretical Studies on the Solid-State Structures of Tribenzotriquinacenes 2 and 3	128
7.2.3. Oligomer-to-Polymer Convergence of Stacked Tribenzotriquinacene 2	135
7.3. Conclusions	137
7.4. Experimental Section	139
8. Crystal Packing Induced Bond Isomerization	141
8.1. Introduction	142
8.2. Results and Discussion	143
8.2.1. Preparation and Structural Characterization of the Compounds 2	143
8.2.2. Theoretical Studies	145
8.3. Discussion and Conclusion	152
8.4. Experimental Section	153
8.4.1. Preparation of Complex 2b	153
8.4.2. Solid-State NMR Studies	153
V. Summary and Conclusions	155
Bibliography	161
VI. Appendix	181
A. Supporting Information to Chapter 2	182

Contents

B. Supporting Information to Chapter 3	189
C. Supporting Information to Chapter 4	191
D. Supporting Information to Chapter 5	197
 Acknowledgments	 211
 Curriculum Vitae	 213

Part I.

Introduction

Computational methodologies are supposed to help and guide scientific understanding of various systems in chemistry and physics ideally in an unbiased way. With reliable predictive power, theoretical methods can partly replace tedious and expensive experimental work. Used complementary to quantitative experiments, it is possible to explain and describe phenomena in a detailed mechanistic way.

The description of organic solids is a very wide field with many possible applications. Organic solids play an important role in material science, where the crystal band structure determines the electronic properties of organic semi-conductors.¹⁻³ Very subtle geometric changes lead to many energetically close polymorphs, which can exhibit different physical properties. As an every-day example, the taste of chocolate depends on its polymorphic form. Polymorph V is the well tasting dark chocolate, while it is transformed into polymorph VI upon heating (Figure 0.1). Chocolate VI is actually more stable compared to V, but a significant barrier separates the two phases and the “bloomed” form VI will only appear with either additional heating or after a period of some months.⁴ The polymorphism

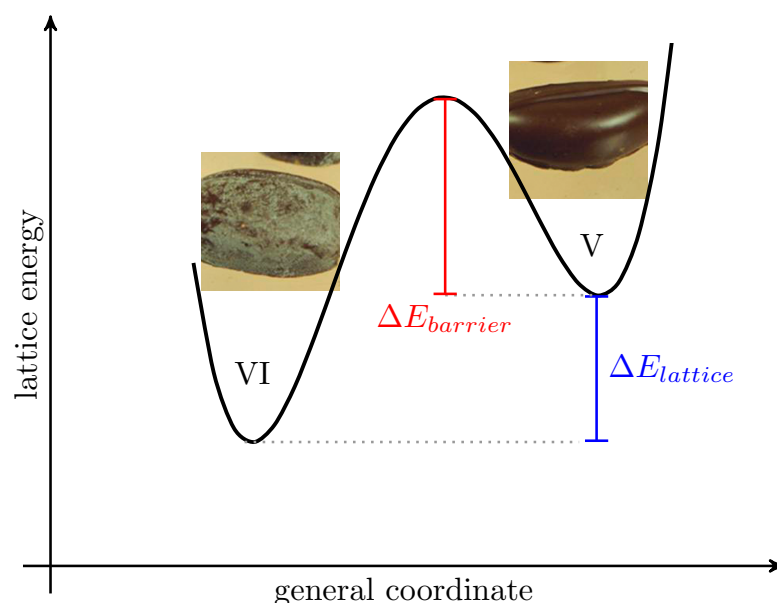


Figure 0.1.: Polymorph VI (left) and V (right) of dark chocolate, a sketch of a hypothetical potential energy surface is given.

of a material can alter other properties, for instance pigments can change their color⁵ and explosives can have different detonation properties.⁶ Further, pharmaceutical compounds are typically applied in solid form, where the polymorphic form is of utmost importance. Other forms may have different undesirable properties. There are many examples of well known drugs with multiple polymorphs like aspirin or paracetamol.⁷ A very dramatic

0. Introduction

effect was seen for the antivirus drug ritonavir. It was used for treating acquired immune deficiency syndromes (AIDS) and was industrially produced in its active form (I) in 1996. Two years later, the effectiveness of the drug decreased significantly. It could be shown that form I transformed into the more stable form II which has a much lower solubility leading to a less potent medicament. Because of its appearance in the manufacturing process, it triggered the crystallization of all follow up processes towards form II.⁸ The

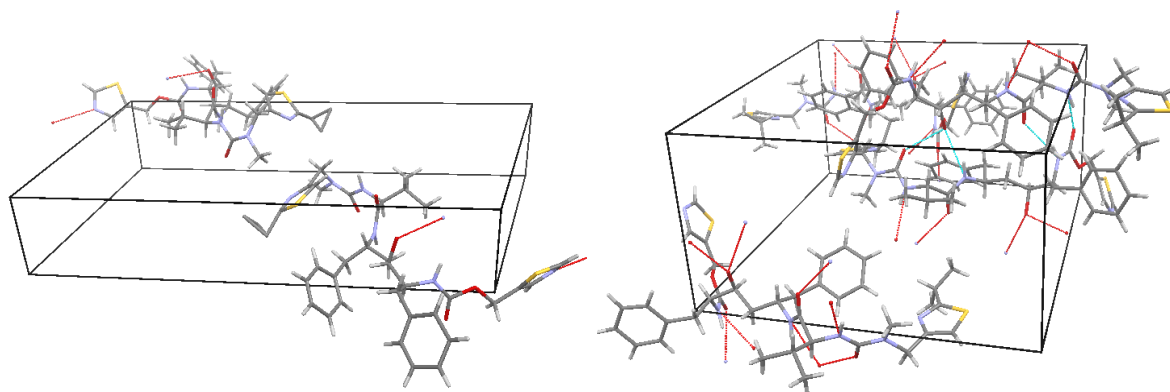


Figure 0.2.: Polymorphic forms I (left) and II (right) of the antivirus drug ritonavir as determined by X-ray experiments.⁸

two forms of ritonavir are shown in Figure 0.2. One can see the more complex hydrogen bonding network in form II, which probably leads to the stronger binding, a.k.a. lower lattice energy. Conclusively, the accurate theoretical description of these systems is of huge importance. Combination with organo-metallic compounds might even give rise to new analytical techniques and theoretical methods can help in their design process. Recent studies showed that these metal-organic frameworks can be used to crystallize small molecules in an environmentally defined lattice structure.⁹

A particularly challenging field is the organic crystal structure prediction (CSP) from "scratch". Solely knowing the chemical formula the corresponding most stable crystal structure shall be determined.^{10,11} As soon as the CSP can be reliably performed, one could theoretically guide the engineering of crystals for specific functions. Two decades back, Gavezzotti answered the question "Are crystal structures predictable?" with a clear "No".¹² This is because CSP has several layers of complexity.

1. The conformations of the single molecule have to be determined. Depending on the intrinsic flexibility of the molecule, several energetically close structures are possible. For example, a simulated annealing guided search for ritonavir revealed 64 conformers within a 10 kcal/mol energy window.

2. All 230 possible crystal space groups with varying number of molecules per unit cell, relative molecular orientation within the unit cell, and molecular conformation (according to step 1) have to be generated.
3. The polymorphs have to be ranked according to their free energy. This is an exceptionally challenging task, because
 - (a) millions of possible structures have to be considered,
 - (b) the energy differences between different polymorphs can be tiny (tenth of kcal/mol).
4. One has to describe the actual crystallization process, because certain structures may not be kinetically feasible due to energy barriers.
5. To design the desired properties of the target material one has to modify the crystallization process and predict the change in certain observables.

The first step is typically done in the gas phase. However, the relative energy between different molecular conformations can change due to the crystal environment. Some of these crystal packing effects (CPE) are discussed in part IV of this thesis. A probably more decent choice is to perform the conformational search in an implicit solvent environment. The combinatorial problem of generating starting structures (step 2) has been more or less solved by different Monte-Carlo sampling techniques and evolutionary or particle swarm algorithms.^{13–15} Concerning step 3, tremendous progress has been made in the last decade by using classical pair-potentials (force fields) to rank the millions of structures.^{16–18} A recently proposed quantum mechanically derived force field (QMDFFF) could be applied for molecular crystals as well.¹⁹ At the same time, it became apparent that quantum mechanical descriptions are needed for a more reliable energy ranking of the pre-sorted polymorphs.

For this aim, an accurate yet efficient description of all inter- and intramolecular non-covalent interactions is needed.^{20–22} Electrostatic (ES), induction (IND), Pauli exchange repulsion (EXR), and London dispersion (DISP) effects have to be treated accurately (compare with Figure 0.3 for a sketch of the different interaction types). The exact solution of the many-particle Schrödinger equation would seamlessly cover all energy contributions. However, this is prohibitive for realistic systems but effective one-particle models, such as the Hartree-Fock (HF) approximation, or the density functional approximation (DFA), can be employed. Density functional theory directly maps the three-dimensional electron density $\rho(\mathbf{r})$ to the electronic energy, which makes this ansatz especially cost-effective. The electron correlation is typically expressed in a local framework. Semi-local

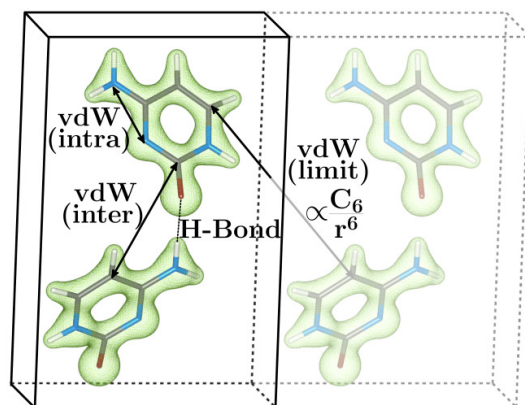


Figure 0.3.: Prototypical molecular crystal with charge density isosurface and highlighted noncovalent interactions.

DFAs can accurately describe ES, IND, and EXR effects. Since London dispersion is a nonlocal, long-range electron correlation effect, standard semilocal DFAs cannot describe it properly.^{23–25} Significant progress has been made in recent years to solve the dispersion problem. Various approaches providing very good accuracy are available and are now becoming the standard in the field.^{26–30} For overviews on these dispersion-corrected, so-called DFT-D methods, see e.g. Refs.^{31–33}

The Cambridge structural database organizes CSP blind challenges every few years to test and trigger new methodological progress.^{34,35} In the 4th crystal structure prediction blind test, a combination of tailor made force fields and periodic DFT-D methods could successfully predict the crystal structure of all four target molecules.^{36,37} Though DFT-D was rather successful and extensive molecular benchmarks exist,³⁸ detailed benchmarking regarding the accuracy of absolute binding energies in organic solids was still missing. For this aim, E. Johnson compiled a set of organic solids with experimental references, which was later refined by A. Tkatchenko (dubbed X23).^{39,40} In part II, the DFT-D methodology is benchmarked concerning absolute lattice energies and equilibrium geometries of organic solids. For the first time, the most recent D3 dispersion correction is applied to the X23 benchmark set and shown to be competitive to the more expensive XDM and TS schemes (Chapter 1).⁴¹ Because water in its various phases is of utmost importance for many applications, an additional benchmark of ten ice polymorphs has been compiled to judge the accuracy of various DFAs.⁴² Effects of the exchange correlation functional, of the exact Fock exchange, and of the three-body dispersion are analyzed and compared to other dispersion inclusive DFA approaches (Chapter 2). This set, dubbed ICE10, is

proposed to be used complementary to the X23 set.

The 5th CSP blind test showed that two problems remain.⁴³ First, the final lattice energies between polymorphs differ by only a few tenth of a kcal/mol. As shown in a number of studies, the accuracy of DFT-D for absolute lattice energies is within 1–1.5 kcal/mol (see also Chapter 1 and 2).^{39–41} Higher (relative) accuracy is needed to separate the different polymorphs. This can potentially be obtained by correlated wave-function based methods. In the past years, substantial improvements have been made to allow a description of larger systems.⁴⁴ Combined with fragmentation and embedding schemes, these methods can also be applied to molecular crystals with promising results.^{45–51} Another approach is to use quantum Monte Carlo (QMC) techniques. By using the shear force of modern computational power, QMC can potentially compute accurate binding energies. This is possible because of the cubic scaling of QMC with respect to the particle number and the possible parallelization on million computer nodes.^{52–55} The second problem revealed in the 5th CSP blind test is the gap in both accuracy and computational efficiency between the classical force fields and the DFT-D (or related) methods.

In part III, methodologies to fill this gap are developed and examined. The reduction of the single-particle basis set (in the HF or DFT framework) can lead to substantial speed-ups (factor of 50–100 compared to a converged basis set calculation). However, basis set superposition errors (BSSE) are large and have to be corrected. A new geometrical counterpoise correction (gCP) has been suggested recently,⁵⁶ which is extended to periodic boundaries⁵⁷ and applied to several benchmark systems (Chapter 3). For additional speed up, many-center terms have to be neglected which is the strategy of all molecular orbital (MO) based semiempirical methods.^{58–61} A modern third order density functional tight-binding Hamiltonian is coupled with the D3 dispersion correction and successfully applied to molecular crystals (Chapter 4).⁶² A valence variant of the HF-3c⁶³ method is developed and extensively evaluated together with competing low-cost semiempirical schemes on several noncovalent benchmark sets (Chapter 5).⁶⁴

Several applications of the developed methodologies have shown their robust and accurate performance. As indicated above, the packing of molecules in the crystal can change their conformation. Some selected examples of these CPEs are presented in part IV of this thesis. The most stable gas phase conformer of ethyl acetate is the *gauche* form. In the solid state, the *trans* conformer can form more hydrogen bonds leading to a significant stabilization (Chapter 6). The relative stabilization can be correctly described via modern DFT-D3 methods, while older variants yield larger errors.⁶⁵ The packing can also modify the relative orientation of neighboring molecules. A subtle rotation of stacked tribenzotriquinacene molecules could be assigned to the three-dimensional packing (Chapter 7).⁶⁶ An even stronger packing effect could be identified in a zirconocene

complex. Intermolecular London dispersion forces transform a double bond in the gas phase ($-\text{MeC}=\text{C}=\text{C}(\text{Zr})-\text{CMe}_3$) to a triple bond in the solid state ($-\text{MeC}\equiv\text{C}-\text{C}(\text{Zr})-\text{CMe}_3$, Chapter 8).⁶⁷ The correct description of these multitudes of subtle CPEs demonstrates the promising perspective of the applied methods.

0.1. Summary of Methodologies

The systems that have to be described are classified by the non-relativistic electronic Hamiltonian in the Born-Oppenheimer approximation

$$\mathcal{H} = \mathcal{V}_{ext} + \mathcal{T}_e + \mathcal{V}_{en} + \mathcal{V}_{ee} . \quad (0.1)$$

consisting of a fixed external potential \mathcal{V}_{ext} generated by the nuclei, the kinetic energy of the electrons \mathcal{T}_e , the Coulomb interaction between the electrons and the nuclei \mathcal{V}_{en} , and the Coulomb interaction between the electrons. For the target systems and target accuracy, the quantum effects of the nuclei and relativistic effects of the valence electrons are not needed. For periodic systems, the translation symmetry can be used to factorize the wave function in a product of translation eigenfunctions and a periodic function reflecting the symmetry of the crystal (known as Bloch theorem). The translation eigenfunctions have to represent the first Brillouin zone of the reciprocal lattice, which can be efficiently done by numerical k -space grids.⁶⁸ A formally exact solution of the many-particle Schrödinger equation corresponding to the Hamiltonian \mathcal{H} can be given, for instance, via a full configuration interaction expansion of a reference state. However, its explicit solution is only possible for systems with very few atoms. To convert the many-particle Schrödinger equation into an algebraic problem, the wavefunction is expanded into a basis set of one-particle functions. While for periodic systems plane-waves are the natural choice, the correct description of localized densities requires a huge number of basis functions (specifically tested in Chapter 6). Atom-centered Gaussian functions may be competitive to describe molecular crystals. This will be analyzed and discussed in more detail in Chapter 3.

A wide range of different methodologies have been used in this thesis. A general classification of various quantum chemical and classical methods are sketched in Figure 0.4. They are sorted according to their computational cost and their general applicability (both decreasing from bottom to top). One typically distinguishes between *ab-initio* wavefunction theory, *first principles* density functional theory (DFT), *semiempirical* molecular orbital (MO) based methods, and classical methods that do not treat electrons explicitly (atomistic and coarse grained force fields and continuum mechanics). As mentioned above,

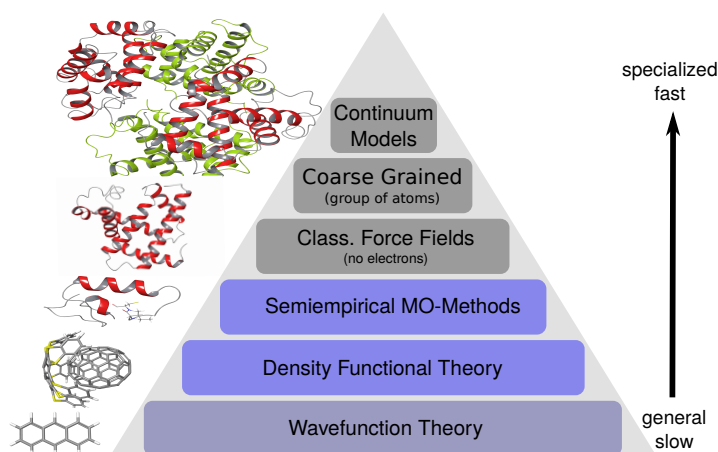


Figure 0.4.: Classification of quantum chemical methods according to their computational cost and generality. Reprinted (adapted) from chapter 5

wavefunction based methods can be used to calculate highly accurate reference energies for small to medium sized systems. The noncovalent benchmark sets used in Chapter 4 and Chapter 5 are mostly performed at the CCSD(T)/CBS(est.) level of theory (coupled cluster level with singles, doubles and perturbative triples at the estimate basis set limit). Effective one-particle DFT and HF methods are introduced in Chapter 1. The main focus is on the inclusion of London dispersion interactions. The DFT-D methodology is shortly reviewed in Section 1.2. Computational method with lower computational costs like minimal basis set HF calculations or MO based semiempirical methods are reviewed in Chapter 5. They present a low-cost alternative when DFT-D methods are not feasible. Especially for the CSP tasks introduced previously, force field potentials cannot be avoided due to the huge sampling space. However, these classical methods do not treat the electronic structure explicitly are not further investigated in this thesis.

Part II.

Benchmarking Density Functional Theory

The first part of the thesis explores the state-of-the-art in density functional based description of noncovalent interactions in organic solids. Because DFT methods rely on certain approximations, their accuracy has to be benchmarked for specific applications. The general methodology is presented and reviewed to some extent in Chapter 1. The main focus lies on the description of London dispersion interactions. The recently introduced X23 benchmark set is evaluated with D3 dispersion corrected DFT and compared to the competing exchange dipole model (XDM) and Tkatchenko-Scheffler (TS) methods. Additionally, a case study on the crystal structure of tribenzotriquinacene molecules is recapitulated. More details on this study can be found in Chapter 7. Some methods with smaller computational costs are tested in comparison to the well established DFT-D methods. They were partly developed during this thesis and presented in more detail in Chapter 3. A perspective for dispersion corrected HF and DFT based organic crystal structure prediction is given.

In Chapter 2, a new benchmark consisting of ten ice polymorphs is compiled. Because water in its various phases is of utmost importance in biological processes, specific benchmarking seems necessary. Further, strong hydrogen bonded systems are under represented in the X23 set, which is basically the only commonly used benchmark set in the field of CSP. Experimental references for the lattice energies of seven ice polymorphs are available. In this study, the experimental unit cells have been corrected for zero-point and thermal effects. In this way, one can directly compare the optimized structures from free optimizations with the back-corrected references. Dispersion inclusive DFT as well as alternative low-cost methods are evaluated and yield promising accuracy.

1. Dispersion Corrected Hartree-Fock and Density Functional Theory for Organic Crystal Structure Prediction

Jan Gerit Brandenburg* and Stefan Grimme*

Keywords: Crystal Structure Prediction, Density Functional Theory, Hartree-Fock, Dispersion Correction, Counterpoise Correction

Received 15th of April 2013, Published online 13th of November 2013

Reprinted (adapted) with permission from

J. G. Brandenburg and S. Grimme, *Top Curr Chem* **2014**, *345*, 1–23.

— Copyright ©2014, Springer-Verlag Berlin Heidelberg. DOI 10.1007/128_2013_488

Own manuscript contribution

- Performance of DFT-D and HF-3c calculations
- Interpretation of data
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstraße 4, 53115 Bonn, Germany

Abstract We present and evaluate dispersion corrected Hartree-Fock (HF) and Density Functional Theory (DFT) based quantum chemical methods for organic crystal structure prediction. The necessity of correcting for missing long-range electron correlation, also known as van der Waals (vdW) interaction, is pointed out and some methodological issues such as inclusion of three-body dispersion terms are discussed. One of the most efficient and widely used methods is the semi-classical dispersion correction D3. Its applicability for the calculation of sublimation energies is investigated for the benchmark set X23 consisting of 23 small organic crystals. For PBE-D3 the mean absolute deviation (MAD) is below the estimated experimental uncertainty of 1.3 kcal/mol. For two larger π -systems, the equilibrium crystal geometry is investigated and very good agreement with experimental data is found. Since these calculations are carried out with huge plane-wave basis sets they are rather time consuming and routinely applicable only to systems with less than about 200 atoms in the unit cell. Aiming at crystal structure prediction, which involves screening of many structures, a pre-sorting with faster methods is mandatory. Small, atom-centered basis sets can speed up the computation significantly but they strongly suffer from basis set errors. We present the recently developed geometrical counterpoise correction gCP. It is a fast semiempirical method, which corrects for most of the inter- and intramolecular basis set superposition error. For HF calculations with nearly minimal basis sets, we additionally correct for short-range basis incompleteness. We combine all three terms in the HF-3c denoted scheme which performs excellently for the X23 sublimation energies with an MAD of only 1.5 kcal/mol, which is close to the huge basis set DFT-D3 result.

1.1. Introduction

Aiming at organic crystal structure prediction, two competing requirements for the utilized theoretical method exist. On the one hand, the calculation of crystal energies has to be accurate enough to distinguish between different polymorphs. This involves an accurate account of inter- as well as intramolecular interactions in various geometrical situations. On the other hand, each single computation (energy including the corresponding derivatives for geometry optimization or frequency calculation) has to be fast enough to sample all space groups under consideration (and possibly different molecular conformations) in a reasonable time^{37,69–72}. Typically, one presorts the systems with a fast method and investigates the energetically lowest ones with a more accurate (but at the same time more costly) method. For the inclusion of zero point vibrational energy (ZPVE) contributions a medium quality level is often sufficient. A corresponding algo-

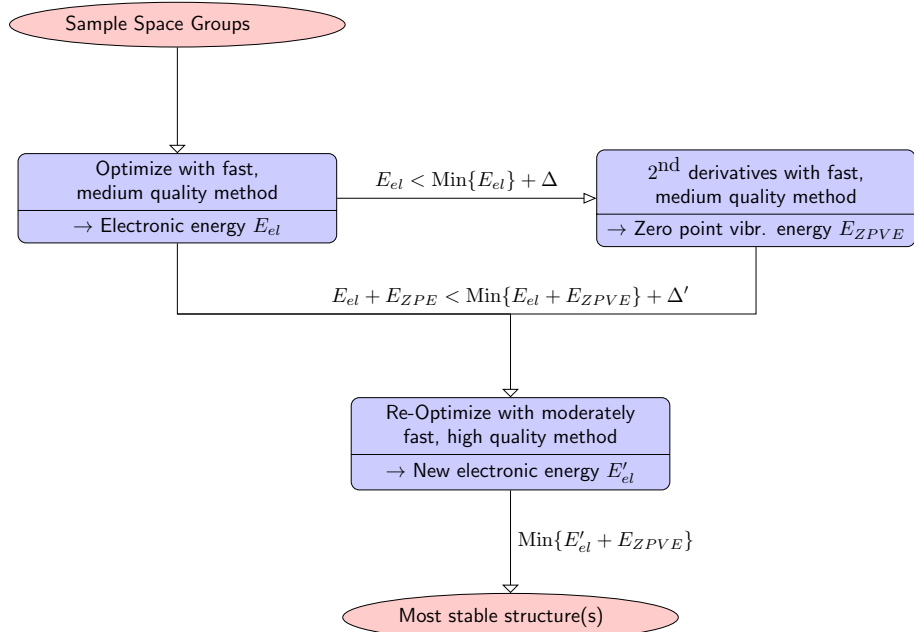


Figure 1.1.: A typical crystal structure prediction algorithm³⁷. First, the optimum electronic crystal energy E_{el} is calculated with a fast, medium quality method. Secondly, the more costly second derivatives for the electronically lowest structures in a certain energy interval (Δ) are calculated to get the zero point vibrational energy E_{ZPVE} . Finally, the electronic energy E'_{el} is re-calculated for the energetically lowest structures in a (different) energy interval (Δ') with a more accurate method. The data from step two can be finally used also to estimate thermal and entropic corrections.

algorithm is sketched in Fig. 1.1. The generation of the start structure (denoted as sample space groups) is an important issue, but will not be regarded by us. Here, we focus on the different electronic structure calculations, denoted by the blue shaded steps in Fig. 1.1. We present dispersion corrected Density Functional Theory (DFT-D3) as a possible high-quality method with medium computational cost and dispersion corrected Hartree-Fock (HF) with semiempirical basis error corrections (HF-3c) as a faster method with medium quality.

Density Functional Theory (DFT) is the 'work horse' for many applications in chemistry and physics and still an active research field of general interest^{73–76}. In many covalently bound (periodic and non-periodic) systems, DFT provides a very good compromise between accuracy and computational cost. However, common generalized gradient approximated (GGA) functionals are not capable of describing long-range electron correlation, a.k.a. the London dispersion interaction^{23–25,77}. This dispersion term can be empirically defined as the attractive part of the van der Waals-type interaction between atoms and molecules that are not directly bonded to each other. For the physically correct description

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

of molecular crystals, dispersion interactions are crucial^{20,78}. In the last decade, several well-established methods for including dispersion interactions into DFT were developed. For an overview and reviews of the different approaches, see e.g. Refs.^{30,32,33,40,79–84} and references therein. Virtual orbital dependent (e.g. random phase approximation, RPA⁸⁵) and fragment-based (e.g. symmetry adapted perturbation theory, SAPT⁸⁶) methods are not further discussed here because they are currently not routinely applicable to larger molecular crystals. For the alternative combination of accurate molecular quantum chemistry calculations for crystal fragments with force-fields and subsequent periodic extension see e.g. Ref.^{7,45}.

Here, we focus on the basically atom-pairwise dispersion correction D3^{26,87} coupled with periodic electronic structure theory. The D3 scheme incorporates non-empirical, chemical environment-dependent dispersion coefficients, and for dense systems a non-additive Axilrod-Teller-Muto three-body dispersion term. We present the details of this method in section 1.2.1. Compared to the self-consistent solution of the Kohn-Sham (KS) or HF equations, the calculation of the D3 dispersion energy requires practically no additional computation time. Although it does not include information about the electron density, it provides good accuracy with typical deviations for the asymptotic dispersion energy of only 5 %³². The accuracy for non-covalent interaction energies with current standard functionals and D3 is about 5-10 %, which is also true for small relative energies⁸⁸. Therefore, it is an ideal tool to fulfill fundamental requirements of crystal structure prediction. We evaluate the DFT-D3 scheme with huge plane-wave basis sets in section 1.2.2 and compare it to competing pairwise-additive methods, which partially employ electron density information.

Because the calculation of the DFT or HF energy is the computational bottleneck, a speed-up of these calculations without losing too much accuracy is highly desirable. The computational costs mainly depends on the number of utilized single particle basis functions N with a typical scaling behavior from N^2 to N^4 . The choice of the type of basis functions is also an important issue. Bulk metals have a strongly delocalized valence electron density and plane-wave based basis sets are probably the best choice⁸⁹. In molecular crystals, however, the charge density is more localized and a typical molecular crystal involves a lot of 'vacuum'. For plane wave based methods, this can result in a huge (inefficient) basis sets. In a recently studied typical organic system (tribenzotriquinacene, $C_{22}H_{16}$), up to $1.5 \cdot 10^5$ projector augmented plane-wave (PAW) basis functions must be considered for reasonable basis set convergence⁶⁶. For this kind of system, atom-centered Gaussian basis functions as usually employed in molecular quantum chemistry could be more efficient. However, small atom-centered basis sets strongly suffer from basis set errors (BSE), especially the basis set superposition error (BSSE) which leads

to overbinding and too high computed weight densities (too small crystal volumes) in unconstrained optimizations. Because different polymorphs often show various packings with different densities, correcting for BSSE is mandatory in our context. Also in order to get reasonable absolute sublimation energies and good crystal geometries, these basis set errors must be corrected. A further problem compared to plane-wave basis sets is the non-orthogonality of atom-centered basis functions which can lead to near-linear dependencies and bad self-consistent field (SCF) convergence. Recently, we have mapped the standard Boys & Bernardi correction⁹⁰, which corrects for the BSSE, onto a atom-pairwise repulsive potential. It was fitted for a number of typical Gaussian basis sets and depends otherwise only on the system geometry and is therefore denoted gCP⁵⁶. Analytic gradients are problematic in nearly all other counterpoise schemes, but are easily obtained for gCP. Especially for the calculation of second derivatives, analytic first derivatives are crucial. Periodic boundary conditions are included and the implementation has been tested in Ref⁵⁷. We present the gCP scheme here together with an additional short-range basis (SRB) incompleteness correction in section 1.3.1. In section 1.3.2, the combination of small (almost minimal) basis set DFT and HF, dispersion correction D3, geometrical counterpoise correction gCP, and short-range incompleteness correction SRB is evaluated for typical molecular crystals. The plane-wave, large basis PBE-D3 results are briefly discussed and used for comparison.

1.2. Dispersion Corrected Density Functional Theory

1.2.1. London Dispersion Correction

At short inter-atomic distances, standard density functionals (DF) describe the effective electron interaction rather well because of their deep relation to the corresponding electron density changes. Long-range electron correlation cannot be accurately described by the local (or semilocal) DFs in inhomogeneous materials. To describe this van der Waals (vdW)-type interaction, one can include nonlocal kernels in the vdW-DFs as pioneered by Langreth and Lundquist^{91,92} and later improved by Vydrov and van Voorhis (VV10³⁰). For the total exchange-correlation energy E_{xc} of a system, the following approximation is employed in all vdW-DF schemes

$$E_{xc} = E_X^{GGA} + E_C^{GGA} + E_c^{NL}, \quad (1.1)$$

where standard exchange (X) and correlation (C) components (in the semilocal generalized gradient approximation GGA) are used for the short-range parts and E_c^{NL} represents the

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

nonlocal correlation term describing the dispersion energy. In the vdW-DF framework it takes the form of a double-space integral

$$E_c^{NL} = \frac{1}{2} \iint \rho(\vec{r}) \Phi^{NL}(\vec{r}, \vec{r}') \rho(\vec{r}') d^3r d^3r'. \quad (1.2)$$

The electron density ρ at positions \vec{r} and \vec{r}' is correlated via the integration kernel $\Phi^{NL}(\mathbf{r}, \mathbf{r}')$. It is physically approximated by local approximations to the frequency dependent dipole polarizability $\alpha(\mathbf{r}, \omega)$. The VV10 kernel has been successfully used in various molecular applications^{93–96} by us but is not discussed further in this work.

The famous Casimir-Polder relationship⁹⁷ connects the polarizability with the long-range dispersion energy, which scales as C_6/R^6 where R is the distance between two atoms or molecules. The corresponding dispersion coefficient C_6^{AB} for interacting fragments A and B is given by

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty \alpha^A(i\omega) \alpha^B(i\omega) d\omega, \quad (1.3)$$

where $\alpha^A(i\omega)$ is the averaged dipole polarizability at imaginary frequency ω . In vdW-DF (but not in DFT-D3) dispersion can be calculated self-consistently and changes the density in turn. Because this change is normally insignificant^{30,91,94}, E_c^{NL} is typically added non-self-consistently to the SCF-GGA energy. The main advantage of vdW-DF methods is that dispersion effects are naturally included via the system electron density. Therefore, they implicitly account for changes in the dispersion coefficients due to different 'atoms-in-molecules' oxidation states in a physically sound manner. The disadvantage is the raised computational cost compared to pure (semi-)local DFs.

By treating the short-range part with DFs and the dispersion interaction with a semi-classical atom-pairwise correction, one can combine the advantages of both worlds. Semi-classical models for the dispersion interaction like D3 show very good accuracy compared to e.g. the VV10 functional^{96,98} for very little computational overhead in particular when analytical gradients are required.

The total energy E_{tot} of a system can be decomposed into the standard, dispersion-uncorrected DFT/HF electronic energy $E_{\text{DFT/HF}}$ and the dispersion energy E_{disp}

$$E_{\text{tot}} = E_{\text{DFT/HF}} + E_{\text{disp}}. \quad (1.4)$$

We use our latest first-principles type dispersion correction DFT-D3, where the dispersion coefficients are non-empirically obtained from a time-dependent, linear response DFT calculation of $\alpha^A(i\omega)$. The dispersion energy can be split into two- and three-body con-

tributions $E_{\text{disp}} = E^{(2)} + E^{(3)}$:

$$E^{(2)} = -\frac{1}{2} \sum_{n=6,8} \sum_{A \neq B}^{\text{atom pairs}} \sum_{\vec{T}} s_n \frac{C_n^{AB}}{\left\| \vec{r}_B - \vec{r}_A + \vec{T} \right\|^n + f(R_0^{AB})^n} \quad (1.5)$$

$$E^{(3)} = \frac{1}{6} \sum_{A \neq B \neq C}^{\text{atom triples}} \sum_{\vec{T}} \frac{C_9^{ABC} (3\cos\theta_a \cos\theta_b \cos\theta_c + 1)}{r_{ABC}^9 \cdot (1 + 6(r_{ABC}/R_0)^{-\alpha})}. \quad (1.6)$$

Here, C_n^{AB} denotes the averaged (isotropic) n^{th} -order dispersion coefficient for atom pair AB , and $\vec{r}_{A/B}$ are their Cartesian positions. The real-space summation over all unit cells is done by considering all translation invariant vectors \mathbf{T} inside a cut-off sphere. The scaling parameter s_6 equals unity for the here employed DFs and ensures the correct limit for large interatomic distances, and s_8 is a functional-dependent scaling factor. The rational Becke and Johnson damping function $f(R_0^{ab})$ is⁹⁹

$$f(R_0^{ab}) = a_1 R_0^{ab} + a_2, \quad R_0^{ab} = \sqrt{\frac{C_8^{ab}}{C_6^{ab}}}. \quad (1.7)$$

The dispersion coefficients C_6^{AB} are computed for molecular systems with the Casimir-Polder relation (Equation 1.3). We use the concept of fractional coordination numbers (CN) to distinguish the different hybridization states of atoms in molecules in a differentiable way. The CN is computed from the coordinates and does not use information from the electronic wavefunction or density but recovers basic information about the bonding situation of an atom in a molecule, which has a dominant influence on the C_6^{AB} coefficients²⁶. The higher order C_8 coefficients are obtained from the well-known relation¹⁰⁰

$$C_8 = \frac{3}{2} C_6 \frac{\langle r^4 \rangle}{\langle r^2 \rangle}. \quad (1.8)$$

With the recursion relation $C_{i+4} = C_{i-2} \left(\frac{C_{i+2}}{C_i} \right)$ and $C_{10} = \frac{49}{40} \frac{C_8^2}{C_6}$, one can in principle generate also higher orders, but terms above C_{10} do not improve the performance of the D3 method. The three parameters s_8 , a_1 , and a_2 are fitted for each DF on a benchmark set of small, non-covalently bound complexes. This fitting is necessary to prevent double counting of dispersion interactions at short range and to smoothly interpolate between short- and long-range regimes. These parameters are successfully applied to large molecular complexes and to periodic systems^{98,101}. In the non-additive Axilrod-Teller-Muto three-body contribution (Equation 1.6)^{26,102}, r_{ABC} is an average distance in the atom-triples and $\theta_{a/b/c}$ are the corresponding angles. The dispersion coefficient C_9^{ABC} describes

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

the interaction between three virtually interacting dipoles and is approximated from the pairwise coefficients as

$$C_9^{ABC} = -\sqrt{C_6^{AB}C_6^{AC}C_6^{BC}}. \quad (1.9)$$

The applicability of this atom-pairwise dispersion correction with three-body corrections in dense molecular systems was shown in a number of recent publications^{33,103,104}.

For early precursors of DFT-D3 also in the framework of HF theory see Ref^{105–109}. Related to the D3 scheme are approaches that also compute the C_6 coefficients specific for each atom (or atom pair) and use a functional form similar to Equation 1.5. A system dependency of the dispersion coefficients is employed by all modern DFT-D variants. We explicitly mention the works of Tkatchenko and Scheffler^{28,29} (TS: 'atom-in-molecules' C_6 from scaled atomic volumes), Sato *et. al.*¹¹⁰ (use of a local atomic response function), and Becke and Johnson^{99,111,112} (XDM: utilizes a dipole-exchange hole model). The TS and XDM methods are used routinely in solid-state applications^{113–116}.

1.2.2. Evaluation of Dispersion Corrected DFT

X23 Benchmark Set

A benchmark set for non-covalent interactions in solids consisting of 21 molecular crystals (dubbed C21) was compiled by Johnson⁴⁰. Two properties for benchmarking are

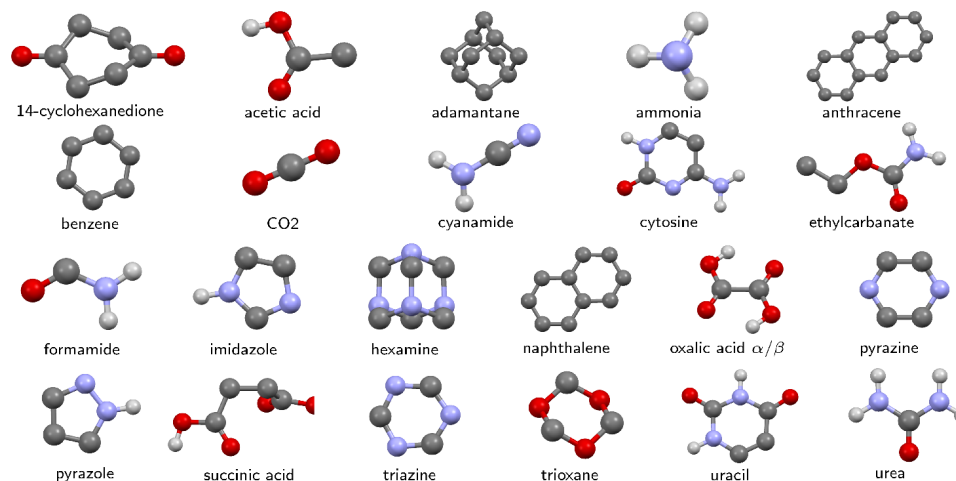


Figure 1.2.: Geometries of the 23 small organic molecules in the X23 benchmark set for non-covalent interactions in solids. Hydrogen atoms at carbons are omitted for clarity. Carbons are denoted by dark gray balls, hydrogens are light gray, oxygens are red, and nitrogens are light blue.

provided: (1) thermodynamically back-corrected experimental sublimation energies and (2) geometries from low-temperature X-ray diffraction. The error of the experimental

Table 1.1.: Mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) of the calculated, zero-point exclusive sublimation energy from reference values for the X23 test set. The energies and geometries refer to the PBE/1000 eV, PBE-D3/1000 eV, PBE-D3/1000 eV+ $E^{(3)}$ levels. Values for the XDM and TS method are taken from Ref⁴⁰ and the data for 16 systems on the PBE-MBD level from Ref¹¹⁷. Negative MD values indicate systematic underbinding. All energies are in kcal/mol per molecule.

Method	X23 Sublimation Energy		
	MAD	MD	SD
PBE/1000 eV	11.55	-11.55	6.20
PBE-D3/1000 eV	1.07	0.43	1.34
PBE-D3/1000 eV+ $E^{(3)}$	1.21	-0.49	1.65
PBE-XDM/1088 eV	1.50	-0.45	2.12
B86b-XDM/1088 eV	1.37	-0.33	1.91
PBE-TS/1088 eV	3.54	3.50	2.32
PBE-MBD/1000 eV	1.53	1.53	0.95

sublimation energies was estimated to be 1.2 kcal/mol¹¹⁸. Recently, the C21 set was extended and refined by Tkatchenko *et al.*¹¹⁷. The X23 benchmark set (16 systems from Ref¹¹⁷ and data for seven additional systems were obtained from these authors) includes two additional molecular crystals, namely hexamine and succinic acid. The molecular geometries of the X23 set are shown in Fig. 1.2. The thermodynamic back-correction was consistently done at the PBE-TS level. Semi-anharmonic frequency corrections were estimated by solid-state heat capacity data. Further details of the back-correction scheme are summarized in Ref¹¹⁷. The mean absolute deviation (MAD) between both data sets is 0.55 kcal/mol. Because the X23 data seem to be more consistent, we use these as reference. If we take the standard deviation (SD) between both thermodynamic corrections as statistical error measure, the total uncertainty of the reference values is about 1.3 kcal/mol. In the following, all (sublimation) energies and their deviations consistently refer to one molecule (and not the unit cell).

The calculations are carried out with the Vienna Ab-initio Simulation Package VASP 5.3^{119,120}. We utilize the GGA functional PBE¹²¹ in combination with a projector-augmented plane wave basis set (PAW)^{122,123} with a huge energy cut-off of 1000 eV. This corresponds to 200% of the recommended high-precision cut-off. We sample the Brillouin zone with a Γ -centered k -point grid with four k -points in each direction, generated via the Monkhorst-Pack scheme⁶⁸. To simulate isolated molecules in the gasphase, we compute the Γ -point energy of a single molecule in a large unit cell (minimum distance

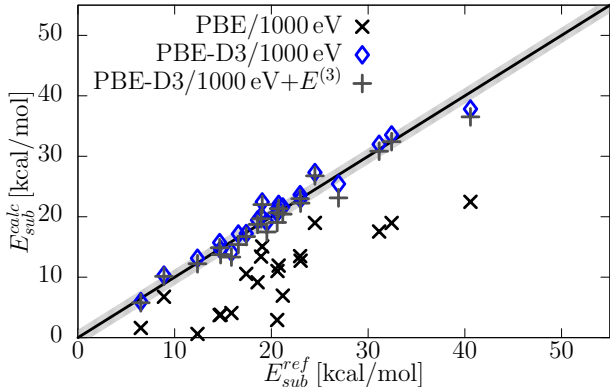


Figure 1.3.: Correlation between experimental and PBE computed sublimation energy with and without dispersion correction. The gray shading along the diagonal line denotes the experimental error interval. All energies are calculated on optimized structures but with experimental lattice constants.

between separate molecules of 16 Å, e.g. adamantane is calculated inside a $19 \times 19 \times 19$ Å³ unit cell). In order to calculate the sublimation energy, we optimize the single molecule and the corresponding molecular crystal. The unit cells are kept fixed at the experimental values. The atomic coordinates are optimized with an extended version of the approximate normal coordinate rational function optimization program (ANCOPT)¹²⁴ until all forces are below 10^{-4} Hartree/Bohr. We compute the D3 dispersion energy in the Becke-Johnson damping scheme with a conservative distance cut-off of 100 Bohr. The three-body dispersion energy is calculated always as a single-point on the optimized PBE-D3/1000 eV structure. The results for X23 are summarized in Table 1.2.2. Fig. 1.3 shows the correlation between experimental sublimation energies and the calculated values on the PBE/1000 eV, PBE-D3/1000 eV, and PBE-D3/1000 eV + $E^{(3)}$ levels. The uncorrected functional yields unreasonable results. Because of the missing dispersion interactions, the attraction between the molecules is significantly underestimated which results in too small sublimation energies. Some systems are not bound at all on the PBE/1000 eV level. For PBE-D3 all results are significantly improved. The MAD is exceptionally low and drops below the estimated experimental error of 1.3 kcal/mol. The mean deviation of +0.4 kcal/mol indicates a slight overbinding on the PBE-D3/1000 eV level. The three-body dispersion correction is always repulsive and therefore decreases the sublimation energy. At the PBE-D3/1000 eV + $E^{(3)}$ level the MAD and SD is slightly raised but these changes are within the uncertainty of the reference data and we hence cannot draw definite conclusions about the importance of three-body dispersion effects from this comparison. Because inclusion of three-body dispersion has been shown to improve the description of binding in large supramolecular structures⁹⁸ and is here not deteriorating the results,

we recommend to always include the term. However, the many-body effect (i.e., adding $E^{(3)}$ to the PBE-D3 data) is smaller than found in recent studies by another group^{29,125} employing a general many-body dispersion scheme. We compare our results to the pairwise dispersion corrections XDM and TS and show the normal error distributions in Fig. 1.4. The XDM model works reasonably well with an MAD of 1.5 kcal/mol, while the TS scheme is significantly overbinding with an MAD of 3.5 kcal/mol. The overbinding of the TS model is partially compensated by large many-body contributions and the MAD on the PBE-MBD level drops to 1.5 kcal/mol. A remarkable accuracy with an MAD of 0.9 kcal/mol was reported with the hybrid functional PBE0-MBD¹¹⁷, but unfortunately only 16 systems of the X23 set were investigated. The XDM model works slightly better in combination with the more repulsive B86b functional. However, the mean deviation of -0.5 kcal/mol and -0.3 kcal/mol reveals a systematic underbinding of the XDM method consistent with results for supramolecular systems¹²⁶. This will lead to a worse result when a three-body term is included.

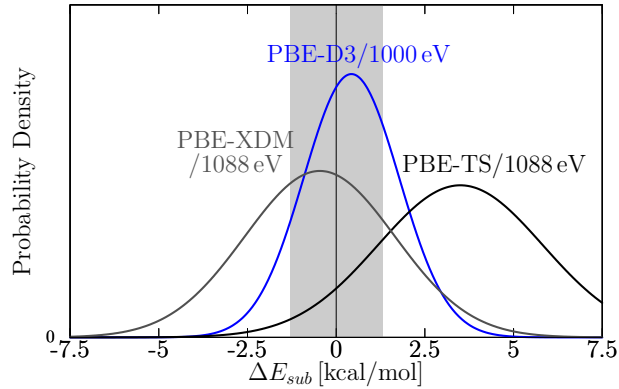


Figure 1.4.: Deviations between experimental and theoretical sublimation energies for the X23 set. We convert the statistical data into standard normal error distributions for visualization. The gray shading denotes the experimental error interval. The quality of the theoretical methods decreases in the following order: PBE-D3/1000 eV, PBE-XDM/1088 eV, and PBE-TS/1088 eV.

As a further test we investigate the unit cell volume for the same systems. We perform a full geometry optimization and compare with the experimental low-temperature X-ray structures. The unit cell optimization is done with the VASP quasi-Newton optimizer with a force convergence threshold of $0.005 \text{ eV}/\text{\AA}$. Without dispersion correction significantly too large unit cells are obtained. On the PBE/1000 eV level, the volumes of the orthorhombic systems are overestimated by 9.7%. We compare the theoretical zero Kelvin geometries with low-temperature X-ray diffraction data at approximately 100 K. Therefore, the calculated values should always be smaller than the measured ones due to thermal expansion effects. After applying the D3 correction, the unit cells are sys-

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

tematically too small by 0.8 % which is reasonable considering typical thermal volume expansions assumed to be approximately 3 %. In passing it is noted that the geometries of isolated organic molecules are systematically too large in volume by about 2 % with PBE-D3¹²⁷, which is consistent with the above findings. In summary, PBE-D3 or PBE-D3+ $E^{(3)}$ provide a consistent treatment of interaction energies and structures in organic solids. Screening effects on the dispersion interaction as discussed in Ref^{29,125} seems to be unimportant in the D3 model.

Structure of Tribenzotriquinazene (TBTQ)

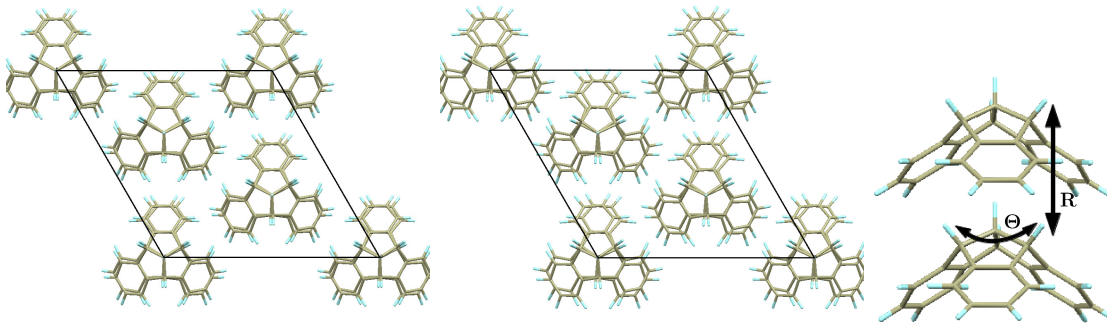


Figure 1.5.: X-ray (left) and PBE-D3/1000 eV (middle) crystal structure of TBTQ. The computed structure was obtained by an unconstrained geometry optimization⁶⁶. The right figure highlights the analyzed geometry descriptors.

As an example for a larger system where London dispersion is even more important, we re-investigate the recently studied tribenzotriquinacene (TBTQ) compound⁶⁶ which involves π -stacked aromatic units. We utilized the GGA functionals PBE¹²¹ and RPBE¹²⁸, a PAW basis set^{122,123} with huge energy cut-off of 1000 eV within the VASP program package. The crystal structures of TBTQ and its centro-methyl derivate (Me-TBTQ) was measured and a space group R3m was found for both TBTQ and Me-TBTQ. However, a refined analysis revealed the true space group of TBTQ to be R3c (an additional c -glide plane), while the space group of Me-TBTQ is confirmed. The structure in Fig. 1.5 shows the tilting between neighboring TBTQ layers. With dispersion corrected DFT (PBE-D3/1000 eV), we were able to obtain all subtle details of the structures as summarized in Table 1.2. The unusual packing induced torsion between vertically stacked molecules was computed correctly as well as an accurate stacking distance. The deviations from experimental unit cell volumes of 1.4 % for TBTQ and 1.5 % for Me-TBTQ are within typical thermal volume expansions. The agreement between theory and experiment is excellent but necessitated a huge basis set with $1.46 \cdot 10^5$ plane-wave basis functions. A calculation of the crystal structure of Me-TBTQ on the same theoretical level confirms

the measured untilted stacking geometry.

Table 1.2.: Comparison of experimental X-ray and computed PBE-D3/1000 eV structures. The first block corresponds to the TBTQ crystal, the second one to the Me-TBTQ crystal. As important geometrical descriptors the vertical stacking distance R , the tilting angle Θ , and the unit cell volume Ω are highlighted. All lengths are given in Å.

				X-ray	PBE-D3/1000 eV		
R				4.75	4.67		
Θ				6.2°	9.8°		
Ω				2075	2046		
a, b, c	15.96,	15.96,	9.48		15.92,	15.92,	9.32
α, β, γ	90.0,	90.0,	120.0		90.0,	90.0,	120.0
R				5.95	5.91		
Θ				0.0°	0.0°		
Ω				2306	2272		
a, b, c	14.96,	14.96,	11.90		14.90,	14.90,	11.82
α, β, γ	90.0,	90.0,	120.0		90.0,	90.0,	120.0

The dispersion correction is also crucial for the correct description of the sublimation energy. For PBE negative values and hence no net bindings are obtained. On the PBE-D3 level reasonable ZPVE-exclusive sublimation energies of 35 kcal/mol and 29 kcal/mol are calculated, which fit the expectations for molecules of this size. In Fig. 1.6, we show the potential energy surface (PES) with respect to the vertical stacking distance for Me-TBTQ. In addition to the PBE functional, we applied the by Hammer *et al.* modified version, dubbed RPBE¹²⁸, to investigate the effect of the short-range correlation kernel. For each point, we perform a full geometry optimization with a fixed unit cell geometry. The curves for both uncorrected functionals show no significant minimum in agreement with the wrong sign of the sublimation energy. Furthermore, we see significant deviations between the two functionals, i.e., PBE is much less repulsive than RPBE. With the inclusion of the D3 correction the differences between both functionals nicely diminishes and the PES are nearly identical. This strongly indicates that the D3 correction provides a physically sound description of long- and medium-range correlation effects. In fact, RPBE-D3 reproduces the equilibrium structure even slightly better than PBE-D3. This confirms previous observations from different groups that dispersion corrections are ideally coupled to inherently more repulsive (semilocal) functionals^{32,129,130}.

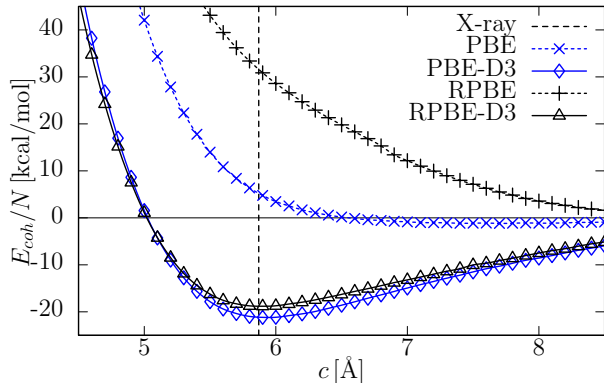


Figure 1.6.: Dependence of the cohesive energy E_{coh} per molecule on the vertical cell parameter c (the dashed line denotes the experimental value). The results refer to the PBE and RPBE functional with a PAW basis set and an energy cut-off of 1000 eV. The cell parameters a and b are fixed to their experimental value. For each point, we perform a full geometry optimization with a fixed unit cell geometry. The asymptotic energy limit $c \rightarrow \infty$ corresponds to the interaction in one Me-TBTQ layer, approximated by a large distance of 15 Å.

1.3. Dispersion Corrected Hartree-Fock with Basis Set Error Corrections

1.3.1. Basis Set Error Corrections

The previously presented results were obtained with huge plane-wave basis sets and these DFT calculations are rather costly. It seems hardly possible to use fewer plane-wave functions, because the stronger oscillating functions are necessary to describe the relatively localized electron density in molecular crystals. A significant reduction of basis functions seems only possible with atom centered functions, i.e., Gaussian atomic orbitals (AO). In contrast to plane waves, however, small AO basis sets strongly suffer from basis set incompleteness errors, especially the BSSE. Already semi-diffuse AOs can exhibit near linear dependencies in periodic calculations and the reduction of the BSSE by systematic improvement of the basis is often not possible. A general tool to correct for the BSSE efficiently in a semiempirical way was developed in 2012 by us⁵⁶. Recently, we extended the gCP denoted scheme to periodic systems and tested its applicability for molecular crystals⁵⁷.

Additionally, the basis set incompleteness error (BSIE) becomes crucial when near minimal basis sets are used. For a combination of Hartree-Fock with a MINIX basis (combination of valence scaled minimal basis set MINIS and split valence basis sets SV, SVP as defined in Ref⁶³), dispersion correction D3, and geometric counterpoise correction

1.3. Dispersion Corrected Hartree-Fock with Basis Set Error Corrections

gCP, we developed a short-ranged basis set incompleteness correction dubbed SRB. The SRB correction compensates too long covalent bonds. These are significant in a HF calculation with very small basis sets, especially when electronegative elements are present. The HF-D3-gCP-SRB/MINIX method will be abbreviated HF-3c in the following. The HF method has the advantage over current GGA functionals that it is (one-electron) self interaction error (SIE) free^{131,132}. Further, it is purely analytic and no grid error can occur. The numerical noise-free derivatives are important for accurate frequency calculations. In contrast to many semiempirical methods, HF-3c can be applied to almost all elements of the periodic table without any further parametrization and the physically important Pauli-exchange repulsion is naturally included. Here, we extend the HF-3c scheme to periodic systems and propose its use as a cheap DFT-D3 alternative or for cross-checking of DFT-D3 results.

The corrected total energy $E_{\text{tot}}^{\text{HF-3c}}$ is given by the sum of the HF energy $E^{\text{HF/MINIX}}$, dispersion energy $E_{\text{disp}}^{\text{D3}}$, BSSE correction $E_{\text{BSSE}}^{\text{gCP}}$, and short-ranged basis incompleteness correction E_{SRB} :

$$E_{\text{tot}}^{\text{HF-3c}} = E^{\text{HF/MINIX}} + E_{\text{disp}}^{\text{D3}} + E_{\text{BSSE}}^{\text{gCP}} + E_{\text{SRB}}. \quad (1.10)$$

The form of the first term $E_{\text{disp}}^{\text{D3}}$ is already described in section 1.2.1. For the HF-3c method the three parameters of the damping function s_8 , a_1 , and a_2 were refitted in the MINIX basis (while applying gCP) against reference interaction energies¹³³ and is denoted D3(refit). The second correction, namely the geometrical counterpoise correction gCP^{56,57}, depends only on the atomic coordinates and the unit cell of the crystal. The difference in atomic energy e_A^{miss} between a large basis (def2-QZVPD¹³⁴) and the target basis set (e.g. the MINIX basis) inside a weak electric field is computed for free atoms A . The e_A^{miss} term measures the basis incompleteness and is used to generate an exponentially decaying, atom-pairwise repulsive potential. The BSSE energy correction $E_{\text{BSSE}}^{\text{gCP}}$ reads

$$E_{\text{BSSE}}^{\text{gCP}} = \frac{\sigma}{2} \sum_{A \neq B}^{\text{atom pairs}} \sum_{\vec{T}} e_A^{\text{miss}} \frac{\exp\left(-\alpha \cdot \left\| \vec{r}_B - \vec{r}_A + \vec{T} \right\|^\beta\right)}{\sqrt{S_{AB} \cdot N_B^{\text{virt}}}}, \quad (1.11)$$

with Slater-type overlap integral S_{AB} , number of virtual orbitals on atom B in the target basis set N_B^{virt} , and basis set dependent fit parameters σ , α , and β . The Slater exponents of s- and p-valence orbitals are averaged and scaled by a fourth fit parameter η to get a single s-function exponent. For each combination of Hamiltonian (DFT or HF) and basis set, the four parameters were fitted in a least-squares sense against counterpoise correction data obtained by the Boys-Bernardi scheme⁹⁰.

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

Systematically overestimated covalent bond lengths for electronegative elements are corrected by the third term E_{SRB} :

$$E_{\text{SRB}} = -\frac{s}{2} \sum_{A \neq B}^{\text{atom pairs}} \sum_{\vec{T}} (Z_A Z_B)^{3/2} \exp \left(-\gamma (R_{AB}^{0,\text{D3}})^{3/4} \left\| \vec{r}_B - \vec{r}_A + \vec{T} \right\| \right). \quad (1.12)$$

We use the default cut-off radii $R_{AB}^{0,\text{D3}}$ as determined *ab initio* for the D3 dispersion correction and $Z_{A/B}$ are the nuclear charges. The parameters s and γ were determined by fitting the HF-3c total forces against B3LYP-D3/def2-TZVPP¹³⁵ equilibrium structures of 107 small organic molecules. Altogether, the HF-3c method consists of nine empirically determined parameters, three for the D3 dispersion, four in the gCP scheme, and two for the SRB correction. The HF-3c method was recently tested for geometries of small organic molecules, interaction energies and geometries of non-covalently bound complexes, for supramolecular systems, and protein structures⁶³ and good results superior to traditional semiempirical methods were obtained. In particular the accurate non-covalent HF-3c interactions energies for a standard benchmark¹³³ (i.e., better than with the 'costly' MP2/CBS method and close to the accuracy of DFT-D3/'large basis') are encouraging for application to molecular crystals.

1.3.2. Evaluation of Dispersion and Basis Set Corrected DFT and HF

We evaluate the basis corrections gCP and SRB by comparison with reference sublimation energies for the X23 benchmark set, introduced in section 1.2.2. We calculate the HF and DFT energies with the widely used crystalline orbital program CRYSTAL09^{136,137}. In the CRYSTAL code, the Bloch functions are obtained by a direct product of a superposition of atom-centered Gaussian functions and a \mathbf{k} dependent phase factor. We use raw HF, the GGA functional PBE¹²¹ and the hybrid GGA functional B3LYP^{138,139}. The Γ -centered k -point grid is generated via the Monkhorst-Pack scheme⁶⁸ with four k -points in each direction. The large integration grid (LGRID) and tight tolerances for coulomb and exchange sums (input settings: TOLINTEG 8 8 8 8 16) are used. The SCF energy convergence threshold is set to 10^{-8} Hartree. We exploit the polarized split-valence basis set SVP¹⁴⁰ and the near minimal basis set MINIX. The atomic coordinates are optimized with the extended version of the approximate normal coordinate rational function optimization program (ANCOPT)¹²⁴.

Mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) of the sublimation energy for the X23 test set and for the subset X12/Hydrogen (systems dominated by hydrogen bonds) are presented in Table 1.3.2. The dispersion and

Table 1.3.: Mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) of the computed sublimation energy with respect to experimental reference data for the X23 test set and for the subset X12/Hydrogen dominated by hydrogen bonds. We compare the HF-3c method with gCP corrected PBE-D3/SVP and B3LYP-D3/SVP methods. For PBE/SVP level, we also give deviations to the corresponding large plane-wave basis set values in parentheses. All values are in kcal/mol per molecule.

Method	X23			X12/Hydrogen	
	MAD	MD	SD	MAD	MD
PBE-D3/SVP	8.5 (8.1)	8.5 (8.1)	3.5 (3.4)	10.5 (9.7)	10.5 (9.7)
PBE-D3-gCP/SVP	2.5 (2.1)	-1.1 (-1.5)	3.0 (2.6)	2.8 (2.5)	-1.4 (-2.3)
PBE-D3-gCP/SVP+ $E^{(3)a}$	2.9 (2.0)	-2.0 (-1.5)	3.2 (2.5)	3.1 (2.4)	-2.2 (-2.2)
B3LYP-D3/SVP	10.1	10.1	4.1	12.0	12.0
B3LYP-D3-gCP/SVP	2.0	0.5	2.3	1.7	-0.1
B3LYP-D3-gCP/SVP+ $E^{(3)a}$	1.7	-0.4	2.2	1.8	-0.8
HF/MINIX ^b	11.3	-11.3	6.1	10.7	-10.7
HF-D3(refit)/MINIX ^b	6.3	6.3	3.6	7.5	7.5
HF-D3(refit)-gCP/MINIX ^b	1.6	0.5	1.9	1.8	-0.0
HF-3c	1.7	0.6	2.0	1.8	0.0
HF-3c+ $E^{(3)a}$	1.5	-0.2	2.0	2.0	-0.7

^a Three-body dispersion $E^{(3)}$ as single-point energy on optimized structures. ^b Single-point energies on HF-3c optimized structures.

BSSE corrected PBE-D3-gCP/SVP and B3LYP-D3-gCP/SVP methods yield good sublimation energies with MADs of 2.5 kcal/mol and 2.0 kcal/mol, respectively. The artificial overbinding of the gCP-uncorrected DFT-D3/SVP methods is demonstrated by the huge MD of 8.5 kcal/mol for PBE and 10.1 kcal/mol for B3LYP. Adding the three-body dispersion energy changes the MADs for D3-gCP to 2.9 kcal/mol and 1.7 kcal/mol, respectively. As noted before⁵⁷, the PBE functional with small basis sets underbinds hydrogen bonded systems systematically. The HF-3c calculated sublimation energies are of very good quality with an MAD of 1.7 kcal/mol and 1.5 kcal/mol without and with three-body dispersion energy, respectively, which is similar to the previous PBE-D3/1000 eV results. Considering the simplicity of this approach, this result is remarkable. The MD is with 0.6 kcal/mol and -0.2 kcal/mol, respectively, also very close to zero. This indicates that with the three correction terms, most of the systematic errors of pure HF are eliminated. For hydrogen bonded systems the MAD is only slightly higher which indicates an overall consistent treatment. To analyze the HF-3c method in more detail, we investigate the different energy contribution to the sublimation energy on the optimized HF-3c structures as shown in Fig. 1.7.

Plain HF is not capable of describing the intermolecular attraction in the crystals and has the largest MAD of 11.3 kcal/mol. The only significant physical attraction between the

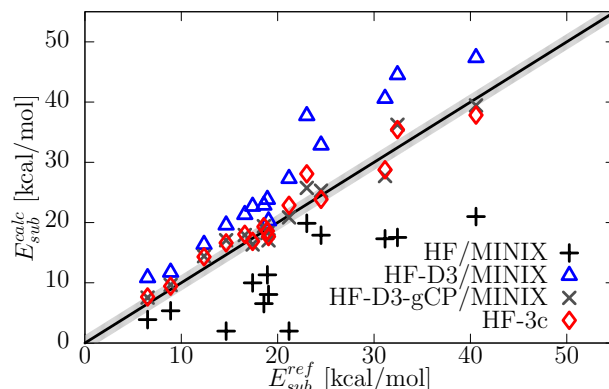


Figure 1.7.: Correlation between experimental sublimation energy and HF results with subsequent addition of the three corrections. All sublimation energies are calculated on optimized HF-3c structures for experimental lattice constants. The gray shading along the diagonal line denotes the experimental error interval.

molecules arises in hydrogen bonded systems which are dominated by electrostatics which is properly described by HF. By inclusion of dispersion, the MAD drops to 6.3 kcal/mol on the HF-D3(refit)/MINIX level, but the sublimation energy is significantly overestimated. This too strong attraction can be efficiently and accurately corrected with the gCP scheme. The MAD on the HF-D3(refit)-gCP/MINIX level is 1.6 kcal/mol and very similar to the MAD of the full HF-3c method. This demonstrates that the SRB correction mainly affects geometries as intended. Because the energy decomposition analysis is done for fixed geometries, we cannot investigate the importance of the E_{SRB} contribution in more detail. In conclusion, the computationally very cheap HF-3c method provides encouraging energies. However, for a few systems we encounter convergence problems of the SCF procedure with the CRYSTAL09 code. This can be sometimes avoided with tighter tolerances for Coulomb and exchange integral sums with the side effect of increased computational cost. Zero point vibrational energies are not analyzed here, but numerically stable second energy derivatives of HF-3c were reported in Ref⁶³.

1.4. Conclusions

We presented and evaluated dispersion corrected Hartree-Fock and Density Functional Theory for their potential to computed organic crystals and their properties. For a correct description of molecular crystals, semilocal (hybrid) density functionals have to be corrected for London dispersion interactions. A variety of modern DFT-D methods, namely D3, TS/MBD, and XDM can calculate sublimation energies of small organic crystals with errors close to the experimental uncertainty. For the X23 test set we found that the D3

scheme gives the best performance of the tested additive dispersion corrections with an MAD of 1.1 kcal/mol, which is well below the estimated error range of 1.3 kcal/mol. In

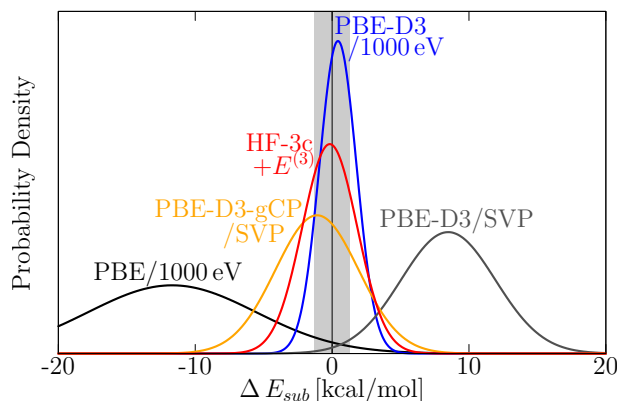


Figure 1.8.: Deviations between experimental and theoretical sublimation energies for the X23 set. We convert the statistics into standard normal error distributions for visualization. The gray shading denotes the experimental error interval. The quality of the theoretical methods decreases in the following order: PBE-D3/1000 eV, HF-3c+ $E^{(3)}$, PBE-D3-gCP/SVP, PBE-D3/SVP, and PBE/1000 eV.

the DFT-D3 scheme the three-body dispersion energy corrections are approximately 5% of the sublimation energy. The finding that the method, which has been developed originally for molecules and molecular complexes, can be applied without further, solid-state specific modifications is encouraging. It was furthermore shown that DFT-D3 can calculate the π -stacking of tribenzotriquinacene and its centro-methyl derivative with all subtle geometry details. This example demonstrates that larger molecules routinely considered in organic chemistry can be treated also in their solid state accurately by DFT based methods.

In addition to these calculations with huge plane-wave based basis sets, we exploited Gaussian atom-centered orbitals. We demonstrated the large basis set errors on the DFT-D3/SVP and HF-D3/MINIX levels and presented and evaluated two semiempirical basis set corrections. The resulting DFT-D3-gCP/SVP and HF-3c methods perform well and especially the MAD of 1.5 kcal/mol (with three-body dispersion) for HF-3c is remarkable. However, the SCF convergence with unscreened Fock-exchange is sometimes problematic and despite of a larger basis used, the PBE-D3-gCP/SVP calculations converge faster and yield an acceptable MAD of 2.5 kcal/mol for the X23 sublimation energies.

In Fig. 1.8, we summarize the results of the various theoretical methods for the X23 benchmark set by converting the statistical data into standard normal distributions. The best results are calculated with the D3 dispersion corrected PBE functional in a huge PAW basis set. HF-3c+ $E^{(3)}$ and PBE-D3-gCP/SVP can also be recommended.

1. Dispersion Corrected HF and DFT for Organic Crystal Structure Prediction

In future work, the description of energy rankings of polymorphs on the different theoretical levels has to be investigated systematically. Furthermore, coupling of the D3 dispersion correction to different GGA, meta-GGA, and hybrid GGA functionals might provide even better performance. In any case, the future for fully quantum chemical based first principles crystal structure prediction seems bright.

2. Benchmarking DFT and Semiempirical Methods on Structures and Lattice Energies for Ten Ice Polymorphs

Jan Gerit Brandenburg*, Tilo Maas*, and Stefan Grimme*

Keywords: Ice, Polymorphism, Non-Covalent Interaction, London Dispersion Correction, Density Functionals

Received 2nd of February 2015, Published online 26th of March 2015

Reprinted (adapted) with permission from

J. G. Brandenburg, T. Maas, and S. Grimme, J. Chem. Phys. **2015**, *142*, 124104.

— Copyright ©2015 American Institute of Physics. DOI 10.1063/1.4916070

Own manuscript contribution

- Supervision of all computations
- Calculation of structures and energies by low-cost methods
- Calculation of energies with hybrid DF
- Performing the quasi-anharmonic back-correction of X-ray geometries
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

Abstract Water in different phases under various external conditions is very important in bio-chemical systems and for material science at surfaces. Density functional theory methods and approximations thereof have to be tested system specifically to benchmark their accuracy regarding computed structures and interaction energies. In this study, we present and test a set of ten ice polymorphs in comparison to experimental data with mass densities ranging from 0.9 to 1.5 g/cm³ and including explicit corrections for zero-point vibrational and thermal effects. London dispersion inclusive density functionals at the generalized gradient approximation (GGA), meta-GGA, and hybrid level as well as alternative low-cost MO methods are considered. The widely used PBE functional systematically overbinds and overall provides inconsistent results. All other tested methods yield reasonable to very good accuracy. BLYP-D3^{atm} gives excellent results with mean absolute errors for the lattice energy below 1 kcal/mol (7 % relative deviation). The corresponding optimized structures are very accurate with mean absolute relative deviations (MARD) from the reference unit cell volume below 1%. The impact of Axilrod-Teller-Muto (atm) type three-body dispersion and of nonlocal Fock exchange is small but on average their inclusion improves the results. While the density functional tight-binding model DFTB3-D3 performs well for low density phases, it does not yield good high density structures. As low-cost alternative for structure related problems we recommend the recently introduced minimal basis Hartree-Fock method HF-3c with MARD of about 3%.

2.1. Introduction

Computationally efficient electronic structure methods are nowadays extensively used in (bio)chemistry, solid state physics, and material science. In this regard, density functional theory (DFT) has emerged as 'work horse' for many applications and is still an active research field of general interest.^{73–76,141} DFT provides an excellent compromise between accuracy and computational cost. Even more efficient semiempirical methods have gained an increased importance for large scale screenings of numerous conformers and in the field of molecular dynamics beyond the classical force field approximation. However, both semilocal DFT and semiempirical approximations thereof are not capable of describing long-range electron correlation effects leading to the important London dispersion interactions.^{23–25,77} In the last decade this flaw and its correction was an intense research topic and an explicit account of London dispersion is now standard in DFT and semiempirical frameworks.^{26–30} For further details on these methods we refer to references^{31,32,64} with some review character.

Water in its various phases is of utmost importance in biological systems.¹⁴² The special

physical chemistry of water systems covers thermodynamical properties, critical phenomena, and chemical reactions.¹⁴³ Theoretical methods shall help and guide experimentalists in this regard which requires an accurate treatment of the various condensed phases of water. This primarily involves the description of (mostly non-covalent) interaction energies and resulting structures. Recently, some efforts were undertaken to perform relatively high-level MP2 and random phase approximation (RPA) (energy and gradient) calculations of liquid water and ice, but these are presently only possible with huge computational resources.^{144,145} Especially the dynamics of biomolecules in water and of water at solid surfaces is an active research field with many challenges.^{146–148}

Because DFT and semiempirical methods by construction contain some empirical elements, their careful benchmarking is mandatory. In the past, most DFT benchmarks focused on isolated molecules, dimers and small clusters (e.g. Refs.^{21,38,127}). Presently, only one common benchmark set for organic solids exists.^{39,40} In this so-called X23 set, only two polymorphs are included and systems with strong hydrogen bonds are under-represented. Previous studies investigated some ice polymorphs with mostly PBE based density functionals with and without corrections for London dispersion effects.^{149,150} Recently, Kresse and coworkers applied RPA to various ice modifications and for some structures embedded many-body expansions at the 'gold standard' CCSD(T) level have been used.^{50,51}

In the present study we investigate a selection of ten experimentally studied ice polymorphs. The performance of various density functionals (DF) at the generalized gradient approximation (GGA), meta-GGA, hybrid, and (range separated) hybrid level is investigated. Both, the structure and corresponding lattice energy are analyzed. Additionally, some low cost molecular orbital (MO) based methods which showed promising accuracy previously⁶⁴ are tested. Furthermore, the importance of an accurate treatment of London dispersion even in systems dominated by hydrogen-bonding is highlighted.

We first present the ten ice polymorphs under consideration with the experimental references in section 2.2.1. Section 2.2.2 shortly summarizes the computational details. We correct the X-ray structures for zero-point and thermal effects as described in section 2.2.3. The main results of this study are given in section 2.3, separated into a potential surface analysis of the high density ice VIII (2.3.1), a structure benchmark (2.3.2), a lattice energy benchmark (2.3.3), and a comparison to results for gas phase water clusters (2.3.4). Finally, a conclusion with recommended methods is given in section 5.4.3.

2.2. Benchmark Setup

2.2.1. Systems under Consideration

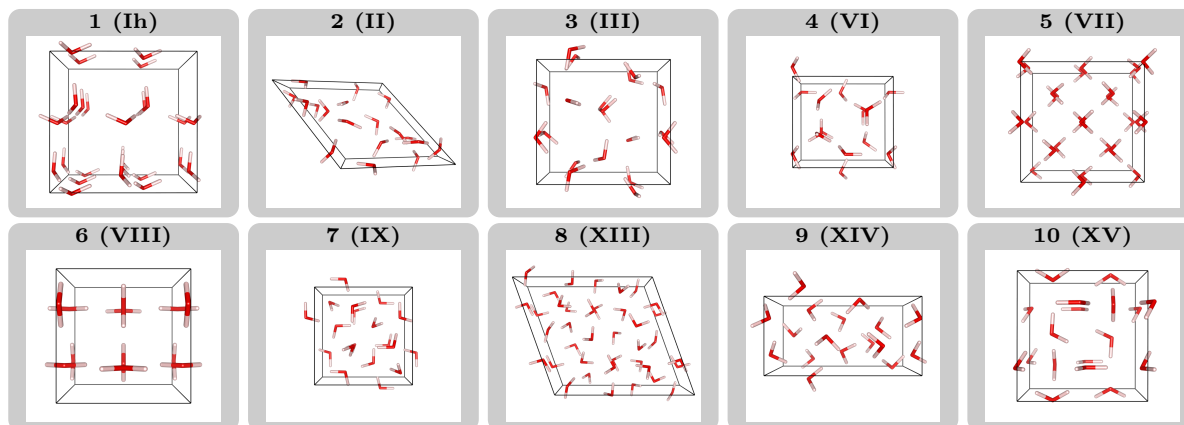


Figure 2.1.: Systems contained in the ICE10 benchmark set. The perspective projection of a single unit cell is shown.

We have compiled the ICE10 set by combining ten different ice polymorphs as summarized in Table 2.1. Their structures were determined by low temperature neutron diffraction experiments. The systems contain 8 to 28 water molecules per crystal unit cell and all obey the 'ice rules' (Bernal-Fowler rules).¹⁵¹ We consider four proton ordered and six proton disordered systems. The proton ordered crystals typically occur at higher densities, where the hydrogens are at fixed positions with low entropy. The unit cell volumes (per molecule) varies between 18 and 32 Å³. Because the experimental detection and accurate placement of hydrogen atoms is challenging, we use the theoretical unit cell volume (which corresponds to a certain mass density) as most sensitive and reliable structure quality criterion. The sublimation enthalpies of systems 1-7 were determined experimentally and extrapolated to electronic (zero-point vibrational exclusive) lattice energies at 0 K. These reference energies can be directly compared to the calculated ones in order to judge the quality of a theoretical method. For systems 7-10 no experimental sublimation data are available. We give theoretical estimates in Table 2.1 at our 'best' theoretical level, which can be used as reference for methods at a lower theoretical level like GGA density functionals, small basis set calculations, or semiempirical methods.

In figure 2.1 we show a single unit cell for each crystal. One can see the variety of different hydrogen bond networks with proton ordered conformations and proton unordered structures. Table 2.1 summarizes all important properties of the ten crystals. When multiple experimental data are available, we give both volumes and mass densities and use

Table 2.1.: Systems contained in the ICE10 benchmark set. Crystallographic specifications, experimental densities (measured at temperature T_{exp}), and lattice energies extrapolated to 0 K are given.

No.	polymorph	#H ₂ O	Bravais lattice	spacegroup	protons ^a	T_{exp}	vol.	ρ	E_{lat}
1	Ih ^{152,153}	16	hexagonal	P6 ₃ /mmc	disordered	10	32.05, 32.50	0.93, 0.92	14.07
2	II ^{154,155}	12	rhombohedral	R $\bar{3}$	ordered	0	24.97, 24.63	1.20, 1.21	14.05
3	III ¹⁵⁶	12	tetragonal	P4 ₁ 2 ₁ 2	disordered	90	25.69	1.16	13.85
4	VI ¹⁵⁷	10	tetragonal	P4 ₂ /nmc	disordered	98	22.84	1.31	13.68
5	VII ¹⁵⁶	16	cubic	Pn $\bar{3}$ m	disordered	90	20.26	1.48	13.07
6	VIII ^{158,159}	8	tetragonal	I4 ₁ /amd	ordered	0	20.09, 18.61	1.49, 1.61	13.31
7	IX ^{160,161}	12	tetragonal	P4 ₁ 2 ₁ 2	ordered	30	25.63, 25.80	1.17, 1.16	13.97
8	XIII ¹⁶²	28	monoclinic	P2 ₁ /a	ordered	80	23.91	1.25	13.95 ^b
9	XIV ¹⁶²	12	orthorhombic	P2 ₁ 2 ₁ 2 ₁	ordered	80	23.12	1.29	13.74 ^b
10	XV ¹⁶³	10	tetragonal	P1	ordered	80	22.45	1.33	13.48 ^b

^a Unit cell volume per molecule is given in Å³, molecular mass density in g·cm⁻³, lattice energy (at 0 K) in kcal·mol⁻¹, and temperature in K. ^b Theoretical estimate at the PBE0-D3^{atm} level with removed systematic shift.

the latest published values in comparisons with theory.

2.2.2. Computational Details

The DFT calculations are mostly conducted with the VASP program package.^{119,164} The projector augmented plane-wave method (PAW) is used with hard pseudo-potentials constructed by Blöchl and Kresse.^{122,123} In order to approach the basis set limit, a huge PAW energy cutoff of 1000 eV is used. Detailed convergence tests showed that this is required in unconstrained geometry optimizations. In smaller basis sets, artificial Pulay stress can lead to too small unit cell volumes. For instance, an optimized Ih structure with PAW cutoff of 600 eV has a 2 % smaller unit cell compared to the 1000 eV basis set calculation. That the basis set limit is indeed reached was confirmed by a corresponding potential energy scan. Similar effects have been observed before in organic crystals.⁶⁶ In the following, the 1000 eV PAW basis is used if not mentioned otherwise.

We apply several GGA functionals (PBE¹²¹, RPBE¹²⁸, revPBE¹⁶⁵, BLYP^{166,167}), the meta-GGAs TPSS¹⁶⁸ and M06L¹⁶⁹, two global hybrid functionals (PBE0¹⁷⁰ and B3LYP^{138,139}), and the range-separated hybrid functional HSE06.⁸⁹ Because of the significantly higher computational demands, the hybrid functionals are only used for single-point energy calculations. The single-point energy calculations were consistently done on the PBE-D3 structures. The non-covalent geometries are typically less sensitive to the inclusion of Hartree-Fock exchange compared to the impact on the lattice energy. Though some func-

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

tionals are shown to yield better geometries, we choose the PBE-D3 level. This is the standard procedure conducted by our group and it is consistent with other approaches, e.g., Tkatchenko and coworkers calculate the PBE0-MBD energies on PBE-TS geometries.^{39,149}

Though ice forms complex networks of hydrogen bonds with large electrostatic and induction contributions to the binding energy, London dispersion forces cannot be neglected. In order to investigate its importance, all methods are applied with and without the D3 London dispersion correction.²⁶ It is applied in the most recent Becke-Johnson damping variant.⁸⁷ Only M06L is used with the zero damping scheme (denoted by D3(0)) to minimize the double counting of short-range dispersion effects, which are covered by the meta-GGA. For final single point energies, our standard three-body dispersion term (of Axilrod-Teller-Muto type^{102,171}) is included and will be indicated by the superscript atm.

Due to the increased importance of low-cost methods as a bridge between first-principles DFT and classical force fields, some alternative approaches are tested as well. We conduct plain B3LYP/6-31G* calculations with atom-centered Gaussian basis sets of double-zeta quality with the CRYSTAL14 program.^{172,173} Additionally, a minimal basis set Hartree-Fock approach with corrections for dispersion (D3), basis set superposition error (gCP^{56,57}) and short-range basis set error (SRB), dubbed HF-3c, is applied.^{41,63} HF-3c frequencies are scaled by 0.86 as suggested in its original publication.

As the computationally fastest here considered method, the dispersion corrected density functional tight-binding DFTB3-D3^{59,62} is applied. We use the latest third order version with empirical damping of hydrogen containing pair-potentials and self-consistent charge redistribution as implemented in dftb+.^{174,175} The 3OB Slater-Koster files constructed by Elstner and coworkers are used.¹⁷⁶ The latter two methods showed excellent performance on benchmark sets for general non-covalent interactions.⁶⁴ The D3, gCP, and HF-3c methods are implemented in a CRYSTAL14 developer version and will be generally available in its next release.

The Brillouin zone is sampled with dense k grids of approximately 0.02 \AA^{-1} generated via the Monkhorst-Pack scheme. The Γ -centered number of k -points are given in the SI (A). Structures are fully optimized without symmetry constraints until all forces are below 0.005 eV/\AA . Especially for flat potential energy surfaces, tight convergence thresholds are necessary.⁶⁵ In the PAW calculations for a single (isolated) water molecule a large unit cell (12 \AA) is employed to minimize the interaction with its periodic images.

2.2.3. Correction for Zero-Point Energies

The measurements of structure and lattice energy have been conducted at finite temperature (up to 100 K). The experimental lattice energies provided by Whalley were extrapolated to 0 K and zero-point vibrational energies (ZPVE) have been removed.¹⁵⁶ The accuracy of the extrapolated energies has been verified by high-level diffusion quantum Monte-Carlo (DMC) and embedded many-body CCSD(T) calculations.^{51,177} We further assume that the thermal contribution to the ice density (from 0 to 98 K) is rather small and the measured structure can be treated as equilibrium (R_e) geometry as discussed further below.

However, the ZPVE is substantial and can not be neglected. In order to provide an easy use-able benchmark, we estimate its effect on the unit cell volume (and the mass density, respectively) for all ten systems. We transform the experimental volumes V_0 into back-corrected reference equilibrium volumes V_e . These can be directly compared to the structures of free optimizations on the electronic energy surface. For each system, we perform constrained (constant volume) optimizations around the electronic equilibrium geometry with scaled unit cell volumes of 80, 90, 95, 100, 105, 110, 120, and 130% V_e . On these optimized structures, the vibrational frequencies are computed in the harmonic approximation. The ZPVE together with the Bose-Einstein occupied phonon modes lead to the free energy

$$F(V) = E(V) + \sum_q \left(\frac{\hbar\omega_q(V)}{2} + \frac{\hbar\omega_q(V)}{e^{\frac{\hbar\omega_q(V)}{k_B T}} - 1} \right). \quad (2.1)$$

The phonon modes with frequencies $\omega_q(V)$ depend on the volume V . Therefore, the correction to the electronic energy also depends on the volume. To the data points $F(\{V_i\})$ and $E(\{V_i\})$, we fit a Murnaghan equation of state

$$E, F(V) = E_{e,0} + \frac{B_{e,0}V}{B'_{e,0}(B'_{e,0} - 1)} \times \left(B'_{e,0} \left(1 - \frac{V_{e,0}}{V} \right) + \left(\frac{V_{e,0}}{V} \right)^{B'_{e,0}} - 1 \right) \quad (2.2)$$

and extract the equilibrium volumes V_e and the free energy volumes V_0 , respectively. The bulk modulus B and its derivative B' are not further analyzed.

As shown below, the HF-3c method provides very reasonable potential energy surfaces at rather low computational costs. This makes it an ideal choice for the free energy calculations. Therefore, the total energy $E_{tot}(V)$ in the above scheme is evaluated at the HF-3c^{atm} level (ATM three-body term included). In order to judge the accuracy of this HF-3c based back-correction, we compare it for system Ih and VIII with values from

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

Mullay and Galli.¹⁷⁸ They calculate vibrational corrections to the volume at the vdW-DF2 level of 1.9 and 5.8 %, respectively. The deviation to the corresponding HF-3c corrections (3.6 and 5.1 %) is rather small. While they also compute PBE based corrections, we believe that the comparison to the apparently wrong PBE potential (see below) should be avoided. Because the two examples cover both extremes (high and low density) of the ice phases, we expect the results to be transferable to the other systems. In our experience, the steepness of the intermolecular DFT potential typically increases with higher HF exchange, which is systematically removed by the frequency scaling. The HF-3c potential energy surface of ice Ih is shown in figure 2.2. Apparently, the equation

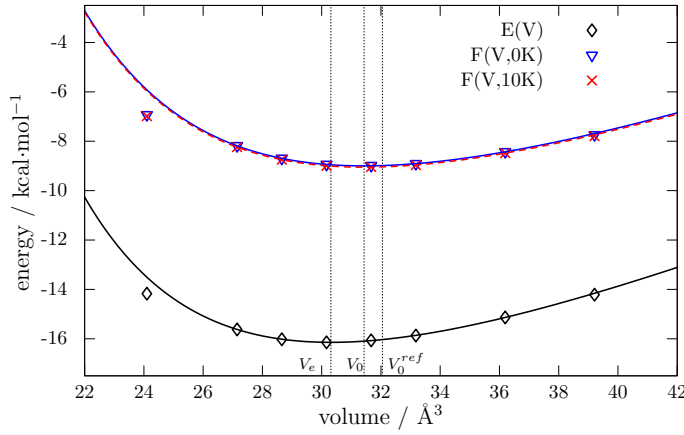


Figure 2.2.: Electronic energy surface $E(V)$ and free energy surface $F(V)$ on the HF-3c level for the ice Ih polymorph compared to the experimental volume V_0^{ref} . The free energy F includes zero-point and thermal contributions in the harmonic approximation.

of state is a good model around the equilibrium. The ZPVE has two effects. First, it shifts the minimum to higher energies by approximately 8 kcal/mol. Additionally, the free energy minimum occurs at $\approx 4\%$ larger unit cell volume. Compared to the significant ZPVE effect, the impact of the three-body dispersion (ATM) term is with 0.3% rather small. The impact of the thermal-vibrational contributions at these low temperatures is tiny ($< 0.1\%$) and the construction of a large supercell for the phonon calculation is not necessary. The unit cell volume extracted from the free energy surface V_0 is very close to the experimental value with a deviation below 2 %.

In Table 2.2, we summarize the equilibrium volumes calculated by HF-3c together with the measured volumes V_0 and the 'experimental' equilibrium volumes V_e^{ref} . If different experimental measurements are available, we use the latest published values (first value

Table 2.2.: Correction of equilibrium volumes V_e to free-energy volumes V_0 due to ZPVE and thermal energies calculated at the HF-3c^{atm} level.

No.	theory			experiment		
	V_0	V_e	$\Delta V/V_0$	V_0	$V_e^{ref} (\pm 1.2\%)$	ρ_e^{ref}
1	31.43	30.31	3.6%	32.05	30.91	0.96
2	24.67	23.68	4.0%	24.97	23.97	1.25
3	26.97	26.55	1.6%	25.69	25.29	1.18
4	21.85	21.03	3.8%	22.84	21.98	1.36
5	18.86	17.96	4.8%	20.26	19.28	1.55
6	18.93	17.96	5.1%	20.09	19.06	1.57
7	25.81	25.14	2.6%	25.63	24.97	1.20
8	23.29	22.43	3.7%	23.91	23.03	1.30
9	22.56	21.72	3.7%	23.12	22.26	1.34
10	22.01	21.15	3.9%	22.45	21.58	1.38

^aExperimental V_e estimated as $V_e^{ref} = V_0^{ref} \left(1 + \frac{V_e^{calc} - V_0^{calc}}{V_0^{calc}} \right)$ with V^{calc} at the HF-3c level.

given for volume and density in Table 2.1). We propose the back-corrected volumes V_e^{ref} as reference benchmark values, which can be compared to free optimizations on the electronic energy surface. The increase of the volume due to free energy contributions is on average 3.7% with a standard deviation of 1.0%. We observe an increased correction with increasing density (linear correlation with correlation coefficient $R^2 = 0.7$), which is expected. However, this is only a rough trend and the specific volume expansion depends non-trivially on the geometry.

2.3. Results and Discussion

2.3.1. Potential Energy Surface of Ice VIII

We first investigate the potential energy landscape of the high density ice polymorph VIII in some detail. We calculate the electronic potential energy surface (PES) by scaling the lattice vectors of the minimum geometry and performing a constraint volume optimization as described in the previous paragraph. This illustrative example is shown in figure 6.1 for PBE, PBE-D3, RPBE-D3, revPBE-D3, TPSS-D3, TPSS-D3^{atm}, HF-3c^{atm}, and DFTB3-D3^{atm} to demonstrate the typical behavior of these methods. For an easier comparison, we show in each plot the PBE and PBE-D3 potential for comparison.

The minima of all methods determined via the PES fit agree well with the free optimizations. Especially for the calculations in the PAW basis this is an important test for basis set completeness and the applied geometry convergence criteria. Plain PBE gives significantly too large lattice parameters resulting in an unit cell volume which is over-

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

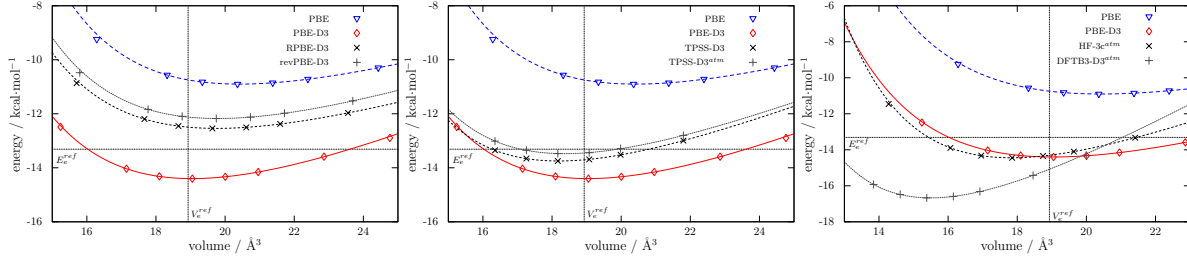


Figure 2.3.: Electronic energy surface of the ice VIII polymorph on various theory levels compared to the experimental references. Zero-point and thermal effects are explicitly excluded from the reference energy E_e^{ref} and reference volume V_e^{ref} .

estimated by 7 %. The corresponding lattice energy of 10.9 kcal/mol is too small by 2.4 kcal/mol. The D3 dispersion correction leads to much better structures, i.e., the unit cell volume of PBE-D3 deviates from the reference by less than 0.2 %. However, the minimum is significantly too low (overestimated lattice energy). The PBE functional is known to overestimate hydrogen bonding¹⁷⁹ which is very pronounced in the ice crystals with many strong hydrogen bonds. The physically correct inclusion of London dispersion lowers the energy further and leads to the observed effect. The revised versions RPBE and revPBE were constructed to give more reasonable energies for these kind of systems. Both potentials are rather similar. The lattice energy is indeed improved (error of RPBE-D3 below 1 kcal/mol). However, the minimum is shifted to 3.7 % too large unit cell volumes. The behavior that revPBE sometimes deteriorates the good PBE structures was already recognized by the authors of PBE for covalent bonding.¹⁸⁰ Of all PBE variants and successors, respectively, TPSS-D3 performs best with accurate and consistent lattice energy and structure. The inclusion of the ATM term has small effects, but overall improves the results. For the geometries, we estimate the three-body contribution by a potential energy scan (see section 2.2.3) and scale the corresponding TPSS-D3 unit cell volumes. The lattice energy of TPSS-D3^{atm} deviates by only 0.2 kcal/mol from the reference and the cell is 3.6 % too small, which we consider reasonably good. Interestingly, the internal structures of covalently bound, medium-sized molecules are systematically too large at the TPSS-D3 level, and this also holds for the X23 set of molecular crystals.^{127,181} Apparently, this cannot be directly transferred to the ice phases where the density mainly depends on the non-covalent hydrogen bond lengths.

The computational cost of all non-hybrid DFT methods is practically identical and only the meta-GGA TPSS is slightly more expensive than the GGAs. Two low cost methods, namely HF-3c and DFTB3-D3, are included in figure 6.1. In the HF-3c method, the

Table 2.3.: Comparison of calculated unit cell volumes of the ICE10 benchmark set with the experimental reference.

	1	2	3	4	5	6	7	8	9	10	MRD	MARD ^a
Exp. reference (V_e^{ref})	30.9	24.0	25.3	22.0	19.3	19.1	25.0	23.0	22.3	21.6	–	–
DFT												
PBE	30.2	24.5	26.4	22.3	20.4	20.4	26.2	23.6	22.9	22.4	3.1	3.6
RPBE	33.7	29.0	30.5	26.9	26.2	26.2	30.2	28.0	27.4	27.0	23.7	23.7
revPBE	33.1	28.0	29.8	25.9	24.6	24.5	29.8	27.1	26.4	26.0	19.1	19.1
BLYP	32.2	26.3	29.0	23.9	22.0	22.0	27.5	25.3	24.5	24.1	10.9	10.9
TPSS	30.5	24.8	26.7	22.5	20.0	19.9	27.0	23.9	23.2	22.6	3.9	4.1
DFT-D												
PBE-D3	29.1	23.2	24.4	20.9	18.9	19.1	24.1	22.2	21.5	21.1	-3.2	3.2
RPBE-D3	30.6	24.4	25.6	22.0	19.7	19.6	25.4	22.2	22.6	22.1	1.0	1.8
revPBE-D3	30.4	24.2	25.4	21.7	19.7	19.7	25.3	23.1	22.3	21.8	0.7	1.3
BLYP-D3	30.4	23.9	24.9	21.4	19.4	19.4	24.5	22.8	22.0	21.5	-0.8	1.3
TPSS-D3	29.3	23.2	24.6	20.8	18.2	18.2	24.5	22.2	21.5	21.0	-3.8	3.8
DFT-D^{atm}												
PBE-D3 ^{atm}	29.2	23.3	24.6	21.0	19.0	19.2	24.3	22.4	21.6	21.2	-2.7	2.8
RPBE-D3 ^{atm}	30.7	24.6	25.8	22.1	19.8	19.8	25.6	22.4	22.7	22.2	1.6	2.3
revPBE-D3 ^{atm}	30.5	24.3	25.7	21.9	19.9	19.8	25.5	23.2	22.4	22.0	1.3	1.7
BLYP-D3 ^{atm}	30.5	24.0	25.2	21.5	19.5	19.5	24.7	22.9	22.1	21.6	-0.3	1.0
TPSS-D3 ^{atm}	29.4	23.3	24.8	21.0	18.3	18.3	24.7	22.4	21.6	21.1	-3.2	3.2
Low-Cost												
B3LYP/6-31G*	30.1	23.7	24.9	21.3	19.1	19.1	24.7	22.9	21.9	21.4	-0.5	2.1
HF-3c ^{atm}	30.3	23.7	26.6	21.0	18.0	18.0	25.1	22.4	21.7	21.2	-2.2	3.3
DFTB3-D3 ^{atm}	30.1	21.8	18.7	18.7	15.5	15.5	20.6	19.4	19.4	18.6	-15.1	15.1

^a Volumes are given in Å³. Mean relative deviations (MRD) and mean absolute relative deviations (MARD) are given in %.

Hartree-Fock part is evaluated in a minimal GTO basis set, which leads to a speed-up of ≈ 50 compared to the DFT/PAW calculations. Notably, however, the corresponding PES is close to the reference. The lattice energy is overestimated by 1.1 kcal/mol and the unit cell is by 6.4 % too small. Especially, the good structure and reasonable PES shape is important for the low-cost methods which are often used for geometry optimization and frequency calculations. In the tight-binding model DFTB3, no three-, and four-center integrals have to be evaluated. This yields an additional speed up of two orders of magnitude compared to HF-3c but on the other hand introduces some significant errors. The important many-center contributions are missing, which leads to a wrong repulsive potential, especially at high densities. The lattice energy is overestimated, but the deviation of 3.4 kcal/mol is acceptable. The system specific results discussed so far are supported by the statistical analysis of the whole ICE10 benchmark set given in the next two sections.

2.3.2. Equilibrium Structures for ICE10

In the following, we analyze the equilibrium structures of the full ICE10 benchmark set. Correct structures are in general important for the application of more involved

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

electronic structure methods. For instance an embedded cluster approach with high-level correlated wavefunction methods relies on accurate input geometries. Especially for localized correlation methods, analytical gradients are not feasible and therefore good DFT structures are needed. The ICE10 systems are all optimized with the (meta-)GGA density functionals and the three low-cost methods. Note that the optimization with a hybrid functional in a nearly complete PAW basis is currently not feasible on standard work stations. Optimizations with the meta-GGA M06L are not performed because of numerical problems (SCF convergence).

In Table 2.3, the experimental and theoretical unit cell volumes are summarized. We additionally give the mean relative deviation (MRD) and the mean absolute relative deviation (MARD) with respect to the reference using 10 data points. For selected methods, we show a graphical representation of the unit cell volume in figure 2.4.

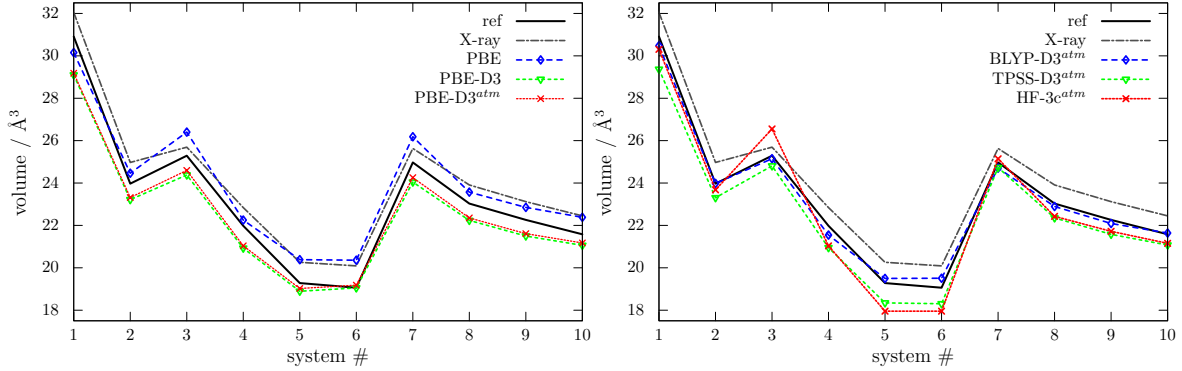


Figure 2.4.: Unit cell volume in \AA^3 of the ten ice polymorphs for selected methods compared to the reference. Note the significant deviation between the raw X-ray data and the back-corrected reference values.

All plain semilocal density functional approximations (PBE, RPBE, revPBE, BLYP, TPSS) produce significantly too large volumes with MARDs ranging from 4 to 24 %. The behavior of the PBE variants is very different, though their mathematically similar functional form. Inclusion of the D3 dispersion correction clearly improves the results and all MARDs are below 4 %. This demonstrates the importance of long-range London dispersion effects even though a significant contribution to the binding in ice is due to electrostatic and induction effects. Interestingly, the PBE-D3 volumes are underestimated, while the revised PBE variants overestimate them. The effect of including the three-body dispersion is rather small, but on average improves the results. In general, all dispersion corrected density functionals evaluated in a large PAW basis set can be recommended. BLYP-D3^{atm} performs exceptionally good with an MARD of only 1 % without systematic

shift. This is in agreement with recent results for the properties of liquid water^{182,183}

However, the PAW based free optimizations can be computationally expensive. Therefore, we tested the low-cost methods introduced in the previous paragraph. Somewhat surprisingly, B3LYP in a small 6-31G* basis set performs well. This is due to an error compensation between the missing long-range London dispersion and the artificial basis set superposition error¹⁸⁴. The good result for the ice crystals can not be generally transferred to other systems. For instance, plain B3LYP in a similar SVP basis set was shown to perform badly on various molecular crystals.⁵⁷ HF-3c is evaluated in an even smaller minimal basis set, but performs only slightly worse compared to the dispersion corrected functionals. The MARD is only 3.3 % and at the same time a speed up of 50–100 is achieved compared to the TPSS-D3/PAW calculations. We have recently shown that the combination of the DFTB3 model with the D3 correction can describe various organic molecular crystals reasonably well.⁶² Unfortunately, the geometries of the ice crystals are not satisfactory. The volumes are on average too small by 15 %, and the largest errors occur for the high density phases. Only the structure of the most stable low density ice Ih can be reproduced with DFTB3-D3 to within 3 %. The missing three- and four-center integrals and corresponding empirical pair potentials in the method are especially important for short distances, which explains the bad performance of the tight-binding model for the various high density phases.

Because the unit cell volume as a quality measure could hide systematic error compensations between the different cell dimensions, we additionally investigate the individual cell parameters. The cell data and statistics are given in the SI (A). The deviations from the references are distributed very uniformly in all directions. The MARDs of the unit cell lengths are approximately 1/3 (to 1/2) of the corresponding unit cell volume MARDs and the ranking of the various methods persists.

2.3.3. Electronic Lattice Energy

We have shown in the last paragraph that London dispersion forces are crucial to get correct cell volumes. The three-body dispersion effects are small but on average improve the results. BLYP-D3^{atm} is the best performing method concerning the structures. However, if one aims at screening of very many structures (polymorphs), the electronic lattice energy is the most important property.

In Table 2.4, we give the lattice energies of all systems together with the mean deviation (MD) and mean absolute deviation (MAD) in kcal/mol. The statistical data correspond to the seven systems with experimental references. In addition to the previously studied (meta-)GGAs, we show values for the meta-GGA M06L and the hybrid functionals PBE0,

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

B3LYP, and HSE06, which are evaluated at the PBE-D3 geometries. The statistics of selected methods are shown in figure 2.5.

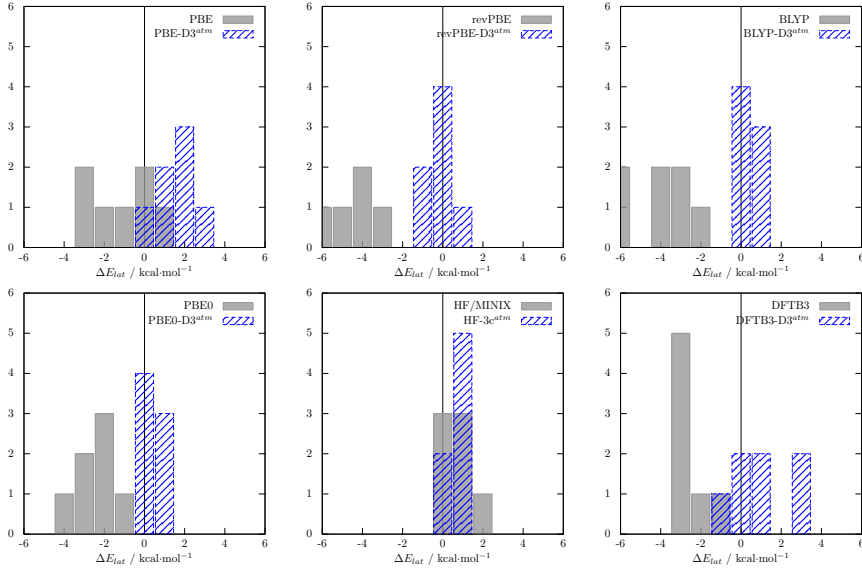


Figure 2.5.: Deviation of selected methods from the reference lattice energy. Histograms are given with 1 kcal/mol bin width. Error distributions of all functionals are given in the SI (A).

Analysis of the energy data supports the trends already observed. Plain PBE is surprisingly good due to error cancellation between the overestimated hydrogen bond energy and missing long-range dispersion. The other plain functionals fail to describe the systems properly. M06L includes to some extent short-range dispersion effects, but has an unsatisfactory MAD of 1.6 kcal/mol with a systematic underbinding tendency. Apparently, the correct long range behavior is important in ice. If these contributions are included by the D3(0) scheme, the MAD diminishes to 1.3 kcal/mol and the MD is close to zero. Apart from PBE (which has the intrinsic hydrogen bond errors), all functionals are significantly improved by the dispersion correction. Both revised PBE versions have an exceptionally small MAD of about 0.5 kcal/mol. The effect of the three body dispersion is small, but again its inclusion improves overall the results. The best performing GGA is revPBE-D3^{atm}. It has a tiny MAD of 0.5 kcal/mol, which corresponds to a MARD of 3.8%. BLYP-D3^{atm} also performs well with an MAD of 1 kcal/mol (MARD of 6.9%).

Inclusion of nonlocal exchange is rather costly in converged basis sets. However, all hybrid functionals have consistently smaller errors compared to their GGA counterparts with MADs usually below 1 kcal/mol. For instance PBE0-D3^{atm} has an MAD of only 0.8 kcal/mol with MARD of 6.1% which is significantly better than PBE-D3. We use

Table 2.4.: Comparison of calculated lattice energies of the ICE10 benchmark set to the experimental reference.

	1	2	3	4	5	6	7	8	9	10	MD	MAD ^a
Exp. reference (E_{lat}^{ref})	14.1	14.1	13.9	13.7	13.1	13.3	14.0	–	–	–	–	–
DFT												
PBE	15.3	13.5	14.1	12.6	10.9	10.9	14.2	13.3	14.0	12.6	-0.6	1.1
RPBE	11.3	9.4	10.2	8.6	6.8	6.8	10.3	9.3	9.0	8.5	-4.7	4.7
revPBE	11.5	9.5	10.3	8.5	6.7	6.7	10.4	9.3	8.9	8.5	-4.7	4.7
BLYP	12.8	10.8	11.5	9.9	8.0	8.0	11.5	10.6	10.3	9.8	-3.3	3.3
TPSS	14.0	11.9	12.6	10.9	9.1	9.1	12.8	11.7	11.3	10.8	-2.2	2.2
M06L	11.2	12.0	11.6	12.4	13.5	13.4	11.6	12.0	12.2	12.3	-1.5	1.6
DFT-D												
PBE-D3	17.4	16.3	16.6	15.7	14.4	14.4	16.7	16.2	16.0	15.6	2.2	2.2
RPBE-D3	15.1	14.1	14.4	13.7	12.5	12.5	14.5	14.0	13.9	13.6	0.1	0.5
revPBE-D3	15.2	14.4	14.6	14.0	12.8	12.8	14.7	14.3	14.1	13.9	0.4	0.6
BLYP-D3	15.9	15.2	15.3	14.8	13.7	13.7	15.5	15.1	15.0	14.7	1.1	1.1
TPSS-D3	16.4	15.2	15.6	14.7	13.5	13.4	15.7	15.1	14.9	14.6	1.2	1.2
M06L-D3(0)	12.0	12.9	12.5	13.3	14.6	14.5	12.5	13.0	13.1	13.3	-0.5	1.3
DFT-D^{atm}												
PBE-D3 ^{atm}	17.2	16.1	16.4	15.5	14.2	14.2	16.5	16.0	15.8	15.4	2.0	2.0
RPBE-D3 ^{atm}	15.0	14.0	14.3	13.5	12.3	12.3	14.4	13.8	13.7	13.4	0.0	0.6
revPBE-D3 ^{atm}	15.1	14.2	14.5	13.8	12.6	12.6	14.6	14.1	13.9	13.7	0.2	0.5
BLYP-D3 ^{atm}	15.8	15.0	15.2	14.6	13.4	13.4	15.3	15.0	14.8	14.5	1.0	1.0
TPSS-D3 ^{atm}	16.3	15.1	15.4	14.4	13.2	13.2	15.6	15.0	14.7	14.4	1.0	1.1
M06L-D3(0) ^{atm}	11.9	12.7	12.3	13.1	14.3	14.2	12.4	12.8	12.9	13.1	-0.7	1.3
DFT-D^{atm} Hybrid												
PBE0-D3 ^{atm}	15.8	14.9	15.0	14.4	13.4	13.3	15.2	14.8	14.6	14.3	0.8	0.8
B3LYP-D3 ^{atm}	15.5	14.9	14.9	14.5	13.6	13.6	15.1	14.8	14.7	14.4	0.8	0.8
HSE06-D3 ^{atm}	15.7	14.9	15.2	14.4	13.3	13.3	15.4	14.9	14.6	14.4	0.9	0.9
Low-Cost												
B3LYP/6-31G*	21.1	20.2	20.5	19.7	18.1	18.1	20.5	18.3	19.9	19.6	6.0	6.0
HF-3c ^{atm}	16.0	15.3	15.3	14.6	14.2	14.2	15.6	15.2	14.8	14.6	1.3	1.3
DFTB3-D3 ^{atm}	13.4	14.4	14.9	14.7	16.7	16.7	14.4	14.7	14.5	14.9	1.3	1.5

^a Energies, mean deviations (MD), and mean absolute deviations (MAD) are given in kcal/mol.

the PBE0-D3^{atm} lattice energies for systems 8-10 shifted by the negative MD on systems 1-7 as best theoretical estimates for these systems where experimental data are lacking. These values should be robust enough to benchmark less accurate small basis DFT, semiempirical, and classical force fields.

The low-cost methods have a larger error spread. Plain B3LYP/6-31G* has a large error of 6 kcal/mol and can not be recommended. Both HF-3c^{atm} and DFTB3-D3^{atm} are rather good with MADs only slightly above the DFT-D methods (MARD of 9.4 and 11.2%, respectively).

Other dispersion inclusive DFT methods perform very similar. The Tkatchenko-Scheffler (TS) method²⁸ is constructed similarly to D3 (use of pre-calculated C₆ coefficients) and can also be applied as a correction to standard density functionals. PBE+TS results in similarly overbound crystals. In combination with PBE0, it is slightly worse compared to PBE0-D3^{atm}. The corresponding many-body dispersion has a similar magnitude as

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

Table 2.5.: Statistical data of the best DFT-D3 methods with other dispersion inclusive DFT methods for the ICE10 lattice energies.

	MD	MAD	RMS ^a
revPBE-D3 ^{atm}	0.18	0.52	0.65
PBE0-D3 ^{atm}	0.85	0.85	1.00
PBE+TS ^b	1.92	1.92	0.66
PBE0+TS ^b	1.16	1.16	1.24
PBE0+MBD ^b	1.02	1.02	1.12
optPBE-vdW ^b	1.36	1.36	1.36
RPA@PBE ^c	1.63	1.63	1.63

^a Mean deviations (MD), mean absolute deviations (MAD) and root-mean-square deviation (RMS) are given in kcal/mol. ^b Values take from Ref. ¹⁴⁹. Note that only four data points are available. ^c Values take from Ref. ⁵⁰. Note that only four data points are available.

the three-body ATM term and leads to minor improvements, the MAD of PBE0+MBD is close to 1 kcal/mol. The van der Waals density functional optPBE-vdW includes a nonlocal kernel combined with an adjusted semilocal DF part. The performance on the ice systems is reasonable, but worse than the DFT-D methods. Also the explicitly correlated RPA method evaluated on PBE orbitals performs worse than DFT-D. This could be significantly improved by replacing PBE with its hybrid variant (including 25 to 50% HF exchange) as already recognized by Kresse and coworkers.⁵⁰

2.3.4. Comparison to Gas Phase Water Clusters

In order to put the above results into a broader perspective, we compare our results to the neutral systems of the WATER27 set compiled by Goddard, dubbed WATER14 in the following.¹⁸⁵ It consists of water oligomers with up to 20 molecules. The reference energies are at the estimated CCSD(T)/CBS level of theory, while for the largest clusters (20H₂O) MP2 values are used. Recently, new reference energies at the incremental CCSD(T) level with tighter basis set convergence have been published.¹⁸⁶ Especially for the neutral systems, the deviations are small and we decided to use the original references. In figure 2.6 we show the MARDs on both benchmark sets for a selection of methods. The typical functional behaviors are consistently found for both sets. The artificially overly attractive PBE functional leads to the bad performance of PBE-D3^{atm} in both phases. The revised version revPBE-D3^{atm} performs significantly better. BLYP-D3^{atm} is the best performing GGA on the WATER14 set and similar good results are obtained for the solid state with an MARD below 7%. Similarly, both meta-GGAs perform consistently well. Note that for good results with M06L, the correct long-range dispersion contribution by the D3(0) treatment is important. The two hybrid functionals have MARDs between 5

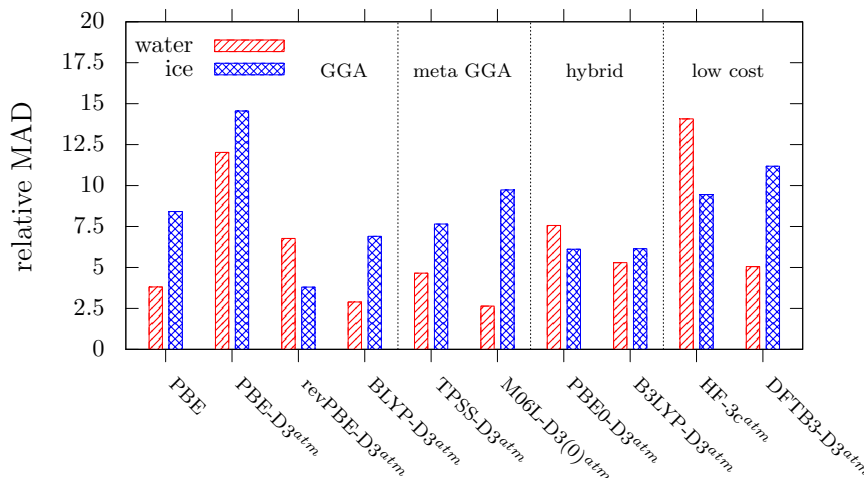


Figure 2.6.: Performance of various methods for interaction energies of water in the gas and solid state evaluated as mean relative absolute deviation (MARD) on the benchmark sets WATER14 and ICE10.

and 7% for both sets. Especially for the gas phase clusters the low-cost methods show more scatter. HF-3c has a rather large error of 14% while DFTB3-D3 has a very low error of 5%. These differences are not observed for the ICE10 set where HF-3c and DFTB3-D3 both have a reasonable MARD around 10%, respectively. Despite some differences, a similar performance of the tested methods for both the gas and the solid state is noted. If the DFTB3-D3 value is excluded, the linear correlation coefficient between the gas/solid MARDs of the shown methods is 0.73.

2.4. Conclusions

We presented a set of ten ice polymorphs ranging from low density (0.9 g/cm^3) to high density (1.5 g/cm^3) phases. The X-ray structural data were back-corrected for zero-point and thermal effects at the HF-3c level in order to get equilibrium structures on the electronic energy surface for convenient benchmarking. The experimental sublimation energies for seven systems have been extrapolated to 0 K lattice energies by Whalley.¹⁵⁶ On these “experimental” equilibrium geometries and electronic lattice energies several dispersion inclusive density functional approximations as well as some selected low-cost methods were benchmarked.

An accurate treatment of nonlocal London dispersion interaction is shown to be mandatory for an accurate description of both, the structures and energies. All dispersion cor-

2. Benchmarking DFT and Semiempirical Methods on Ice Polymorphs

rected GGA functionals yield very reasonable structures with MARD from the reference unit cell volume of approximately 2–3%. The corresponding lattice energies are accurate and close to or below the 'chemical accuracy' of 1 kcal/mol. Especially BLYP-D3^{atm} performs excellently with MARD of the unit cell volumes and MAD of the lattice energy below 1% and 1 kcal/mol, respectively. Compared to corresponding GGAs, the hybrid density functionals improve the lattice energies slightly and can be recommended for obtaining best (routinely) possible energies. While dispersion uncorrected PBE provides reasonable (but not very good) results, the overall picture is more consistent with other (inherently more repulsive) dispersion corrected GGAs. From the investigated low-cost methods, we recommend HF-3c for geometry optimization including frequency calculations. This conclusion is in agreement with the good performance of HF-3c for noncovalently bound organic complexes and solids.⁶⁴ For an analysis of the individual non-covalent interaction terms and their compensations in dispersion-corrected minimal basis set HF calculations see Ref.⁷ The structures of the tight-binding model DFTB3-D3 have to be taken with care. The lattice energies at this level are rather good, especially when considering the tremendous speed up of approximately 2-3 orders of magnitude compared to full DFT calculations.

In summary, we have shown a hierarchy of methods which are ideal suited to describe ice at various densities. Some comparisons with water clusters furthermore indicate that the conclusions are transferable to the liquid state where detailed benchmarking studies are hampered by the sampling problem. With the best performing theoretical models, also e.g. large solid water interfaces should be described quantitatively. Especially the cheaper methods can be used in molecular dynamic simulations, where full DFT calculations are often prohibitive in terms of computational cost.

Part III.

Development of Low-Cost Methods

The benchmarking of DFT-D methods in part II showed that absolute lattice energies of small organic crystals can be calculated within (or close to) the chemical accuracy of 1 kcal/mol. This corresponds to a relative error of approximately 5–8%. The unit cell geometries deviate by 1–3% from the experimental reference. The highest accuracy can only be achieved in huge single-particle basis sets close to the CBS. These calculations are rather costly, and in the CSP context, faster methods with computational cost and overall reasonable accuracy (between the DFT-D/'CBS' and classical force fields) are required.

The easiest way to speed-up a DFT or HF calculation is to decrease the one-particle basis set. Molecular crystals have rather localized electron densities, which can be represented most efficiently by atom-centered Gaussian functions (in contrast to the plane-waves in the PAW approach). A basis set reduction introduces errors, which are most pronounced when calculating binding energies of noncovalently bound complexes in the supermolecular approach. The additional basis functions in the complex (or crystal) increase the variational space which artificially lowers the energy. This is most pronounced for medium sized basis sets of double- ζ quality as sketched in Figure 2.7. In a minimal basis set, the neighboring fragment has only few (to zero) virtual orbitals and the extension of the variational freedom is minor (small BSSE). In a complete basis set, the virtual space is huge, but the energy gain is zero because it is already converged in the single fragment basis (zero BSSE). With medium sized basis sets, the increase in virtual functions and the corresponding lowering of the energy can be substantial. In general,

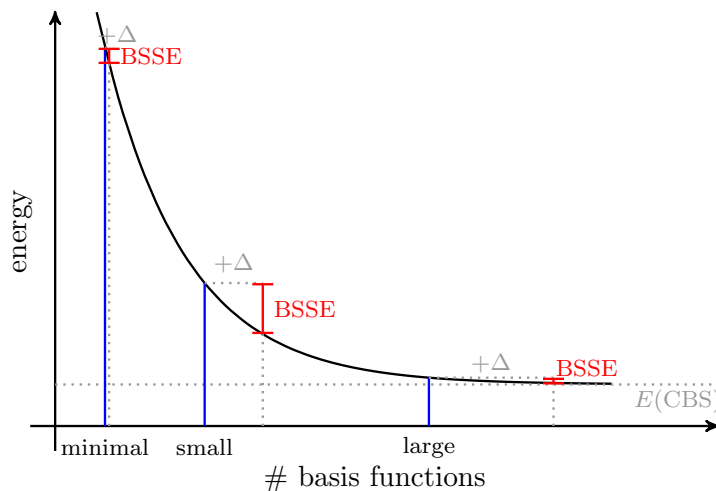


Figure 2.7.: Sketch of the BSSE effect on binding energies. While a quadruple- ζ basis set can be considered as large (nearly complete for SCF methods), a double- ζ basis set is considered as rather small with substantial BSSE.

2. Development of Low-Cost Methods

BSSE leads to overestimated binding energies and underestimated binding distances.

There are several ways to correct the BSSE. An exceptionally efficient approach, the geometrical counterpoise correction (gCP) was introduced for molecular systems in 2012⁵⁶ and is explained in Chapter 3. The scheme is extended to periodic boundaries and evaluated on a number of benchmark sets. A new basis set has been parametrized and applied to selected systems. A minimal basis set HF or DFT calculation is also the theoretical foundation for the even faster semiempirical MO methods. Additionally, many-center terms are neglected and replaced by empirical pair-potentials. In Chapter 4 we present a well established third order density functional tight-binding model (DFTB3). It is coupled to the D3 London dispersion correction resulting in excellent performance for organic solids. Finally, an established HF method evaluated in a minimal basis with corrections for London dispersion, BSSE, and basis set incompleteness error (BSIE) is further tuned to give an even larger speed-up. This is done by replacing the core electrons of all elements by effective core potentials (ECPs) and reducing the contraction of the valence shells. The method, termed HF-3cv, is presented in Chapter 5. Its performance is compared to competing semiempirical methods with focus on noncovalent binding energies in large and periodic systems. The different methodologies are classified according to their specific approximations and corresponding computational costs in Figure 2.8. It is shown, that the correct asymptotic behavior of the energy contributions, specifically the C_6/r^6 limit of the London dispersion energy, is most important in periodic systems.

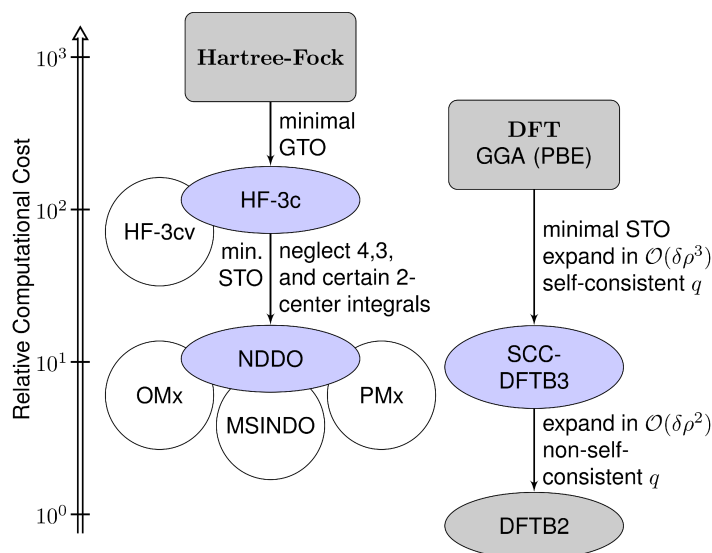


Figure 2.8.: Classification of low-cost, minimal basis set based quantum chemical methods according to their specific approximations. Reproduced from Chapter 5.

3. Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic Density Functional Theory Calculations

Jan Gerit Brandenburg*, Maristella Alessio[†], Bartolomeo Civalleri[†], Michael F. Peintinger*, Thomas Bredow*, and Stefan Grimme*

Keywords: Quantum Chemical Calculations, Organic Crystals, Counterpoise Correction, Dispersion Correction

Received 5th of July 2013, Published online 15th of August 2013

Reprinted (adapted) with permission from

J. G. Brandenburg, M. Alessio, B. Civalleri, M. F. Peintinger, T. Bredow, and S. Grimme, J. Phys. Chem. A **2013**, *117*, 9282–9292.

— Copyright ©2013 American Chemical Society. DOI 10.1021/jp406658y

Own manuscript contribution

- Implementation of the method into CRYSTAL
- Benchmark calculations and basis set parametrization

[†]Department of Chemistry and Centre of Excellence NIS, University of Turin, Via P. Giuria 7, 10125 Torino, Italy

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstraße 4, 53115 Bonn, Germany

3. *Correction for the Inter- and Intramolecular BSSE in Periodic DFT*

- Computation on graphite system
- Writing main parts of the manuscript

Abstract We extend the previously developed geometrical correction for the inter- and intra-molecular basis set superposition error (gCP) to periodic Density Functional Theory (DFT) calculations. We report gCP results compared to those from the standard Boys-Bernardi counterpoise correction scheme and large basis set calculations. The applicability of the method to molecular crystals as the main target is tested for the benchmark set X23. It consists of 23 non-covalently bound crystals as introduced by Johnson *et. al.* (*J. Chem. Phys.* **2012**, *137*, 054103) and refined by Tkatchenko *et. al.* (*J. Phys. Chem. Lett.* **2013**, *4*, 1028). In order to accurately describe long-range electron correlation effects, we use the standard atom-pairwise dispersion correction scheme DFT-D3. We show that a combination of DFT energies with small atom-centered basis sets, the D3 dispersion correction, and the gCP correction can accurately describe van der Waals and hydrogen bonded crystals. Mean absolute deviations of the X23 sublimation energies can be reduced by more than 70% and 80% for the standard functionals PBE and B3LYP, respectively, to small residual mean absolute deviations of about 2 kcal/mol (corresponding to 13 % of the average sublimation energy). As a further test we compute the interlayer interaction of graphite for varying distances and obtain good equilibrium distance and interaction energy of 6.75 Å and -43.0 meV/atom at the PBE-D3-gCP/SVP level. We fit the gCP scheme for a recently developed pob-TZVP solid-state basis set and obtain reasonable results for the X23 benchmark set and the potential energy curve for water adsorption on a nickel (110) surface.

3.1. Introduction

The theoretical description of periodic systems using Density Functional Theory (DFT) or Hartree-Fock (HF) with moderate computational costs is highly desirable. Especially the computation of reliable sublimation energies, geometries and relative energies of molecular crystals and their polymorphic forms is of utmost importance, i.e., aiming at crystal structure prediction^{37,71,72}. The theoretical evaluation of molecular crystals and their polymorphs is an active research field^{7,30,33,45,80,83,84}.

Bulk metals have a strongly delocalized valence electron density and therefore, originless plane-wave based basis sets are probably the best choice in orbital based methods⁸⁹. In molecular crystals, on the other hand, the charge density is more localized. Thus, for plane-wave based methods huge basis sets are needed. We have recently found that for a typical system of stacked organic π -systems, up to $1.5 \cdot 10^5$ projector augmented plane-wave (PAW) basis functions must be considered for accurate results⁶⁶. For this and similar systems, atom-centered basis functions could be much more efficient^{187–190}.

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

Small atom-centered basis sets strongly suffer from basis set errors (BSE), especially the basis set superposition error (BSSE). For molecular systems, a number of different correction schemes exist. The Ref.¹⁹¹ gives a good review of the various approaches. Recently, we have mapped the standard Boys-Bernardi counterpoise correction (BB-CP) onto a semiempirical atom-pairwise potential⁵⁶. This repulsive potential was fitted for a number of typical basis sets and depends only on the system geometry and was therefore denoted as geometrical counterpoise correction (gCP). Analytic gradients are problematic in nearly all other counterpoise schemes, but are easily and efficiently obtained within the gCP. Moreover, the accidental error cancellation in standard small basis set calculations was recently demonstrated for molecular thermochemistry and results could be significantly improved by applying gCP¹⁸⁴.

In this work we include periodic boundary conditions and unit cell gradients in the gCP code. We correct the (semi) local density functionals for missing long-range electron correlation, also known as van der Waals or London dispersion interaction, with the DFT-D3 method²⁶. The theory of both correction schemes, D3 and gCP, is briefly discussed. The proposed approach could also be applied analogously to HF calculations but we restrict ourselves to DFT in this work. Computational details are summarized and a short comparison of gCP and BB-CP energies is given. We test the applicability of the DFT-D3-gCP (or 'functional'-D3-gCP) denoted method on the X23 benchmark set^{40,117} for non-covalent interactions in molecular crystals and for the interaction between graphite layers. Furthermore, gCP parameters are calculated for a new solid-state optimized basis (pob-TZVP) and tested on the X23 sublimation energies, the graphite stacking, and for water adsorption on a nickel (110) surface. Finally, we give a conclusion.

3.2. Theory

We split the BSE into basis set incompleteness error (BSIE) and BSSE. The BSIE arises always, when finite incomplete basis sets are used. The BSSE is due to an inhomogeneous basis, i.e., the variational freedom varies in different spacial regions. In the description of periodic systems, two types of single particle basis sets are the most prominent, (1) a superposition of orthonormal plane-waves spread uniformly in space which is BSSE-free and (2) a direct product of a superposition of atom-centered functions and a momentum (\mathbf{k}) dependent phase factor enforcing the correct translational symmetry. The atom-centered orbitals are typically contracted from primitive Gaussian functions with different exponents to decrease the computational effort. In the dissociation limit, each atom can be accurately described by this set. It is possible to choose the 'most important' atom-

centered basis functions and to minimize the BSIE with a fixed number of basis functions. Moreover, the use of Gaussian basis sets presents some advantages: core electrons can be easily treated, the algorithms are similar to those used in efficient molecular quantum chemistry codes where Gaussian orbitals are generally adopted, and the quasi-local character of electronic structure (at least for insulators) is more naturally exploited. However, incomplete, atom-centered basis sets always favor certain spatial locations.

Two major manifestations of the BSSE arise: (1) In the calculation of cohesive energies the presence of surrounding molecules increases the basis relative to that in the dissociation limit leading to an artificial overbinding; (2) In geometry optimizations the atomic centers move and therefore the basis changes, too. Unit cell volumes are artificially underestimated. A complete basis set is BSE-free, so the best way to avoid both BSIE and BSSE, is to increase the basis set and this is always recommended if technically possible.

Unfortunately, in most periodic systems large atom-centered basis sets are not commonly affordable. This is particularly relevant when diffuse functions with small Gaussian exponents are included. While they are important to describe the tails of molecular density distributions, they are less needed in periodic systems because the overlap between Bloch functions is higher than between AOs in finite systems. Diffuse functions may lead to quasi-linear basis set dependencies which can cause an instability in the convergence of the iterative solution of the Kohn-Sham or HF equations (SCF procedure). Practically, it turns out to be difficult to enlarge the basis to the quadruple- ζ level, even in not very dense solids, although polarized triple- and quadruple- ζ basis sets have been employed occasionally in calculations on molecular crystals^{83,192–194}. Recently, a consistent triple- ζ basis set for periodic systems was developed by some of us, which aims to overcome some of the mentioned problems¹⁹⁵. Therein, basis functions with very small exponents are consistently replaced by more localized contracted orbitals. However, we show below that the BSSE of this basis is still very large for typical cases and must be corrected.

Another problem of HF and common (semilocal) DFT functionals is that they are not capable of describing long-range electron correlation, a.k.a. London Dispersion interaction. In order to get physically reasonable results, the methods have to be properly dispersion corrected³². We decompose the total energy E_{tot} of a system into DFT/HF energy $E_{\text{DFT/HF}}$, dispersion energy E_{disp} , and additional counterpoise correction E_{gCP} :

$$E_{\text{tot}} = E_{\text{DFT/HF}} + E_{\text{disp}} + E_{\text{gCP}}. \quad (3.1)$$

To accurately describe the London dispersion interaction, we use our latest semi-classical *ab-initio* dispersion correction DFT-D3^{26,87}. It incorporates non-empirical, pairwise-specific, chemical environment-dependent dispersion coefficients, a physically sound damp-

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

ing function according to Becke and Johnson⁹⁹, and optionally a non-additive Axilrod-Teller-Muto three-body dispersion term¹⁰². The dispersion energy can be split into two- and three-body contributions $E_{\text{disp}} = E^{(2)} + E^{(3)}$:

$$E^{(2)} = -\frac{1}{2} \sum_{n=6,8} \sum_{A \neq B}^{\text{atom pairs}} \sum_{\mathbf{T}} s_n \frac{C_n^{AB}}{\|\mathbf{r}_B - \mathbf{r}_A + \mathbf{T}\|^n + f(R_0^{AB})^n} \quad (3.2)$$

$$E^{(3)} = -\frac{1}{6} \sum_{A \neq B \neq C}^{\text{atom triples}} \sum_{\mathbf{T}} \frac{C_9^{ABC} (3\cos\theta_a\cos\theta_b\cos\theta_c + 1)}{r_{ABC}^9 \cdot (1 + 6(r_{ABC}/R_0)^{-\alpha})}. \quad (3.3)$$

Here, C_n^{AB} denotes the averaged (isotropic) n^{th} -order dispersion coefficient for atom pair AB , and $\mathbf{r}_{A/B}$ are their Cartesian positions. The real-space summation over all unit cells is done by considering all translation invariant vectors \mathbf{T} inside a cut-off sphere. The scaling parameter s_6 equals unity for the here applied functionals and ensures the correct limit for large distances, and s_8 is a functional-dependent scaling factor. The rational damping function $f(R_0^{ab})$ is

$$f(R_0^{ab}) = a_1 R_0^{ab} + a_2, \quad R_0^{ab} = \sqrt{\frac{C_8^{ab}}{C_6^{ab}}}. \quad (3.4)$$

The dispersion coefficients are obtained from first-principle calculations for molecular systems with time-dependent DFT and application of the Casimir-Polder relation⁹⁷. The three parameters s_8 , a_1 , and a_2 are fitted for each functional (or HF) on a benchmark set of small, non-covalently bound molecules. The fitting is necessary to prevent double counting of dispersion interactions at short range and to adjust repulsive and attractive parts. This procedure has an impact on the short- to medium-ranged part of the dispersion energy, but does not affect the long-range regime, which is most important for periodic systems. In the three-body contribution, r_{ABC} corresponds to an averaged distance between the three pairs and $\theta_{a/b/c}$ are the corresponding angles. The general applicability of this atom-pairwise dispersion correction was shown in a number of recent publications by us^{98,103,179} and other groups^{33,196,197} and also for solids^{66,101,104,198,199}. The importance of three- and many-body dispersion effects is currently not clear but this question is intensively investigated^{29,200,201}.

The most frequently used method to correct for the BSSE is the Boys-Bernardi counterpoise (BB-CP)⁹⁰ method. For a dimer AB with basis functions a and b , respectively, it reads

$$E_{\text{BB-CP}} = \{E(A)_a - E(A)_{ab}\} + \{E(B)_b - E(B)_{ab}\}, \quad (3.5)$$

where $E(A)_a$ is the energy of fragment A with basis functions a and $E(A)_{ab}$ denotes the

energy of the same fragment A with the enlarged basis ab . In the geometric counterpoise correction scheme, we map the BB-CP onto a semiempirical, repulsive pair potential $V_A(r)$, which decays exponentially with the inter-atomic distance r :

$$V_A(r) = e_A^{\text{miss}} \frac{\exp(-\alpha \cdot r^\beta)}{\sqrt{S \cdot N_B^{\text{virt}}}}. \quad (3.6)$$

The energy difference between a large and the generally smaller target basis set for an atom A inside a weak electric field e_A^{miss} is computed for all atoms. This quantity measures the incompleteness of the atomic target basis. The potential is normalized by the Slater-overlap S , the number of virtual orbitals N_B^{virt} , and the empirical parameters α and β . The Slater exponents of s- and p-valence orbitals ζ_s and ζ_p , respectively, are averaged to get a single s-function exponent ζ_s^*

$$\zeta_s^* = \eta \frac{\zeta_s + \zeta_p}{2}, \quad (3.7)$$

with fit parameter η . In order to satisfy the periodic boundary conditions (PBC), the summation runs over all distinct pairs AB inside a large super-cell of certain radius. Using a real-space cut-off is possible because of the exponential decay, which converges rapidly with increasing distance. We sum over all atom pairs and get the energy E_{gCP} :

$$E_{\text{gCP}} = \frac{\sigma}{2} \sum_{A \neq B}^{\text{atom pairs}} \sum_{\mathbf{T}} V_A(\|\mathbf{r}_A - \mathbf{r}_B + \mathbf{T}\|), \quad (3.8)$$

with a global scaling parameter σ .

Altogether, gCP involves four fit parameters ($\sigma, \alpha, \beta, \eta$) for each basis set. They have been fitted to reproduce the BB-CP energy of the large S66 \times 8 non-covalent interaction benchmark set¹³³ by minimizing the root-mean-square deviation. No parameter re-fit is considered here nor needed here for the applied standard basis sets (the new pob-TZVP basis¹⁹⁵ is discussed separately) because a short-range potential should not be affected by PBC. Parameters are available for a variety of basis sets, namely the Ahlrichs-type basis sets SV, def2-SVP, and def2-TZVP^{135,202,203}, the minimal basis set MINIS^{204,205}, and the Pople-style basis 6-31G*²⁰⁶.

3.3. Computational Details

We calculate the HF and DFT energies mostly with the widely used crystalline-orbital program CRYSTAL09^{136,137}. In the CRYSTAL code, the Bloch functions are obtained by

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

a direct product of a superposition of atom-centered Gaussian functions and a \mathbf{k} dependent phase factor. We use the generalized gradient approximated (GGA) functional PBE^{121,180} and the hybrid GGA functional B3LYP^{138,139}. The Γ -centered \mathbf{k} -point grid is generated via the Monkhorst-Pack scheme⁶⁸ with one \mathbf{k} -point for molecular calculations, four \mathbf{k} -points for molecular crystals, and 12 \mathbf{k} -points in each direction for graphite. The large integration grid (LGRID) and tight tolerances for Coulomb and exchange sums are used. The self-consistent field (SCF) energy threshold is set to 10^{-8} Hartree. We exploit the polarized and unpolarized split-valence basis set SVP and SV and the triple- ζ basis set pob-TZVP. Calculations close to the complete basis set limit (CBS) are carried out with the Vienna Ab-initio Simulation Package VASP 5.3^{119,120}. We utilize the GGA functional PBE in combination with a projector-augmented plane wave basis set (PAW)^{122,123} with a huge energy cut-off of 1000 eV. This corresponds to $\sim 200\%$ of the recommended high-precision cut-off and should accurately approximate the CBS for the here investigated properties (denoted as PBE/CBS in the following).

We include the dispersion energy with the *dftd3* program in the Becke-Johnson damping scheme and a conservative distance cut-off radius of 95 Bohr. The BSSE correction is added via the gCP method as discussed above where a smaller cut-off radius of 60 Bohr is used. Compared to the computation of the HF or DFT energies and gradients, the dispersion and counterpoise contributions require practically no additional computation time (even in a small basis, the HF or DFT calculation is one to two orders of magnitudes slower).

The gCP correction will be also compared with the BB-CP scheme. In CRYSTAL, the BB-CP method is applied by supplementing the basis set of a single molecule, as cut out from the crystal structure, with the functions of an increasing number of atoms (ghost atoms) belonging to the surrounding array of molecules within a sphere of a given radius. Here, a sphere of 4.0 Å from each atom of the molecule was used. Tests with a larger radius of the sphere show a change of the lattice energies by only a few tenths of a kcal/mol.

For the geometry optimization, we use an extended version of the approximate normal coordinate rational function optimization program (ANCOPT)¹²⁴. The Cartesian coordinates of all atoms in one unit cell are transformed into approximate normal coordinates. In order to obtain reasonable internal coordinates, we translate all atoms by multiples of unit cell vectors in such a way that all molecular fragments are directly connected. We use a rational function algorithm to calculate the new coordinates from analytic atomic gradients and an interpolated Hessian matrix of second derivatives²⁰⁷.

Common unit cell optimizer rescale all atom positions inside the unit cell to get the new coordinates. For molecular crystals, this procedure is very inefficient because all

intramolecular distances change and therefore all atom gradients rise significantly. We perform a slightly different cell step. We identify all molecular fragments, calculate their center of mass, and transform all these centers according to the new cell matrix. This keeps the intramolecular distances fixed and we need less optimization steps. Full convergence is achieved if the energy change is below 10^{-6} Hartree and if the gradient thresholds for the total atomic gradient (10^{-4} Hartree/Bohr) and total cell gradient (10^{-3} Hartree/Bohr) are fulfilled.

3.4. Comparison of Counterpoise Corrections

We first compare the gCP and BB-CP correction schemes with a large basis set calculation. We start this comparison by calculating potential energy curves for two simple cubic molecular crystals, namely N_2 and NH_3 , with the hybrid functional B3LYP and small (SVP) and large (QZVP²⁰⁸) basis set. By starting from the experimental crystal structures, the lattice parameter has been varied in a range of $-10\% < a < +15\%$ by fixing the internal coordinates of the molecules at their experimental values. In 3.2, the gCP corrected potential energy curves are shown. In the case of N_2 , the curve obtained with the small basis set shows a minimum not far from the experimental lattice constant of 5.649 Å, while the BB-CP corrected one is purely repulsive. This is expected because of the very weak dispersion dominated intermolecular interactions in solid N_2 and of the well known failure of the B3LYP functional in properly describing van der Waals interactions. The artificial minimum is evidently due to the BSSE as confirmed by the results with the larger QVZP basis set which shows only a very small BSSE. The gCP correction removes most of the BSSE and closely reproduces the BB-CP corrected curve. A closer inspection shows that the gCP approach gives a better agreement with BB-CP at shorter lattice parameters, while it overestimates the correction when the cell is expanded. For ammonia (3.1 (b)) similar results are found. However, the gCP corrected lattice energies tend to be underestimated with respect to the BB-CP ones. Larger deviations are observed as the lattice constant enlarges. Notably, for both molecular crystals, the B3LYP/QZVP results are still affected by some BSSE. This indicates that the utilized QZVP basis still suffers from incompleteness, which leads to a significant BSSE.

Additionally, we investigate the urea crystal in the same fashion. In order to have a single variable to scan, we adopt the following procedure. By starting from the experimental crystal geometry (tetragonal) as measured at 12 K by neutron powder diffraction²⁰⁹ the lattice energy is computed as a function of the lattice constant a while the lattice parameter c is varied by fixing the c/a ratio at its experimental value (i.e. 0.8417). We

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

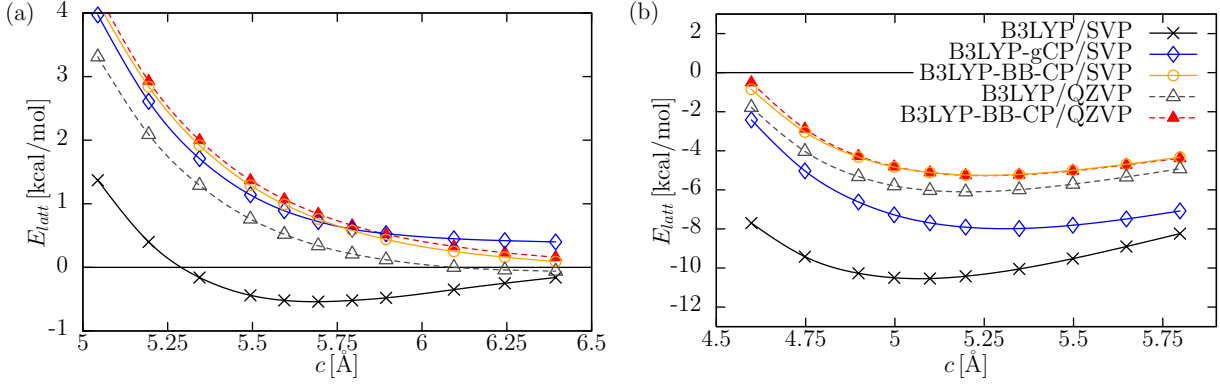


Figure 3.1.: Potential energy curves for (a) N_2 and (b) NH_3 molecular crystals. The lattice parameter is changed around the experimental value. The B3LYP/SVP curves that includes either the gCP and the BB-CP corrections are compared with the B3LYP/QZVP one (with and without BB-CP correction).

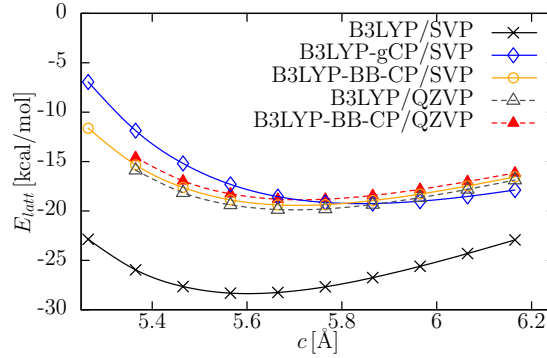


Figure 3.2.: Potential energy curves for urea. The lattice parameter is changed around the experimental value. The B3LYP/SVP curves that includes either the gCP and the BB-CP corrections are compared with the B3LYP/QZVP one (with and without BB-CP correction).

change the lattice parameter in a range between -5.4% and $+10.8\%$. Expansion and contraction of the lattice parameters is carried out in a rigid body approach. This means that during the cell deformation the internal geometry of the urea molecule is kept fixed. As can be seen in 3.2, the crystal is bound on all theoretical levels. This is expected because of the strong hydrogen bonds, which can be properly described by the (semi-)local hybrid functional B3LYP. However, the binding energy is strongly overestimated by 42 % on the B3LYP/SVP level compared to the B3LYP/QZVP result. We correct the small basis calculation with the standard BB-CP correction and with the new gCP scheme. Deviations of the binding energy compared the the B3LYP/QZVP level diminishes to 2 % and 3 % on the B3LYP-BB-CP/SVP and B3LYP-gCP/SVP levels, respectively. As shown in 3.2, the gCP curve is in good agreement with the BB-CP and the large QZVP

basis set. While the gCP corrected energy is very close to the BB-CP corrected one, the minimum is slightly shifted to larger values. This will change the equilibrium in unconstrained geometry optimizations towards too large unit cells. We present a more detailed analysis and work-around in the following section.

Overall, in the three cases, the gCP approach appears to reasonably reproduce the BB-CP correction at least regarding the lattice energies. The shape of the potential energy profile is qualitatively corrected in the right direction but we note a slight inconsistency between crystal energy and geometry. This is important for geometry optimization as will be discussed in the next section. Note that dispersion corrections have not been applied yet which is mandatory when comparisons to experimental data are made as discussed in the next section.

3.5. X23 Benchmark Set

A benchmark set for non-covalent interactions in solids consisting of 21 molecular crystals (dubbed C21) was compiled by Johnson⁴⁰ recently. Two data sets are considered:

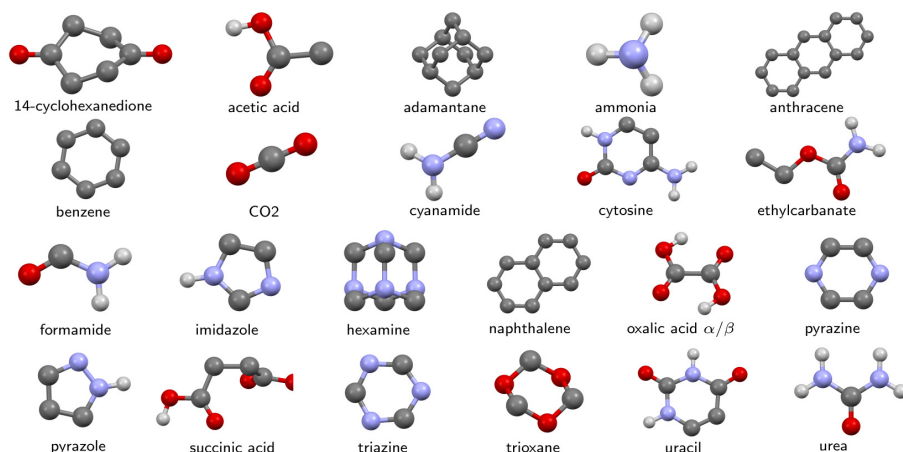


Figure 3.3.: Geometries of the 23 small organic molecules from the X23 benchmark set for non-covalent interactions in solids. H-atoms at carbons are omitted for clarity. Carbons are denoted by dark gray balls, hydrogens are light gray, oxygens are red, and nitrogens are light blue (color online).

(1) thermodynamically back-corrected experimental sublimation energies and (2) structural data from low-temperature X-ray diffraction. The thermal and zero-point effects were explicitly accounted for. Therefore, we can directly compare the electronic energy differences with the back-corrected experimental values. The error bar of experimental sublimation energies was estimated to be 1.2 kcal/mol¹¹⁸. Recently, the C21 set was extended and refined by Tkatchenko *et. al*¹¹⁷. The X23 benchmark set (16 systems are

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

Table 3.1.: Mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) of the sublimation energy for the X23 test set and for the subset X12/Hydrogen dominated by hydrogen bonds. We calculate the energy with different combinations of functionals (PBE and B3LYP), dispersion correction D3, and geometric counterpoise correction gCP and compare with thermodynamically back-corrected experimental sublimation energies. On the PBE/SVP level, we give values based on deviations to the corresponding large plane-wave basis set values in the SI. All values are in kcal/mol.

Method	X23			X12/Hydrogen	
	MAD	MD	SD	MAD	MD
PBE/CBS	11.7	-11.7	6.1	9.7	-9.7
PBE-D3/CBS	1.1	0.4	1.3	1.3	0.8
PBE-D3/CBS+ $E^{(3)}$ ^a	1.2	-0.5	1.7	1.1	0.1
PBE/SVP	5.4	-3.8	7.0	2.6	-0.1
PBE-D3/SVP	8.5	8.5	3.5	10.5	10.5
PBE-D3-gCP/SVP	2.5	-1.1	3.0	2.8	-1.4
PBE-D3-gCP/SVP+ $E^{(3)}$ ^a	2.9	-2.0	3.2	3.1	-2.2
B3LYP-D3/SVP	10.1	10.1	4.1	12.0	12.0
B3LYP-D3-gCP/SVP	2.0	0.6	2.2	1.7	-0.1
B3LYP-D3-gCP/SVP+ $E^{(3)}$ ^a	1.7	-0.3	2.2	1.8	-0.8

^a Three-body dispersion single-point energy $E^{(3)}$ on optimized structures.

presented in Ref.¹¹⁷ and seven additional systems were obtained as a private communication from the authors) includes two additional molecular crystals, namely hexamine and succinic acid. The molecular geometries of the X23 set are shown in 3.3. The details of the back-correction procedure are summarized in Ref.¹¹⁷. The mean absolute deviation (MAD) between both reference data sets is 0.55 kcal/mol. Because the X23 data seem to be more consistent we use it here as reference. If we take the standard deviation (SD) between both thermodynamic corrections as independent error source, we can add the squared errors to the total uncertainty of the reference values and obtain about 1.3 kcal/mol as statistical error. We calculate sublimation energies and crystal geometries utilizing the GGA functional PBE and the hybrid GGA functional B3LYP with a small polarized split-valence basis set SVP. The PBE/CBS values are computed with the VASP program package. In the sublimation energy calculations the isolated molecules are approximated by a large unit cell calculation with a minimum distance between molecule images of 16 Å. In order to calculate the sublimation energy, we optimize the isolated molecule and the corresponding molecular crystal. The unit cells are kept fixed at the experimental values. We summarize the deviations from reference data in Table 3.1. The individual values for all systems are given in the SI. First, the exceptionally small MAD of 1.1 kcal/mol on the PBE-D3/CBS level should be mentioned which is within the es-

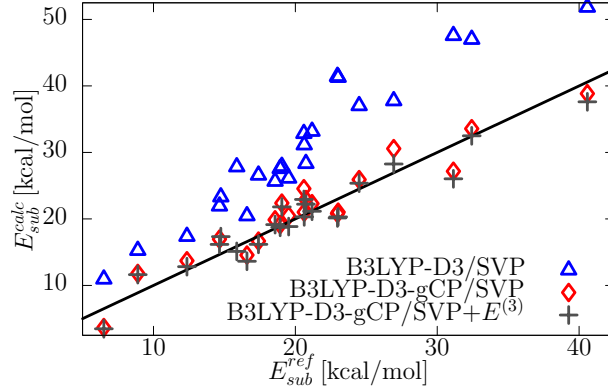


Figure 3.4.: Sublimation energy per molecule for the X23 set on the B3LYP-D3/SVP, B3LYP-D3-gCP/SVP, and B3LYP-D3-gCP/SVP+E⁽³⁾ levels. We compare with thermodynamically back-corrected experimental energies E_{sub}^{ref} .

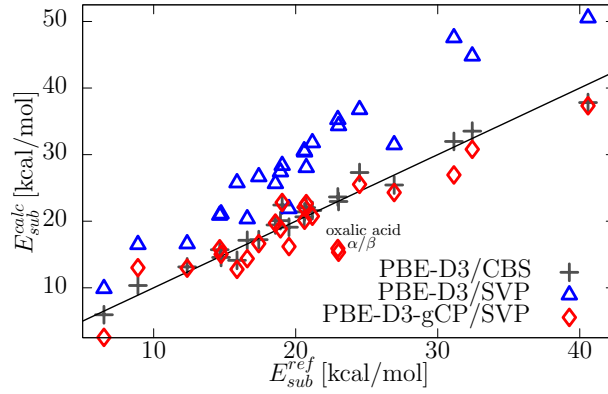


Figure 3.5.: Sublimation energy per molecule for the X23 set on the PBE-D3/SVP and PBE-D3-gCP/SVP levels. We compare with thermodynamically back-corrected experimental energies. The energies are calculated on optimized structures with experimental lattice constants.

estimated experimental error. The importance of the D3 correction is obvious from the huge error of plain PBE/CBS which very strongly underbinds most of the crystals. The D3 uncorrected functional (PBE/SVP) yields a smaller error than the corresponding CBS result which is due to the BSSE. Because of the very encouraging results of the D3 scheme in the estimated CBS limit and the physical significance of dispersion in periodic systems, we will only report and discuss D3 corrected results in the following.

Without BSSE correction, the MADs are 8.5 kcal/mol and 10.1 kcal/mol on the PBE-D3/SVP and B3LYP-D3/SVP level, respectively. The mean deviation (MD) and the MAD are identical meaning that the sublimation energy is artificially overestimated in all tested systems. Utilizing the gCP potential for BSSE correction, the MAD drops drastically to 2.5 kcal/mol and 2.0 kcal/mol, respectively, for the two methods. The BSSE

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

corrected small basis set PBE-D3 results nicely match those at the estimated CBS limit obtained with the BSSE-free plane wave basis. 3.5 shows the correlation between the PBE-D3(-gCP)/SVP energies and the experimental reference values. The remaining BSE is of the same order of magnitude than the error of the corresponding CBS calculation. Additionally we investigate the effect of three-body dispersion for all levels of theory. The correlation plot between the B3LYP-D3(-gCP)/SVP(+ $E^{(3)}$) energies and the reference values is shown in 3.4. The three-body term is repulsive for all systems and significantly improves the results for the B3LYP functional (see Table 3.1). However, the PBE-D3 errors are slightly larger with three-body term. The MAD is 1.7 kcal/mol and 1.2 kcal/mol, respectively, on the B3LYP-D3-gCP/SVP+ $E^{(3)}$ and PBE-D3/CBS+ $E^{(3)}$ level. Tentatively this can be attributed to the very different many-body behavior of the two functionals for overlapping densities as noted already some years ago²¹⁰.

We compute the deviations separately for hydrogen and non-hydrogen bonded systems to investigate the different energetic contributions. Results are also reported in Table 3.1. On the PBE-D3/CBS level, the sublimation energies of hydrogen bonded systems (X12/Hydrogen subset) are systematically too large. The systematic overbinding of PBE for such systems is well-known from molecular complexes^{88,179} and not unexpectedly transfers to molecular crystals. However, for the small SVP basis the sublimation energies of the X12/hydrogen set are systematically underestimated by 2.2 kcal/mol and by 0.8 kcal/mol on the PBE-D3-gCP/SVP+ $E^{(3)}$ and B3LYP-D3-gCP/SVP+ $E^{(3)}$ level, respectively. We explain this underbinding by the lack of a second set of polarization functions with smaller exponents. This prevents the system from a proper polarization in an electric field, which is significant in hydrogen bonds. The effect is very dominant in both oxalic acid polymorphs and in succinic acid and explains their larger error (visible in 3.4 and 3.5). For both functionals, α/β -oxalic acid have the largest deviation from the reference values. For α -oxalic acid the sublimation energy is overestimated with PBE-D3/CBS by 4.7 kcal/mol, while it is underestimated with PBE-D3-gCP/SVP+ $E^{(3)}$ and B3LYP-D3-gCP/SVP+ $E^{(3)}$ by 8.4 kcal/mol and 2.7 kcal/mol, respectively. The errors in describing hydrogen bonds with the PBE functional are significantly larger than with the hybrid functional B3LYP.

For eleven crystals of the X23 set, we perform a full geometry optimization and explore the PBE functional. For eight systems, we additionally investigate the performance of the hybrid functional B3LYP. A more detailed analysis of these structural data is presented in the SI. As most sensitive observable, we compare the unit cell volume with the experimental values in Table 3.2. Due to thermal length expansions, the calculated zero Kelvin cell volumes should be smaller than the measured ones. Thermal expansions of these systems are estimated to be approximately 3 %, while also larger thermal expansion

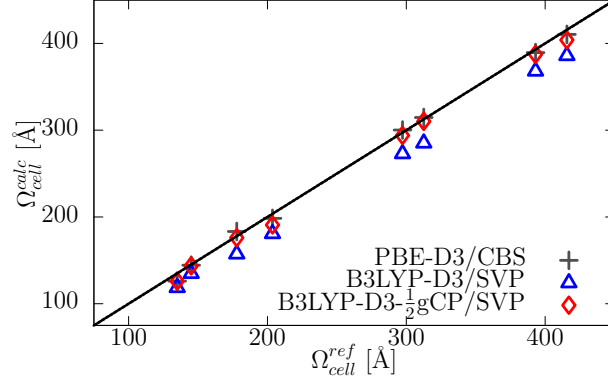


Figure 3.6.: Unit cell volumes for eight representative crystals of the X23 set on the B3LYP-D3/SVP and B3LYP-D3- $\frac{1}{2}$ gCP/SVP levels. We compare with experimental low-temperature X-ray data.

up to 8 % were reported²¹¹. The estimated PBE-D3/CBS results have a small MAD of

Table 3.2.: Mean absolute deviations (MAD) and mean deviation (MD) of the unit cell volumes. We calculate the geometries with different combinations of functionals PBE and B3LYP with SVP basis, dispersion correction D3, and geometric counterpoise correction gCP. Deviations with respect to the CBS values are given in parentheses. Absolute values are in \AA^3 .

Method	MAD	MD	Rel. MD
PBE-D3/CBS	4.2 (0.0)	-1.5 (0.0)	-0.8%
PBE-D3	17.5 (15.9)	-17.5 (-15.9)	-6.1%
PBE-D3- $\frac{1}{2}$ gCP	7.5 (5.8)	-3.3 (4.9)	1.0%
PBE-D3-gCP	23.7 (25.3)	23.7 (25.3)	7.9%
B3LYP-D3	22.1	-22.1	-9.1%
B3LYP-D3- $\frac{1}{2}$ gCP	6.6	-6.6	-2.9%

4 \AA^3 . As expected, the unit cells are systematically too small by 0.8 %. With the SVP basis set, the BSSE artificially shrinks the crystal by on average 6 % on the PBE-D3/SVP level. With the full counterpoise correction we obtain unit cell volumes by 8 % too large. The full BB-CP is known to overestimate the BSSE in large and dense systems²⁰¹. While this seems not to be the case for the sublimation energies, we notice a significant effect on the crystal geometry which, however, at present is not entirely clear. Empirically it has been found before^{201,212–216} that half of the counterpoise correction is a reasonably work around. Indeed, when the gCP correction is reduced to 50 %, a good agreement with the experimental data is obtained. The MAD drops from 17 \AA^3 to 7 \AA^3 on the PBE-D3/SVP and PBE-D3- $\frac{1}{2}$ gCP/SVP levels, respectively. The MADs with respect to the corresponding PBE-D3/CBS values are smaller. This indicates that some of the remaining errors do

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

not arise because of BSE, but rather due to shortcomings of the PBE functional as noted in the previous paragraph. The unit cell volume of the systems dominated by hydrogen bonds has an MD of 5 \AA^3 with respect to the CBS estimate, while the non-hydrogen bonded systems have a significantly smaller MD of 2 \AA^3 . Again, in the description of hydrogen bonds, B3LYP performs better than PBE. In 3.6, we show the correlation with the experimental unit cell volumes. Utilizing BSSE uncorrected B3LYP-D3/SVP, the unit cell volumes are underestimated by 9 % while with 50 % of the gCP correction very reasonable geometries, too small by only 3 %, are obtained.

3.6. Graphite: Interlayer Distance and Energy

In order to demonstrate the applicability of the gCP correction for denser systems, we investigate the interactions between graphite layers. The experimental interlayer equilibrium distance is $3.34 \pm 0.03 \text{ \AA}$ ²¹⁷. The experimental exfoliation energies have a large spread and range from -35 meV to -52 meV per atom^{218–221}. We studied a similar system previously and demonstrated the applicability of the semiempirical dispersion correction DFT-D2^{109,222}. The earlier study was done by an extrapolation of different finite size graphene layers. Significant errors were noted utilizing basis sets below the quadruple- ζ level.

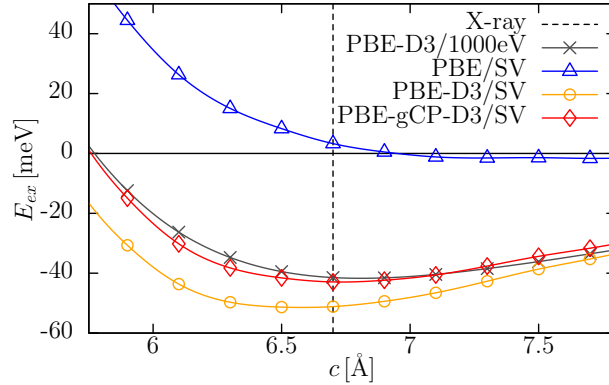


Figure 3.7.: Exfoliation energy E_{ex} of two graphite layers per atom as a function of the stacking distance c . We explore the PBE functional with SV basis and huge PAW (dubbed as CBS) basis sets. The vertical line denotes the experimental value. The cell parameters a and b are fixed to the experimental values.

Here, we describe the system periodically and investigate the effect of BSSE. For our calculations, we use the PBE functional with unpolarized split valence basis set SV. We compute the interlayer binding energy of graphite for different interlayer distances c (PES) shown in 3.7. The plain PBE functional shows no significant minimum in the

potential energy curve and no net bonding is observed. Adding the D3 correction, the minimum distance is underestimated by 0.2 Å due to BSSE. At the gCP corrected level a perfect agreement between theory and experiment is found. Calculations with huge, BSSE free plane-wave basis sets confirm the identified PBE-D3 minimum. Especially in the minimum region, the agreement between the PBE-D3-gCP/SVP and PBE-D3/CBS results is excellent. We calculate an exfoliation energy of -43.0 meV/atom on the PBE-D3-gCP/SVP level. This is reasonable in comparison to the experimental estimates, and it differs by only 3 % from three PBE-D3/CBS value.

We investigate the same system on the PBE/pob-TZVP level. The additional basis functions increase the BSSE significantly resulting in an artificial minimum at an interlayer distance of 3.2 Å. This seemingly good result arises from an accidental error cancellation. The BSSE gives an artificial attraction which simulates, near the minimum, the neglected dispersion attraction. However, the exponentially decaying BSSE can not accurately mimic the dispersion interaction, which decays as r^{-6} with the interatomic distance for each atom pair. This is demonstrated in 3.8, where we analyze the long-range behavior of the interlayer interaction on the PBE-D3-gCP/SV and PBE(-D3-gCP)/pob-TZVP levels. We fit a power law $E_{ex} \propto c^{-n}$ in a least-square sense for distances c in the interval $[10, 20]$ Å.

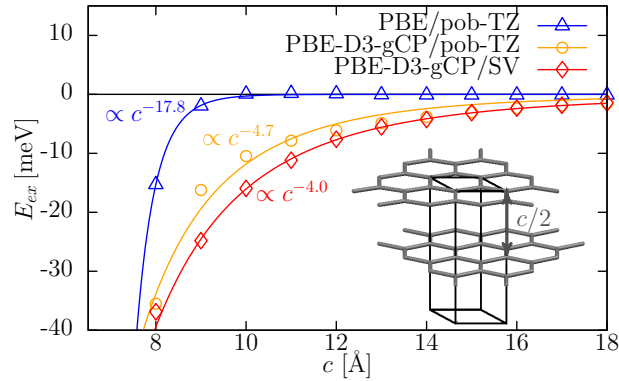


Figure 3.8.: Exfoliation energy E_{ex} of two graphite layers per atom as a function of the stacking distance c in the long-range limit. The continuous lines correspond to a power law fit with critical exponents of 18 and 4 for the PBE/pob-TZVP and PBE-D3-gCP/SV levels, respectively. The cell parameters a and b are fixed to the experimental values.

The combination of PBE/SV, D3, and gCP shows the correct asymptotic behavior of c^{-4} for insulating infinite layers. For the special case of graphite (k -point conductor) the true dependence is $c^{-3} \ln(c/c_0)$ ²²³. We find a critical exponent of 4.03 which nicely agrees with the value of 4.2 as predicted by means of quantum Monte Carlo calculations²²⁴. On the other hand, the raw PBE functional with a triple- ζ basis exhibits a more exponential

behavior (the exponent for the power-law fit is about 18) as expected.

3.7. The gCP Correction for the pob-TZVP Basis

We mentioned the problem of near linear dependencies that arises if basis functions with small exponents are included in periodic calculations of dense systems. Even with the Ahlrichs SVP basis set, we encounter some SCF convergence problems. Recently, a new Gaussian basis set, denoted as pob-TZVP, was developed¹⁹⁵ which provided stable and robust SCF convergence for a wide range of solids. However, we demonstrated above, that its BSSE can be huge. The main target of these basis sets are bulk systems or surfaces where both, dispersion interactions and BSSE artifacts are important. Therefore, we determined the gCP parameters for the pob-TZVP basis in the same way as described in the original Ref⁵⁶ using the TURBOMOLE program suite²²⁵. We calculate the e^{miss} parameters for all elements H-Br (excluding rare gases) between the restricted open shell HF energy E_{ROHF} for the target basis and a large quadruple- ζ basis def2-QZVPD according to

$$e^{\text{miss}} = E_{ROHF}^{\text{target basis}} - E_{ROHF}^{\text{large basis}}|_{F=0.06 \text{ au}}, \quad (3.9)$$

where $F = 0.06 \text{ au}$ denotes an applied weak electric field in order to populate higher angular momentum functions. These values can be used to judge the completeness of an atomic basis set as discussed in the original gCP work. In this regard, the pob-TZVP and the SVP basis sets are rather similar, with mean e^{miss} parameters of 0.24 and 0.22 Hartree, respectively. We fit the four parameters σ , α , β , and η in equation 3.8 to computed BB-CP data at the B3LYP/pob-TZVP level for the S66x8¹³³ benchmark set. The optimized parameters are

$$\begin{aligned} \sigma &= 0.1300 & \alpha &= 1.3743 \\ \beta &= 0.4792 & \eta &= 1.3962. \end{aligned} \quad (3.10)$$

The root mean square deviation (RMSD) of the fit for the pob-TZVP and SVP basis sets are 0.002 and 0.001 Hartree, respectively, i.e., the Ahlrichs basis set SVP seems to be better balanced.

For the X23 benchmark set, presented previously, the MAD of the sublimation energies is significantly reduced from 10.7 kcal/mol on the PBE-D3/pob-TZVP level to 4.7 kcal/mol on the PBE-D3-gCP/pob-TZVP level. The asymptotic behavior of the interlayer interaction in graphite is also improved. The scaling exponent of 4.72 on the PBE-D3-gCP/pob-TZVP level represents a huge improvement compared to the unphysi-

cal value of 17.83 for plain PBE/*pob*-TZVP. Although the dispersion and BSSE corrected *pob*-TZVP basis set works reasonably well, we recommend to use the Ahlrichs type basis set for calculations on molecular crystals.

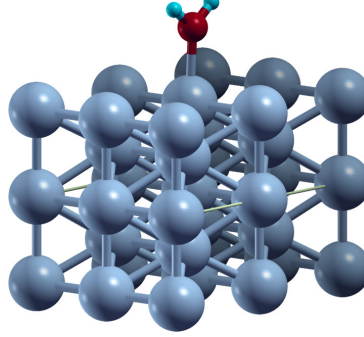


Figure 3.9.: Model system for the adsorption of H_2O on the Ni (110) surface.

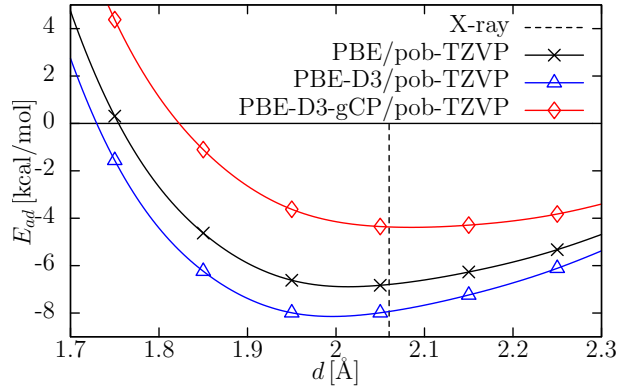


Figure 3.10.: Potential energy surface of the water adsorption on a Ni(110) surface with varying Ni- OH_2 distance d . We utilize the PBE functional with *pob*-TZVP basis set and explore the effects of D3 and *gCP* correction, respectively.

One of the most important applications of bulk-optimized basis sets in combination with correction schemes for dispersion and BSSE is the description of adsorption of molecules on surfaces. Neglecting even one of these effects can lead to significant errors. We therefore demonstrate the application of our periodic BSSE correction scheme for the adsorption of water on Ni (110). Water adsorbs oxygen-down on the Ni (110) surface at a top position. The Ni- OH_2 distance was determined by surface extended X-ray absorption fine structure (SEXAFS) spectroscopy as $2.06 \pm 0.03 \text{ \AA}$ ^{226,227}.

The Ni bulk was optimized employing the PBE functional and the *pob*-TZVP basis set. With 3.451 Å for the unmodified *pob* basis set, the deviation from the experimental lattice constant²²⁸ of 3.524 Å is only 2 %. A 2×2 supercell slab model ($a = 4.881 \text{ \AA}$, $b = 6.902 \text{ \AA}$) with 5 atomic layers was optimized, allowing full relaxation of all atoms. The structure

3. Correction for the Inter- and Intramolecular BSSE in Periodic DFT

of the water molecule was optimized in the gas phase also employing the PBE functional and the pob-TZVP basis set and set on top of a center Ni atom (see 3.10).

We varied the Ni–OH₂ bond distance from 1.7 to 2.3 Å, fixing the orientation of the water molecule. Without any correction, PBE/pob-TZVP seems to give the correct value of 2.05 Å. When applying only the D3 correction, a slight overbinding (2.00 Å) is found, while applying only the gCP correction overestimates the bond distance (2.15 Å). Correcting for dispersion as well as for the BSSE, the Ni–OH₂ bond distance (2.05 Å) is in perfect agreement with the experiment.

3.8. Conclusions

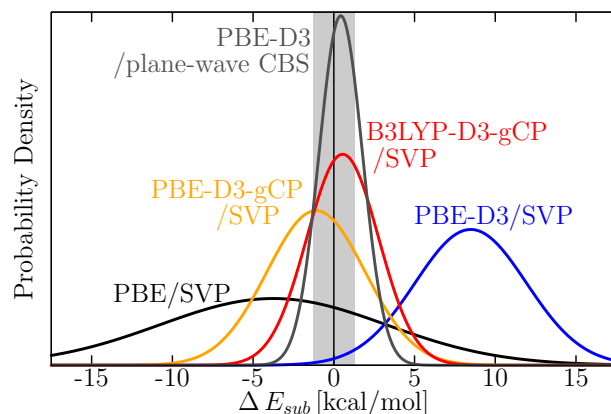


Figure 3.11.: Deviations between experimental sublimation energies of the X23 set and theoretical calculations utilizing the PBE and B3LYP functional and small SVP basis set. The CBS values are estimated with a huge plane-wave basis set with energy cutoff of 1000 eV. The statistical data are converted into normal error distributions. The gray shading denotes the experimental error interval. The quality of the theoretical methods is in the following (declining) order: PBE-D3/CBS, B3LYP-D3-gCP/SVP, PBE-D3-gCP/SVP, PBE-D3/SVP, and PBE/SVP.

We presented and evaluated the semiempirical geometrical counterpoise correction gCP for usage in periodic DFT calculations. The gCP correction is added to the total (ideally dispersion corrected) DFT energy and approximates the Boys-Bernardi counterpoise energy in a fast and differentiable way for various atom-centered Gaussian basis sets. A benchmark set for non-covalent interactions in solids (X23) is exploited to evaluate the performance of various DFT methods also with small Gaussian AO basis sets. The statistical data for the deviations of computed sublimation energies from reference data are converted to normal error distributions shown in 3.11. The plane-wave basis set PBE-D3

results are of high quality with an MAD below the estimated experimental uncertainty. From the small basis set calculations, PBE-D3-gCP/SVP and B3LYP-D3-gCP/SVP can also be recommended. The error spread at the PBE/SVP and PBE-D3/SVP levels is significant due to an incomplete account of dispersion and BSSE. This is also true for the B3LYP/SVP calculations if one of the corrections is neglected.

For the computation of the crystal structures at the SVP level, the BSSE seems to be overestimated by the counterpoise corrections. We suggest the following general strategy when small basis sets are employed. First, a geometry optimization using the atom pairwise dispersion correction D3, and 50 % of the gCP correction (dubbed DFT-D3- $\frac{1}{2}$ gCP) should be conducted. Subsequently, a single-point energy with two- and three-body dispersion energy and full counterpoise correction on the optimized structure (dubbed DFT-D3-gCP+ $E^{(3)}$) should be computed. If hydrogen bonds are present, the B3LYP functional is preferred over PBE. The gCP can be also applied to denser systems as demonstrated by the good results for the graphite stacking on the PBE-D3-gCP/SV level. The interlayer potential agrees remarkably well with BSSE-free, plane-wave results on the PBE-D3/CBS level. The agreement with the experimental interlayer equilibrium interaction energy and distance is satisfactory.

We fitted the gCP parameters for a recently developed solid-state Gaussian basis set (pob-TZVP). For the adsorption of water on a nickel (110) surface as an example, the PBE-D3-gcp/pob-TZVP level was tested. For the Ni-O distance a very good agreement with the experimental value was found.

The gCP method is implemented in a freely available FORTRAN program obtainable from the author's website and will be available in the next release of the CRYSTAL code. Similar to what has been pointed out recently for molecular thermochemistry¹⁸⁴, it is recommend as a default for dispersion corrected, small basis set DFT or HF calculations also for solids.

4. Accurate Modeling of Organic Molecular Crystals by Dispersion Corrected Density Functional Tight-Binding (DFTB)

Jan Gerit Brandenburg* and Stefan Grimme*

Keywords: Dispersion Correction, Noncovalent Interaction, Organic Crystals, Tight-Binding, Semiempirical MO

Received 15th of April 2014, Published online 6th of May 2014

Reprinted (adapted) with permission from

J. G. Brandenburg and S. Grimme, J. Phys. Chem. Lett. **2014**, 5, 1785–1789

— Copyright ©2014 American Chemical Society. DOI 10.1021/jz500755u

Own manuscript contribution

- Parametrization and interfacing of method combination
- Benchmark calculations
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract The ambitious goal of organic crystal structure prediction challenges theoretical methods regarding their accuracy and efficiency. Dispersion corrected Density Functional Theory (DFT-D) in principle is applicable, but the computational demands e.g. to compute a huge number of polymorphs is too high. Here we demonstrate that this task can be carried out by a dispersion-corrected Density Functional Tight-Binding (DFTB) method. The semiempirical Hamiltonian with the D3 correction can accurately and efficiently model both, solid and gas-phase inter- and intramolecular interactions at a speedup of two orders of magnitude compared to DFT-D. The mean absolute deviations for interaction (lattice) energies for various databases are typically 2-3 kcal/mol (10-20%), i.e., only about two times larger than for DFT-D. For zero-point phonon energies small deviations of <0.5 kcal/mol compared to DFT-D are obtained.

4.1. Results and Discussion

Accurate and efficient modeling of inter- and intramolecular interactions both in the gas and solid phase of organic molecules is mandatory for a variety of applications and represent a very active research field.^{6,7,41,69,72,229,230} The theoretical description of gas phase dimers, supramolecular host-guest complexes, and organic crystals mainly relies on the correct description of non-covalent interactions. Crucial parts of these intermolecular forces are the hydrogen bonding and van-der-Waals (vdW) interactions.²⁰ In principle, high level quantum chemical methods can seamlessly describe all the local and nonlocal interactions but are computationally too demanding for large complexes and especially for molecular crystals of larger molecules. Empirical potentials (force fields) and London dispersion-corrected Density Functional Theory (DFT-D) are the mainly used alternatives.³² Recent developments and applications of different DFT-D methods demonstrated their predictive power in both molecular complexes and organic solids.^{38,201} For example our well-established semi-classical DFT-D3 scheme successfully participated recently in the SAMPL4 blind test for prediction of host-guest association free energies.²³¹ It was further shown that semilocal density functionals in a huge plane-wave basis augmented with the D3 London dispersion correction can calculate sublimation energies with an accuracy of 1 kcal/mol.^{41,181} Similarly accurate results are obtained with the Tkachenko-Scheffler (TS) Many-Body-Dispersion correction (MBD)^{39,117} and E. Johnson’s exchange dipole model (XDM)⁴⁰. G. Beran proposed a fragment-based hybrid many-body interaction model, which is also capable to calculate lattice energies with chemical accuracy.^{7,232} These methods have already been used in the field of crystal structure prediction.⁴³

Though having a good accuracy–cost ratio, the DFT-D methods cannot be applied

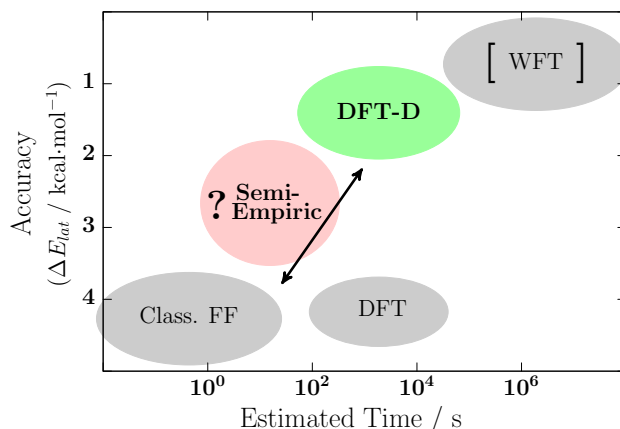


Figure 4.1.: A schematic view on the accuracy–computational cost ratio for different methods is given. The accuracy is exemplary given for the calculation of organic crystal lattice energies. Wavefunction theory methods (WFT) are expected to give the correct result in principle but can not be applied routinely. The gap between force fields and DFT-D is highlighted.

to thousands of large complexes or organic crystals in a reasonable time. Especially for pre-screening of multiple conformations in organic crystal structure prediction (or likewise in crystal structure refinement) faster methods are required. Purely empirical force fields are typically not accurate enough for a reasonable energy ranking of low-lying structures which results in a large number of cases which have to be treated by DFT.³⁷ This obvious gap between force fields and DFT could be covered by semiempirical methods as sketched in Figure C.2. In this context we recently reduced the computational cost of DFT calculations by applying small atomic orbital basis set (of mainly double- ζ quality) and correcting the arising basis set superposition error (BSSE) by a semiempirical pair potential gCP.⁵⁶ For plain Hartree-Fock in a nearly minimal basis set, we additionally corrected empirically for the basis incompleteness and compiled the HF-3c method.⁶³ Both approaches were successfully tested on organic crystals as well.^{41,57}

Here, we investigate the performance of the Density Functional Tight-Binding method DFTB3 for binding (interaction) energies of small to large molecular gas phase complexes (6 to 177 atoms) and organic molecular crystals. This method is based on a third-order expansion of the Kohn-Sham total energy with respect to charge density fluctuations. The arising matrix elements are modified by a self-consistent charge (SCC) redistribution. The modification corresponds to an on-site repulsion for short distances and to a Coulomb interaction at long distances with correct Coulomb limit. In the latest version an additional damping of the pair interactions involving hydrogen atoms is included. This significantly improves the description of hydrogen bonded systems and proton transfer.^{59,174,175,233} We abbreviate this SCC-DFTB3 method as DFTB throughout the article.

4. Modeling Organic Molecular Crystals by Dispersion Corrected DFTB

This model Hamiltonian must be additionally corrected to account for nonlocal electron correlation effects, with the London dispersion interaction as the most dominant contribution. Because the charge density of the DFTB method is (mainly due to its minimal basis) not very accurate, it is ideal to use a correction scheme which does not explicitly depend on the electronic structure and hence such an augmentation in the TB context has been proposed already some time ago²³⁴.

The atom-pairwise D3 correction solely uses the geometry information to calculate the dispersion energy

$$E_{disp} = -\frac{1}{2} \sum_{n=6,8} \sum_{i,j}^N s_n \frac{C_n^{ij}}{\|\mathbf{r}_{ij}\|^n + f(R_0^{ij})^n}, \quad (4.1)$$

where $C_{6/8}^{ij}$ are the leading order dipole-dipole and dipole-quadrupole dispersion coefficients and r_{ij} is the distance between the atom pairs i, j .²⁶ The s_6 scaling coefficient is set to unity to ensure the correct long-range behavior. The Becke-Johnson⁹⁹ rational damping function $f(R_0^{ij})$ is used to match the long- and medium-range dispersion contribution from D3 with the semilocal correlation captured by DFTB.⁸⁷ The C_6 dispersion coefficients depend geometrically on the molecular environment and are pre-calculated by time-dependent DFT and utilizing the Casimir-Polder relation.^{97,235} Because of its small numerical complexity, the D3 correction is ideally suited for a coupling with inherently fast electronic structure methods where more complicated density based schemes (e.g.^{27,30}) would lead to a huge computational overhead.

In the following, two standard benchmark sets are investigated. The molecular S66¹³³ set consists of 66 small to medium sized dimers in their equilibrium geometry. It contains purely vdW bonded, purely hydrogen bonded, and mixed systems. In the S66x8 set, the S66 dimers are considered at eight different center of mass distances. This set is the de-facto standard for testing non-covalent interactions in gas phase dimers. The reference values are basis set extrapolated CCSD(T) energies. Because similar high-level calculations are not affordable for molecular crystals, reliable reference energies can only be extracted from experiment. E. Johnson compiled a test set of molecular crystals, which was extended and refined by A. Tkatchenko.^{39,40,117} Similar to the S66 set, these X23 systems consist of purely vdW bonded, purely hydrogen bonded, and mixed systems. The experimental sublimation energies are explicitly back-corrected to electronic lattice energies. In this way, one can directly compare the electronic energies with the provided reference data in full analogy to S66. Table 4.1 shows the mean absolute deviation (MAD) and mean deviation (MD) of DFTB for both test sets. To put these value into perspective, DFT (PBE) as well as semiempirical PM7²³⁶ values are included. The individual values

Table 4.1.: Mean absolute deviations (MAD) and mean deviation (MD) of the dissociation and lattice energies for the S66x8 and the X23 test set are shown. Data are given for uncorrected as well as dispersion corrected (suffix D3) methods. All values are in kcal/mol and a positive MD denotes on average overbinding.

Method	S66x8		X23	
	MAD	MD	MAD	MD
All systems				
DFTB	2.17	−2.17	12.29	−12.29
DFTB-D3	0.79	−0.42	2.48	−0.22
PBE-D3 ^(a)	0.35	0.24	1.07	0.43
PM7 ^(b)	0.73	−0.13	—	—
vdW bonded				
DFTB	2.22	−2.22	14.31	−14.31
DFTB-D3	0.54	0.37	1.80	0.23
PBE-D3	0.27	0.01	0.86	0.06
H-bonded				
DFTB	2.36	−2.36	7.62	−7.62
DFTB-D3	1.30	−1.29	2.56	1.63
PBE-D3	0.55	0.55	1.27	1.21
mixed				
DFTB	1.88	−1.88	16.06	−16.06
DFTB-D3	0.49	−0.35	4.19	−2.64
PBE-D3	0.20	0.14	1.50	0.13

^(a) PBE in a huge almost complete basis set according to Ref^{41,179}.

^(b) PM7²³⁶ is currently not applicable to crystals.

are given in the SI. The statistical data are separated in the different bonding situations to identify the main error sources. When judging the results, one should keep in mind that the mean S66x8 dissociation energy and the mean X23 lattice energy is 4 kcal/mol (range from 0.0 to 19.5 kcal/mol) and 20 kcal/mol (range from 6.5 to 40.6 kcal/mol), respectively.

The uncorrected DFTB significantly underestimates the binding in the S66x8 dimers by more than 54 %. The MAD is approximately 2 kcal/mol for all binding motives. This (unsurprising) finding indicates that an important contribution is missing in the model Hamiltonian. The dispersion correction D3 reduces the mean error drastically with an residual MAD below 1 kcal/mol. The comparison of the hydrogen bonded with the vdW bonded systems shows that the main error originates in a partially wrong description of the delicate electrostatic and induction contributions in hydrogen bonds. The MAD for the vdW systems of about 0.5 kcal/mol is considered as very accurate and competitive to standard DFT. To cross validate this result we additionally employ the benchmark sets S22 (small gas phase dimers²³⁷), X40 (halogenated gas phase dimers²³⁸), L7 (large gas phase dimers and trimers²³⁹) with new DLPNO-CCSD(T)/ Δ CBS/CP references (unpublished

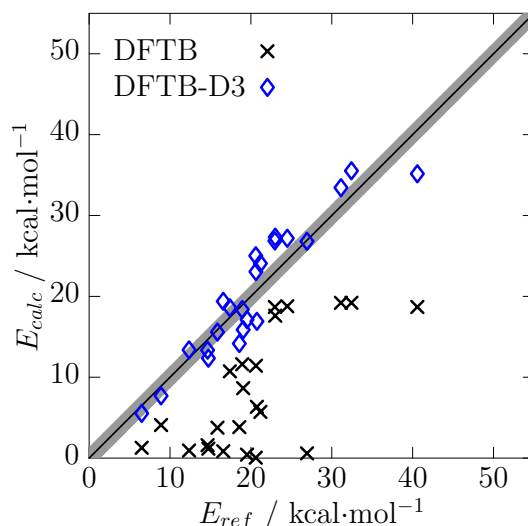


Figure 4.2.: Correlation between the calculated DFTB and DFTB-D3 lattice energies with the experimental reference values. The gray shading denotes the uncertainty of the references of approximately 1.2 kcal/mol⁴¹.

results[†]), S12L (large host-guest complexes⁹⁸). The small DFTB-D3 MADs for these sets of 0.95, 1.66, 1.74, and 5.90 kcal/mol confirm the picture drawn above. This also indicates that the S66x8 test set seems to be representative for a large number of systems.¹⁷⁹ Though the results are worse than for standard density functionals (MAD on the PBE-D3/large basis set level is 0.35 kcal/mol¹⁷⁹ for S66x8), it is an improvement over other semiempirical methods if larger systems are considered. While PM7²³⁶ has a slightly smaller MAD on the S66x8 set, it performs worse for the larger systems with MADs of 0.77, 1.69, 7.61, and 17.51 kcal/mol for the S22, X40, L7, S12L benchmark sets. Due to the large errors for larger complexes (i.e., L7 and S12L), one can not expect good PM7 results for organic crystals. The semiempirical PM6-DH2^{240,241} performs slightly better than PM7. However, the larger deviations for L7 and S12L systems persist.

A similar picture of performance is observed for the lattice energies of the molecular crystals in X23. Periodic systems are most sensitive to the correct treatment of long-range interactions. The underbinding tendency of DFTB is more pronounced for X23 than for S66 and its MAD is more than 60 % of the mean lattice energy. The errors for the X23 lattice energies drop significantly to a small MAD of 2.5 kcal/mol on the DFTB-D3 level which corresponds to only 12 % of the mean lattice energy. Note, that the dispersion corrected DFTB-D3 is used without any electronic re-parametrization, which demonstrates the robustness of the method.

[†]Note that meanwhile the references of the L7 set have been revised. However, the differences are minor and affect the statistical errors of the tested methods by less than 0.1 kcal/mol.

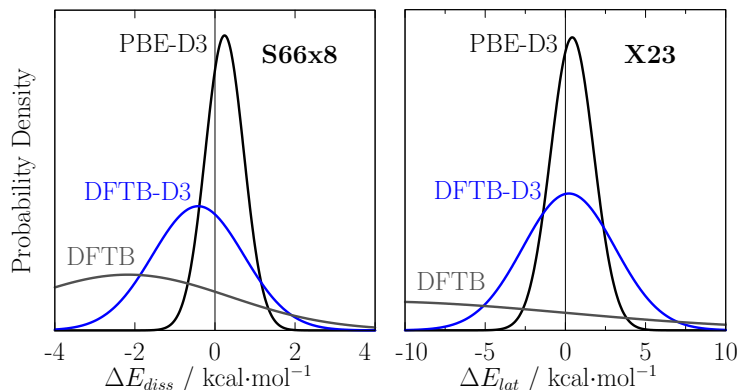


Figure 4.3.: Normal error distributions for the benchmark sets S66x8 and X23. The newly evaluated method DFTB-D3 is highlighted.

Figure 4.2 shows the individual values compared to the reference energies. The error of the back-corrected reference energies is estimated to 1.2 kcal/mol.⁴¹ The linear correlation coefficient is 0.68 and 0.94 for DFTB and DFTB-D3, respectively. Similar to the molecular case, the MAD of the purely vdW bonded systems is lower with 1.8 kcal/mol, i.e., systems containing more complicated electrostatics (hydrogen- and mixed bonding motifs) are slightly worse described. A comparison between the performance for the gas phase (S66x8) and the solid phase (X23) test sets is sketched in Figure 4.3.

Aside from pre-screening and electronic structure calculations of huge systems, DFTB-D3 can also be used as a cheaper alternative for second energy derivatives. These are needed for instance to explicitly calculate the phonon spectrum and to correct the electronic energies to enthalpies or free energies at finite temperatures. We computed the vibrational contribution to the X23 sublimation energies in the harmonic approximation (unscaled frequencies) and compare the resulting energy corrections with the recently published PBE-TS values.³⁹ The mean absolute deviation between the two data sets is only 0.5 kcal/mol, which is comparable to the results known from molecular complexes.⁹⁸ For the calculation of the sublimation energy of a given crystal structure, we propose the following procedure. First optimize the entire crystal structure including cell parameters at the general gradient approximated DFT level PBE-D3 in a large projector-augmented plane wave (PAW) basis. Then calculate a single point energy at the higher hybrid functional level PBE0-D3 at the PBE-D3 structure. Finally correct for vibrational contributions using DFTB-D3 frequencies. For example the frequency calculation of a cytosine crystal (in a supercell with 156 atoms) takes less than one hour on a standard workstation with the `dftb+` and `dftd3` codes. In the DFT single point calculation, other functionals (e.g. BLYP, TPSS, HSE06) or other London dispersion corrections (e.g. TS, MBD, XDM) could be used for comparison.

4. Modeling Organic Molecular Crystals by Dispersion Corrected DFTB

In this work, we have augmented an existing Density Functional Tight-Binding Hamiltonian (including full third-order correction, self consistent charges, and special hydrogen-pair damping) with the latest first-principles London dispersion correction D3. The DFTB-D3 method was evaluated both for mostly organic gas phase as well as solid structures with very promising results. The MAD of 2.5 kcal/mol obtained for the X23 lattice energies is exceptionally small for a semiempirical method. The analysis of different subsets in the benchmarks illustrated that the main error source for the DFTB-D3 method are the hydrogen-bonded systems. Although their description has improved in going from DFTB1 to DFTB3, this point clearly needs further improvements. The potential of the combined DFTB-D3 approach for organic crystal structure prediction and refinement was demonstrated on the X23 test set. Furthermore, the vibrational corrections on the DFTB-D3 level were compared to those obtained at the dispersion corrected DFT level and showed good mutual agreement. In summary, the future for electronic structure based organic crystal structure prediction is bright when a kind of multi-level approach is employed.

4.2. Computational Methods

We utilize the DFTB Hamiltonian with full third-order correction and self consistent charges (SCC). The SCC tolerance is 10^{-7} au. We use the most recent Slater-Koster files provided by the group of M. Elstner. The hydrogen containing pair potentials are additionally damped with an exponent of 4.2, which is the recommended value for proton transfers.^{59,175,233} The Brillouin zone is sampled with a Γ centered grid with at least 0.05 \AA^{-1} k -points, generated via the Monkhorst-Pack scheme.⁶⁸ The London dispersion correction D3 is used in the Becke-Johnson damping variant with parameters $s_8 = 0.5883$, $a_1 = 0.5719$, and $a_2 = 3.6017$. The parameters were fitted on the S66 reference energies similar to the procedure in the original publication.²⁶ A similar DFTB-D3 parametrization was already tested on the S12L set and for calculations of electron impact mass spectra.^{98,242} The X23 geometries are optimized on the DFTB-D3 level with fixed unit cell with the approximate normal coordinate rational function optimizer ANCOPT^{124,243} until the atomic forces are below 10^{-4} au. For all other benchmarks the standard single-point energy approach was applied. Phonon frequencies are calculated at the Γ -point in a supercell approach. The vibrational corrections to the lattice enthalpy are calculated in the harmonic approximation similar to the reference approach.³⁹ In the X40 test set, systems including Br or I are excluded, and the Fe-containing complex in the S12L set is also disregarded due to missing Slater-Koster files.

5. Low-Cost Quantum Chemical Methods for Noncovalent Interactions

Jan Gerit Brandenburg*, Manuel Hochheim*, Thomas Bredow*, and Stefan Grimme*

Keywords: Noncovalent Interaction, Dispersion Correction, Semiempirical MO, Organic Crystals, Supramolecular Complexes

Received 8th of Oktober 2014, Published online 25th of November 2014

Reprinted (adapted) with permission from

J. G. Brandenburg, M. Hochheim, T. Bredow, and S. Grimme, *J. Phys. Chem. Lett.* **2014**, *5*, 4275–4284 .

— Copyright ©2014 American Chemical Society. DOI 10.1021/jz5021313

Own manuscript contribution

- Implementation of the HF-3c and HF-3cv method
- Benchmark calculations (excluding MSINDO methods)
- Interpretation of benchmark results
- Writing main parts of the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstraße 4, 53115 Bonn, Germany

Abstract The efficient and reasonably accurate description of non-covalent interactions is important for various areas of chemistry, ranging from supramolecular host-guest complexes and bio-molecular applications to the challenging task of crystal structure prediction. While London dispersion inclusive density functional theory (DFT-D) can be applied, faster “low-cost” methods are required for large scale applications. In this perspective, we present the state-of-the-art of minimal basis set, semiempirical molecular orbital based methods. Various levels of approximations are discussed based either on canonical Hartree-Fock or on semilocal density functionals. The performance for intermolecular interactions is examined on various small to large molecular complexes and organic solids covering many different chemical groups and interaction types. We put the accuracy of low-cost methods into perspective by comparing with *first principle* density functional theory results. The mean unsigned deviations of binding energies from reference data are typically 10–30% which is only two times larger than those of DFT-D. In particular for neutral or moderately polar systems, many of the tested methods perform very well while at the same time, computational savings of up to two orders of magnitude can be achieved.

5.1. Introduction

The description of non-covalent interactions (NCI) has gained increasing importance in quantum chemistry in the last decade.^{20–22} Although NCIs, such as hydrogen bonding, π - π stacking, and van der Waals interactions, are typically an order of magnitude weaker than covalent bonds, they play an important role in various domains of chemistry and biology. For instance, the structure of bio macromolecules like DNA, RNA, and proteins, molecular recognition processes, and crystal packing are dominated by NCI.^{244–246} In this context, the prediction of organic crystal structures is a very active research field.^{6,10,229}

The development of new or improved methods for the description of NCIs is driven by a compromise between efficiency and routine applicability for large molecular and crystalline systems on the one hand and the necessity to keep a certain accuracy level on the other. Concerning the latter point one can further distinguish between absolute accuracy for basic properties (binding or lattice energies, equilibrium structures) and relative energies (conformational energies, affinity trends, polymorphism). In Figure 5.1, we give a general classification of the different theoretical levels. One typically distinguishes between *ab-initio* wavefunction theory, *first principles* density functional theory (DFT), *semiempirical* molecular orbital (MO) based methods, and classical methods that do not treat electrons explicitly (atomistic and coarse grained force fields and continuum mechan-

ics). Depending on the system size and the available computational resources, one has to choose an appropriate method for the desired application. While wavefunction theory is rather general, it is inapplicable to many interesting systems due to the huge computational cost, although some steps into the '100 atom regime' have been made recently⁴⁴. DFT can routinely describe and optimize larger systems of up to 10^3 atoms. When the electronic structure is still of importance and DFT is too expensive, 'low-cost' methods like various semiempirical MO schemes⁵⁸ represent a viable alternative. At this level huge systems (up to approximately $10^3 - 10^4$ atoms) can be treated routinely in a reasonable time and allowing e.g. a high-throughput screening of different polymorphs.^{247,248}

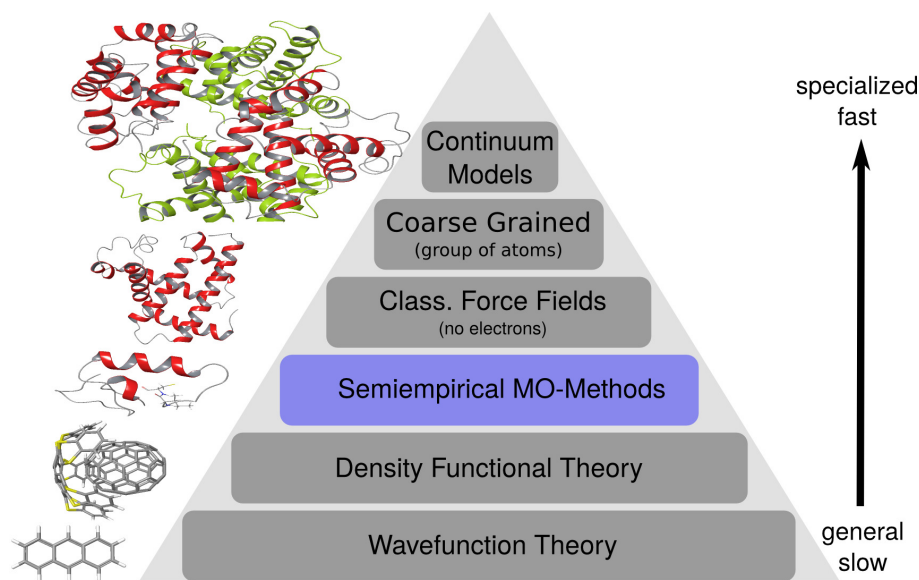


Figure 5.1.: Classification of quantum chemical methods according to their computational cost and generality.

The description of NCIs involves different levels of complexity. Electrostatic (ES), induction (IND), Pauli exchange repulsion (EXR), and London dispersion (DISP) interactions have to be treated accurately. Following the above classification, high-level, electron correlation including wavefunction theory like CCSD(T) seamlessly covers all interactions. In DFT, electron correlation is typically expressed in a local framework. Since London dispersion is a nonlocal, long-range electron correlation effect, standard DFT approximations cannot describe it properly.^{23–25} The same holds for the (mean field) Hartree-Fock approach, which lacks all Coulomb correlation effects. Significant progress has been made in recent years to solve the dispersion problem. Various approaches providing very good accuracy are available and are becoming now the standard in the field.^{26–30} For overviews on these dispersion-corrected, so-called DFT-D methods see e.g. Refs.^{31–33} The semiem-

5. Low-Cost Quantum Chemical Methods for Noncovalent Interactions

pirical MO methods are approximations to DFT or HF. As such, the above mentioned dispersion problem is also present with the same rather general solution.²³⁴ However, the semiempirical integral and basis set approximations lead to further errors in ES, IND, and EXR terms. The repulsion term is typically repaired by empirical, element specific pair potentials. The poor description of ES and IND interactions – leading for instance to a bad description of hydrogen bonds – is another well known problem in all semiempirical methods.^{249–251} In this field, geometry-dependent hydrogen bond corrections became popular.^{241,252}

Two general classes of semiempirical approaches exist in the literature. They have in common that minimal AO basis sets are used to expand the molecular orbitals which is a prerequisite for computationally efficient methods. One can start from a Kohn-Sham DFT total energy expression and expand it with respect to charge density fluctuations. The resulting density functional tight-binding (DFTB) Hamiltonian is actively developed by the groups of Frauenheim, Elstner, and Seifert.^{59–61} Recently, we coupled the third-order variant of the DFTB Hamiltonian to the D3 London dispersion correction and showed its good performance for non-covalent complexes and crystals.⁶²

Another starting point is the Hartree-Fock (HF) approach. We combined it in a small (MINIX) Gaussian basis set with the D3 dispersion model and two basis set corrections. This method (denoted as HF-3c) showed stable and good performance for molecular complexes and also for organic solids.⁶³ For larger speedups, three- and four-center two-electron integrals have to be neglected in the HF formalism as realized in the so-called neglect-of-differential-overlap (NDO) type semiempirical methods⁵⁸. This is done with slightly different strategies in the OMx, PMx, or MSINDO methods.^{253–255} To correct the arising errors, adjusted one-electron contributions and element specific pair-potentials are introduced.

In the present perspective, the presentation of the vast field of low-cost quantum chemical methods cannot be complete. Nevertheless, we try to give a comprehensive view on the available minimal basis set MO methods and highlight both, success and some failures. Fragmentation and similar methods which are applied to the system of interest are not considered here. We first give an overview on the different general strategies and then try to validate the methods on various standard benchmarks. The focus is on the computation of absolute interaction energies. Improving the performance for those more or less automatically makes relative energies better. However, the performance for NCI is not necessarily correlated to those for other properties like molecular structures or chemical reaction energies which should be kept in mind when general conclusions are made.

5.2. Covered Methods

A classification of different semiempirical methods is sketched in Figure 5.2 and will be discussed in the following subsections. The computational cost decreases in the order DFT/'CBS' > HF-3c > HF-3cv > NDDO \approx SCC-DFTB3 > DFTB2. On standard workstations and for typical applications, an HF-3c single-point calculation is a factor of about 10-15 faster than the DFT computation near the basis set limit ('CBS'). HF-3cv gains a factor of 40-50 compared to DFT/'CBS' and both NDDO and SCC-DFTB3 are even a factor of 100 to 1000 faster. The non-self-consistent DFTB2 method reduces the timings by an additional factor of 5-10.

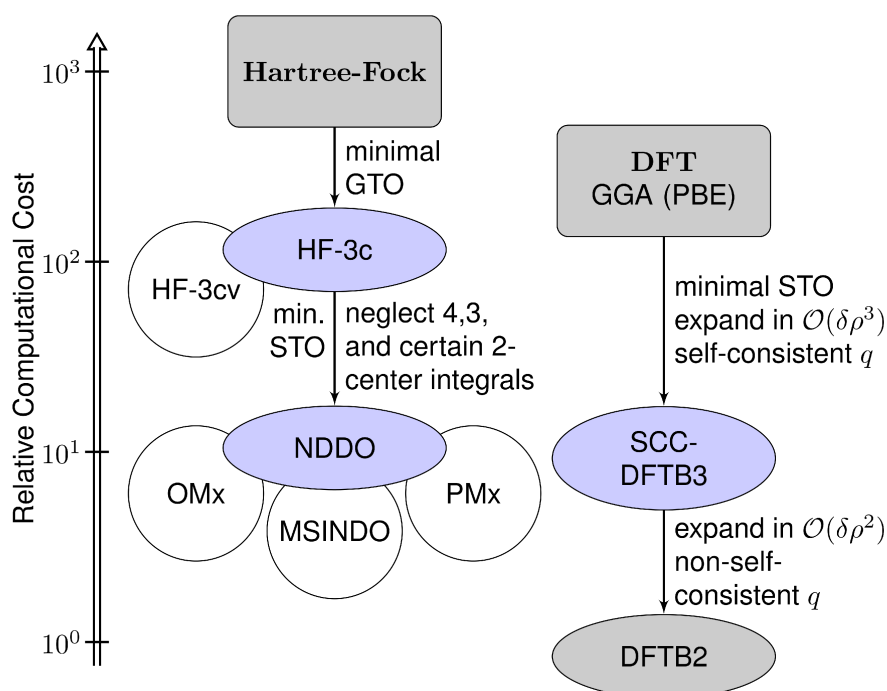


Figure 5.2.: Classification of low-cost, minimal basis set based quantum chemical methods according to their specific approximations.

5.2.1. HF-3c

HF lacks the chemically important Coulomb correlation effects, but has several advantages. It is by construction (one-electron) self-interaction error free, is analytical, and naturally includes reasonably accurate density response leading to correct ES and IND interactions. The Pauli exchange repulsion (EXR) effects, which together with the ES describe the repulsive behavior of the potential at short interatomic distances, are treated correct to first-order in HF. The clear downside is that very many two-electron repulsion

integrals have to be computed. The simplest way to speed up an HF computation is by reducing the one-particle basis set. This introduces errors, which are corrected in the HF-3c^{41,63} (and HF-3cv, see below) methods by two additional (D3 dispersion is already there) physically plausible atom pair-wise correction terms. For correcting the basis set superposition error (BSSE) the geometrical counterpoise correction (gCP) scheme^{56,57} is used and in order to account for short-ranged basis set incompleteness (SRB) effects another correction is added:

$$E_{\text{HF-3c(v)}} = E_{\text{HF}} + E_{\text{D3}} + E_{\text{gCP}} + E_{\text{SRB}}. \quad (5.1)$$

In the gCP scheme, the Boys-Bernardi counterpoise correction is mapped onto a semiempirical, repulsive pair potential, which decays exponentially with the interatomic distance. Due to the small size of the basis set, bond lengths of electronegative elements in molecular structures are systematically too long. We correct this with a short-ranged basis set correction (SRB). In total, HF-3c has nine global empirical parameters, which is comparable to common density functionals.

In order to demonstrate how far one can push this approach, we developed a 'valence variant' denoted HF-3cv. Here, the core electrons of all elements are implicitly described by the standard Stuttgart-Dresden (relativistic) effective core potentials.^{256,257} The HF equations of the valence electrons are evaluated in a minimal basis with only two Gaussian functions per shell. This leads to an additional speedup compared to HF-3c by about 2-4 in typical applications. In this context, the SRB has to be slightly modified and element dependent potentials are introduced. The method is parameterized for all elements H-Rn with eleven global and one element specific parameters. In the progress of this work, the HF-3c method (together with HF-3cv) has been implemented into the CRYSTAL14 software package and will be available in its next release (details can be found in the SI).

5.2.2. Neglect of Differential Overlap Approximations

Neglect of differential overlap (NDO) methods are typically based on HF evaluated in a minimal basis of Slater-type orbitals (STO). A significant speedup compared to e.g. HF-3c is achieved by neglecting many-center two-electron integrals. This rather broad method family was pioneered by Pople, Zerner, Dewar, and Thiel.²⁵⁸⁻²⁶⁰ Different strategies in the integral neglect, such as neglect of diatomic differential overlap (NDDO) and intermediate NDO (INDO), leads to the various variants like MNDO, OMx, PMx, and MSINDO.^{236,253,260,261}

The arising errors are corrected by adjusted one-electron contributions and element-

specific pair potentials. Because some of those NDDO methods are expanded in an orthogonalized AO basis and hence solve a special eigenvalue problem (opposed to DFTB which solves the 'right' problem including overlap), EXR effects are incompletely accounted for. Special-empirical (two-body) potentials are designed to repair this in MNDO and PMx. In OMx^{253,262} and MSINDO^{255,261} explicit orthogonalization corrections are included.

The NDO methods are empirically parameterized on theoretical or experimental reference data. The possible strategies are nicely outlined by Stewart in his latest variant (PM7²⁶³). One focus of the latter method is the inclusion of non-covalent complexes in the fit set. However, as shown here this does not automatically lead to a more consistent description of these systems. While OM2 parameters only exist for the first and second row elements, PM6/7 and MSINDO are more extensively parameterized for larger parts of the periodic table.

While in principle all methods can be used in a Γ -point approach for periodic systems, only MSINDO has an easy usable periodic implementation. The periodicity is accounted for in the real-space cyclic cluster model (CCM).²⁶⁴

5.2.3. Density Functional Tight-Binding

DFT has several advantages over HF and it is reasonable to develop approximations to a certain density functional in a minimal STO basis. An expansion of the total DFT energy in charge density fluctuations to second order $\mathcal{O}(\delta\rho^2)$ results in a standard tight-binding model. The most widespread variant is the density functional tight-binding (DFTB) Hamiltonian of Frauenheim, Seifert, and Elstner derived from the PBE density functional.¹⁷⁴ This Hamiltonian consists of contributions from the free atoms, pre-calculated kinetic energy integrals, and atom-pair potentials. These pair-potentials have to be pre-computed for all element combinations.

More recent developments take into account third-order density fluctuations. Additionally, a self-consistent charge redistribution (SCC) is introduced by coupling of the atomic charges to the Hubbard derivatives (chemical hardness).^{175,233} The long-range limit from the nuclei is a pure Coulomb potential, while at short-range it resembles an on-site repulsion. The set of coupled equations is solved self-consistently. This leads to computational costs comparable to the NDO methods. Up to this point, exclusively pre-computed data are used. However, the description of hydrogen bonds is poor and an additional damping of all hydrogen containing pair potentials with an empirical exponent is introduced. This SCC-DFTB3 (and XH-damping) will be abbreviated as DFTB3 in the following. Periodic boundary conditions are available with a standard k -space sampling and the corresponding solution of the Bloch equations.

5.2.4. London Dispersion Interaction

Being approximations to HF and semilocal DFT, respectively, all the above methods cannot describe long-range electron correlation effects. The London dispersion interaction (also referred to as the attractive part of the van der Waals interaction) is of utmost importance for extended systems. In principle, several London dispersion corrections can be applied. For instance the exchange-dipole model XDM²⁷, the Tkatchenko-Scheffler model TS²⁸, the many-body dispersion (MBD) model²⁹, and the nonlocal DFT kernel VV10³⁰ all provide reasonable accuracies. However, here we focus on intrinsically fast electronic structure methods, where the computationally least demanding D3 approach is probably the best choice. Further, the mentioned methods depend explicitly on the electron density, which is typically not very accurate in a minimal basis set representation. Because of these reasons, the D3 correction is very prominent in the low-cost field of quantum chemistry and is a reasonable choice for methods that employ a minimal basis. Except for PM7, all the tested method use the D3 scheme.

The London dispersion energy between two fragments can be calculated in the large distance limit as frequency integral over their dynamic polarizabilities.

$$E_{\text{disp}}^{\alpha\beta}(r_{\alpha\beta} \gg 1) = \frac{3}{\pi} \int_0^\infty \alpha^\alpha(i\omega) \alpha^\beta(i\omega) d\omega \times \frac{1}{r_{\alpha\beta}^6} \quad (5.2)$$

This is known as Casimir-Polder relation.⁹⁷ In the standard D3 London dispersion correction this is mapped on an atom-pairwise scheme

$$E_{\text{D3}} = -\frac{1}{2} \sum_{n=6,8} \sum_{\alpha,\beta} s_n \frac{C_n^{\alpha\beta}}{\|\mathbf{r}_{\alpha\beta}\|^n + f(R_0^{\alpha\beta})^n}, \quad (5.3)$$

where $C_{6/8}^{\alpha\beta}$ are the leading order dipole-dipole and dipole-quadrupole dispersion coefficients, respectively, and $r_{\alpha\beta}$ is the distance between the atom pair α, β .²⁶ The s_6 scaling coefficient is set to unity to ensure the correct long-range behavior. The Becke-Johnson⁹⁹ rational damping function $f(R_0^{\alpha\beta})$ is used to match the long- and medium-range dispersion contributions from D3 with the basic electronic potential.⁸⁷ The C_6 dispersion coefficients depend geometrically on the molecular environment and are pre-calculated by time-dependent DFT utilizing the Casimir-Polder relation. A non-additive Axilrod-Teller-Muto three-body dispersion (ATM-3B) is also available. The corresponding C_9 coefficients are approximated as averages from the C_6 coefficients. Similar to the two-body term, the ATM-3B contribution is evaluated in an atom-triplewise scheme.

Current research is dedicated to analyze the importance of dispersion contributions

beyond the pairwise model. Tkatchenko and co-workers significantly contributed to this field, which is reviewed in reference²⁶⁵. However, the importance of these contributions (competing with higher multipole or anisotropy effects) is still an active debate.^{266,267}. According to our experience, the three-body dispersion is in fact important for binding energies of extended systems. For instance, the ATM-3B contribution to the lattice energy of typical organic solids is 2-7%, while its effect on structure is negligible^{41,181}. Because the impact on the final energy can be significant and improves the results of nearly all tested methods, the three-body contribution is included in the single-point energy of all large systems (L7, S30L, X23). The influence for the small molecule test sets (S22, S66x8, X40, WATER14) is negligible. The ATM-3B term is the same for all tested methods and not adjusted empirically.

5.2.5. Hydrogen Bonding Correction

As mentioned in the introduction, the neglect of two-electron integrals and the applied minimal basis set in both the NDO and DFTB methods leads to a bad description of ES and IND effects. This is most important in hydrogen bonded systems where the binding is typically underestimated. This can be partially cured in DFTB by introducing an additional damping of certain pair-potentials. In the NDO-based methods, explicit geometry dependent hydrogen bonding corrections are commonly used together with D3 in an additive scheme

$$E_{NDO-D3H+} = E_{NDO} + E_{D3} + E_{H+}, \quad (5.4)$$

where E_{NDO} is the uncorrected SCF energy. Korth has developed a series of slightly different corrections, the final version is dubbed D3H+.²⁴¹ The geometric arrangement of acceptor-, donor-, and hydrogen atoms is used to calculate the hydrogen bond energy

$$E_{H+} = \frac{C_{AD}}{r_{AD}^2} f_{AHD} f_{damp}, \quad (5.5)$$

with the hydrogen bonding strength parameter C_{AD} as the arithmetic mean of the empirical acceptor and donor parameters C_A and C_D and their distance r_{AD} . The other terms are the angle-dependent and damping functions f_{AHD} and f_{damp} , respectively. Within this work, the D3H+ correction was implemented in the CCM model of MSINDO. A slightly modified version was recently published by Hobza to ensure smooth energy surfaces and a more consistent repulsive potential.²⁵² This method in combination with PM6, D3, and a halogen bond correction, dubbed PM6-D3H4X, showed most promising results.²⁶⁸ Re-

sults with older variants of both dispersion correction and hydrogen bonding corrections have been analyzed in the past. Also some additional variants of the here presented semiempirical methods, namely OM3, AM1, and SCC-DFTB2, are covered.³¹

5.3. Technical Setup

The HF-3c(v) calculations are conducted with a developer version of CRYSTAL14.¹⁷² The MSINDO-D3H+ results are obtained with the MSINDO program version 3.7. Tight SCF convergence thresholds ($\Delta E < 10^{-8}$ Eh), geometry convergence (maximum nuclear gradient $< 10^{-4}$ a.u.), and integral neglect thresholds were applied. Some organic crystals with convergence problems are disregarded in the statistics (see SI). All individual binding energies are given in the SI. To put these values into perspective, we compare with other well established semiempirical methods and the PBE density functional¹²¹ evaluated in the estimated complete basis set limit (CBS). The PM6, PM7, OM2, and DFTB energies are calculated with the MOPAC2012, MNDO2005, and DFTB+ program, respectively.^{260,269} For DFTB the 3OB parameters are used.²⁷⁰ For the H4 correction, we use the H_bonds4 program with default parameters. For the PM6-D3H+ and PM6-D3H4X method combinations, the D3 correction is used in its zero-damping variant as parametrized by Korth and Hobza, respectively.^{241,268} All other methods are treated in the rational Becke-Johnson damping scheme.

The molecular systems are treated in a single point energy approach in order to allow a direct comparison with the reference energies. For the WATER14 subset, we give the binding energies per water molecule because of the large spread of interaction energies. Because the resolution of X-ray structures is not accurate enough (especially for the position of hydrogen atoms), we optimized at each level the atomic positions of all X23 solids with fixed (experimental) unit cell parameters. The solid-state HF-3c(v) and DFTB3-D3 calculations were carried out using the standard Bloch function approach with dense k -point grids ($< 0.05 \text{ \AA}^{-1}$). For MSINDO, the CCM approach with supercells with minimum lattice vector length of 10 \AA are used. Note, that the reported interaction energies sometimes but not always include the fragment relaxation energy and we always apply the sets as originally published.

5.4. Results and Discussion

A broad range of test sets, including ones for the interaction of small, medium sized, and large molecular complexes as well as for molecular crystals were investigated to test

the performance of the methods in various situations. Prototypical systems from each analyzed benchmark set are shown in Figure 5.3. We first summarize the computational details and then introduce the benchmark systems and discuss the results.

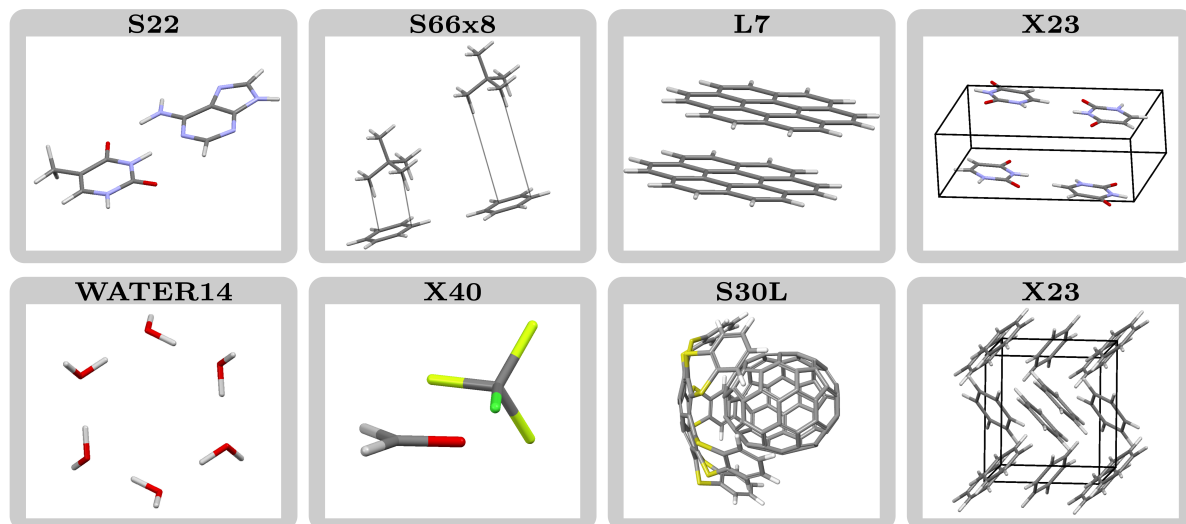


Figure 5.3.: Prototypical system from each investigated benchmark set.

5.4.1. Small Molecular Systems

The S22,²³⁷ S66x8¹³³, and the X40²³⁸ test sets compiled by the group of Hobza served as references for the interaction of small and medium sized molecular gas phase dimers. The S22 set consists of 22 complexes with typical non-covalent interactions like hydrogen bonds, dispersion interactions, and mixed electrostatic-dispersion interactions. The S66x8 test set is similar to S22, but with slightly larger molecules and less emphasis on nucleobases. In addition to the equilibrium structures seven points along the center-of-mass dissociation curve of each complex are available. The X40 test set contains small and medium sized complexes of non-covalently bound organic halides, halohydrides, and halogen molecules. Beside the typical non-covalent interactions, the X40 test set includes halogen bonding. Because water plays a dominant role in many fields, we include the uncharged systems of the WATER27 set compiled by Goddard, dubbed WATER14 in the following.¹⁸⁵ It consists of water oligomers of varying size (up to 20 molecules). The reference energies for these four sets refer to the estimated CCSD(T)/CBS level of theory.

We give the statistical data for the different methods in Table 5.1 and highlight the two best performing low-level methods for each set. In order to put the semiempirical methods into a broader perspective and to judge their overall performance, we included

5. Low-Cost Quantum Chemical Methods for Noncovalent Interactions

Table 5.1.: Performance of the low-cost quantum chemical methods for small to medium sized NCI dimers verified on the benchmark sets S22, S66x8, X40, and WATER14.^a For each set the two best performing methods are highlighted.

	S22		S66x8		X40		WATER14		all
	MAD	MD	MAD	MD	MAD	MD	MAD	MD	MAD
PBE ^b	2.6	-2.6	1.5	-1.5	1.0	-0.9	0.5	0.5	1.4
PBE-D3 ^b	0.6	0.1	0.5	0.2	0.5	0.3	1.3	1.3	0.7
HF-3c ^c	0.6	0.0	0.4	0.0	1.4	0.8	1.1	1.1	0.9
HF-3cv	1.7	1.5	0.5	0.1	2.1	1.4	2.3	2.3	1.7
OM2-D3 ^d	0.9	-0.9	0.8	-0.4	–	–	0.9	-0.9	0.9
PM6-D3H+	0.6	-0.2	0.6	-0.0	1.6	0.1	1.6	1.6	1.1
PM6-D3H4X	0.7	-0.4	0.5	0.1	1.2	-0.9	0.7	-0.7	0.8
PM7 ^b	0.8	-0.0	0.7	-0.1	1.8	-1.4	0.5	0.4	1.0
MSINDO-D3H+ ^d	1.5	-0.5	1.1	-0.2	1.4	0.4	1.5	1.5	1.4
DFTB3-D3 ^b	0.9	-0.8	0.8	-0.2	1.8	0.9	0.4	-0.4	1.0

^aMean absolute deviation MAD and mean deviation MD (> 0 denotes overbinding) given in kcal/mol. For WATER14 the interaction energy is divided by the number of water molecules in each complex.

^bFrom Ref. ⁶² ^cFrom Ref. ⁶³ ^dX40 systems neglected due to missing parameters.

a widely used density functional (PBE) with and without London dispersion correction (and evaluated with single-particle basis sets close to the variational limit) for comparison. In the SI, we additionally give PBE values evaluated in a smaller basis set of double- ζ quality. The arising BSSE is huge (in fact larger than in the HF/MINIX calculation in HF-3c) and even with decent corrections (PBE-D3-gCP) the results are on average worse than for HF-3c.

As expected, the plain PBE functional can not describe the long-range correlation effects and systematically underbinds the NCI complexes. The inclusion of D3 London dispersion correction clearly improves the results. The semiempirical methods perform also well. The mean absolute errors (MAD) on the S66x8 set are 0.5 to 1.2 kcal/mol while the MADs for the S22 set are slightly larger with 0.6 to 1.7 kcal/mol. The larger errors on the S22 set can be seen for nearly all DFT and semiempirical methods. This is due to the emphasis of the stronger bound hydrogen bonded systems in S22 compared to S66. In the following discussion one should also keep in mind that all the discussed methods are (partly) parameterized on S22 or S66 sets. For the newly implemented HF-3cv and MSINDO methods, we show the individual binding energies of the S22 test set in Figure 5.4. The errors for the X40 test set are overall larger, but still acceptable in a 1.2 to 2.1 kcal/mol range. The more involved electrostatic and induction contributions challenge all semiempirical methods. Most of the tested methods bind the X40 systems

too strongly, most clearly for the HF based schemes. In this regard, the good performance of MSINDO-D3H+ for the X40 set with smallest MAD of 1.4 kcal/mol is notable. The reduced WATER14 set is an exception, where the uncorrected PBE functional per-

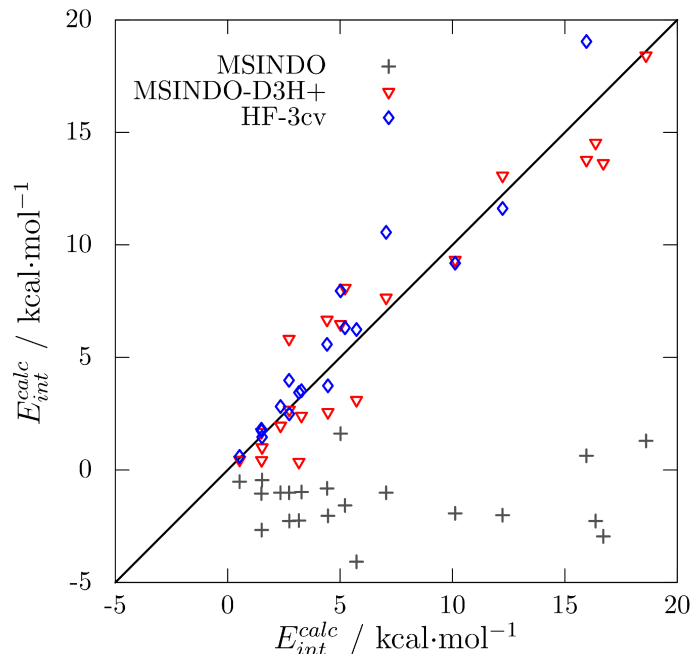


Figure 5.4.: Interaction energies for the S22 set on MSINDO, MSINDO-D3H+, and HF-3cv level in comparison with the CCSD(T)/CBS reference.

forms surprisingly good. This is due to a systematic oversetimation of the hydrogen bond energy, which compensates partly the missing London dispersion. Inclusion of the physically correct dispersion term deteriorates the result due to short-range double-counting effects. The PBE functional should be avoided if strong hydrogen bonds are present. We recommend the meta GGA functional TPSS-D3 in such cases. The performance of the semiempirical methods for the water clusters is good. The best performing methods are DFTB3-D3 and PM7 with MADs of only 0.4 and 0.5 kcal/mol, respectively, which corresponds to about 10 % deviation per water monomer. Interestingly, the methods with explicit hydrogen bonding corrections do not outperform the other schemes.

In summary, all semiempirical methods perform reasonable well on these small systems. HF-3c and PM6-D3H4X have the smallest average overall MAD well below 1 kcal/mol. This is certainly 'good news' but an insufficient condition to be applicable to much larger complexes or solids. This is the 'true' target for such low-cost methods. However, the worse description of halogen bonded systems with a more complicated electronic structure is apparent. In passing it is noted that more explicit halogen-bonding corrections

Table 5.2.: Performance of the low-cost quantum chemical methods for large complexes verified on the benchmark sets L7 and S30L.^{a,b}

	L7		S30L		all
	MAD	MD	MAD	MD	$\overline{\text{MAD}}$
PBE ^b	15.6	-15.6	31.1	-28.6	23.3
PBE-D3 ^b	2.0	-0.8	6.6	-1.0	4.3
HF-3c	1.2	0.4	6.5	1.5	3.9
HF-3cv	2.4	2.3	7.4	-0.4	4.9
OM2-D3 ^c	2.9	-1.8	6.7	1.1	4.8
PM6-D3H+	3.5	-0.2	6.2	1.2	4.8
PM6-D3H4X	2.9	0.0	7.4	1.9	5.2
PM7 ^b	6.5	6.5	16.0	15.7	11.3
MSINDO-D3H+	5.7	3.9	9.5	5.3	7.6
DFTB3-D3 ^b	1.7	0.2	6.2	0.9	3.9

^aMean absolute deviation MAD and mean deviation MD (> 0 denotes overbinding) given in kcal/mol.

^bThe D3 correction includes the ATM-3B energy. ^cFour systems in S30L neglected due to missing parameters.

exist which, however, have not been considered here.^{19,238,271,272}

5.4.2. Large Molecular Complexes and Organic Solids

In this section, first the L7²³⁹ set by Hobza and the S30L set, which both contain large molecular complexes, are examined. The L7 set is composed of seven large non-covalently bound molecular complexes of organic molecules. We use new DLPNO-CCSD(T)/ Δ CBS/CP reference values (to be published elsewhere), which are more consistent than the original values.²³⁹ The S30L test set consists of 30 supramolecular host-guest complexes. It is an extension of the S12L set which was the first benchmark set of supramolecular complexes in the literature.⁹⁸ Together, these systems are certainly challenging for any molecular electronic structure method. In passing it is noted that most empirical-force fields fail badly on these tests either due to missing parametrization or totally inconsistent results as discussed recently¹⁹. The reference interaction energies vary between 20 and 150 kcal/mol, multiple hydrogen bonds are present, and nine systems have a positive charge between one and four. The reference binding energies are obtained from experimental data which are explicitly back-corrected for thermodynamic and solvent effects to ensure direct comparability to electronic energies, for details see Ref.⁹⁸ This procedure has been independently verified by comparison to high-level DMC results^{52,†}

[†]Note that meanwhile the references of the L7 and S30L sets used in the following have been revised. However, the differences are minor and affect the statistical errors of the tested methods by less than 0.1 and 0.2 kcal/mol, respectively.

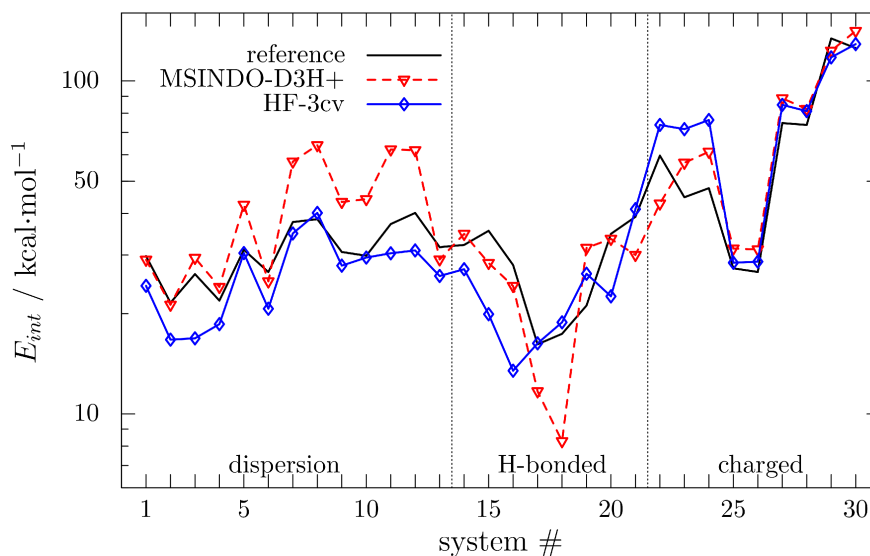


Figure 5.5.: Interaction energies for the S30L set on the MSINDO-D3H+ and HF-3cv level in comparison with the back-corrected experimental interaction energies.

The statistical data for these complexes are shown in Table 5.2. The analysis of the deviations reveals similar trends as discussed in the previous paragraph. The overall interaction energies are larger and therefore also the absolute deviations are larger. Again, almost no binding is obtained by the uncorrected PBE functional. With inclusion of the D3 dispersion correction, a significant improvement is apparent. More sophisticated

Table 5.3.: Performance of the low-cost quantum chemical methods for the X23 set of organic crystals.^{a,b} For each set the two best performing methods are highlighted.

	MAD	MD
PBE	11.7	-11.7
PBE-D3	1.1	0.4
HF-3c	2.2	-1.3
HF-3cv	3.1	-1.4
MSINDO-D3H+ ^c	6.6(3.9)	1.9(-1.6)
DFTB3-D3	2.4	-0.9

^aMean absolute deviation MAD and mean deviation MD (> 0 denotes overbinding) given in kcal/mol.

^bThe D3 correction includes the ATM-3B energy in the final energy. ^cin parentheses: four outliers neglected.

functionals like the meta-GGA TPSS-D3 or the hybrid PBE0-D3 (not discussed here) improve these results even further.

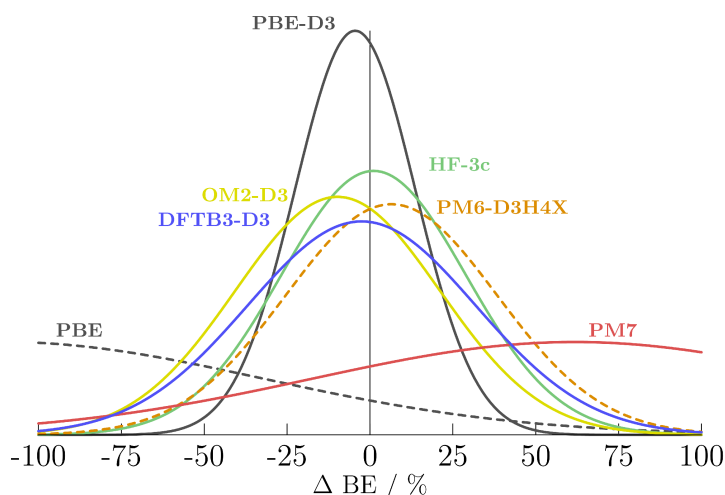


Figure 5.6.: Relative deviations of the computed binding energies ΔBE from the reference for the joint L7 and S30L set. The statistical results are converted into normal error distributions. The deviations for the L7 set were upscaled by a factor of four in order to ensure a balanced weighting.

Most of the low-cost methods work very reasonably well with errors sometimes even below those of PBE-D3. Only PM7 is far off and can not be recommended. Figure 5.5 shows individual interaction energies for the S30L set. At the MSINDO level, some problems with large π systems are apparent (e.g. the bucky 'catcher' 9-12). Except for PM7, the other methods perform better. DFTB3-D3 has the smallest MAD of 6.0 kcal/mol closely followed by PM6-D3H+. For L7, HF-3c and DFTB3-D3 have the smallest errors of 1.2 and 1.7 kcal/mol, respectively. These deviations are below 12% of the mean binding energy, which is even better than the dispersion corrected PBE functional. Somewhat expected, the charged systems are the main error source in all the minimal basis set approaches. These errors can be reduced by applying counter ions as discussed in Ref.²³¹. The relative errors in the binding energies of the joint L7 and S30L sets are shown in Figure 5.6. We converted the statistical data into normal error distributions. While PM7 and plain PBE are far off, the other methods perform rather similar. Especially for PBE-D3, HF-3c, and DFTB3-D3 no systematic shift is found.

As most critical test for long-range interactions the X23³⁹ set of organic solids was considered. The original test set of 21 molecular crystals was set up by Johnson⁴⁰ and later refined by Tkatchenko et al.³⁹ The crystal structures are taken from low-temperature X-ray diffraction and the reference energies are based on experimental sublimation energies, which are back-corrected to get electronic lattice energies. The results are given in Table 5.3 and the corresponding normal error distributions are shown in Figure 5.7.

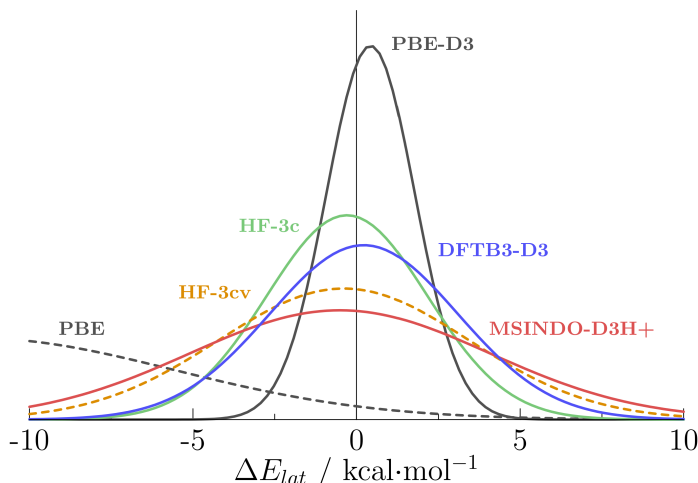


Figure 5.7.: Absolute deviation of the computed lattice energies ΔE_{lat} from the references for the X23 set. The statistical data are converted into normal error distributions.

The remarkable performance of HF-3c and DFTB3-D3 has been noted before^{41,62}. The MADs of 2.2 and 2.4 kcal/mol, respectively, are close to the DFT-D results at significantly lower costs. However, HF-3c is the least empirical and most expensive of the here investigated low-cost methods. Further basis set reduction in the HF-3cv scheme gives slightly worse results without systematic over- or underbinding. The MAD of 3.1 kcal/mol is still reasonable to allow explorative or screening investigations. While HF-3cv has a larger MAD compared to DFTB3-D3 and is even computationally more expansive, its empirical adjustment is much less involved and therefore the method is readily available for the whole periodic table. Though the D3H-corrections lead to a clear improvement for MSINDO-D3H+ of about 70%, its overall error is rather large. The high MAD of 6.6 kcal/mol is particularly caused by four systems (benzene, CO₂, oxalic acid α , β). If these systems are neglected, the MAD diminishes to 3.9 kcal/mol. The bad description of the very strong hydrogen bonded systems and the pure π -systems could already be seen for large molecular complexes of similar structure in S30L.

5.4.3. Summary and Future Directions

We considered several well established low-cost quantum chemical methods ranging from minimal basis set Hartree-Fock to semiempirical MO based schemes for the calculation of general non-covalent interactions. We have developed a computationally more efficient valence variant of the HF-3c method and augmented the MSINDO method with corrections for London dispersion and hydrogen bonding interactions. Both schemes were imple-

	S22	S66x8	X40	WATER14	L7	S30L	X23
PBE-D3/'CBS'	0.6	0.4	0.5	1.3	2.0	6.6	1.1
HF-3c	0.6	0.4	1.4	1.1	1.2	6.5	2.2
HF-3cv	1.7	0.5	2.1	2.3	2.4	7.4	3.1
DFTB3-D3	1.0	0.8	1.8	0.4	1.7	6.2	2.4
OM2-D3	0.9	0.8	—	0.9	2.9	6.7	—
PM6-D3H4X	0.7	0.5	1.2	0.7	2.9	6.7	—
PM6-D3H+	0.6	0.6	1.5	1.6	2.6	6.2	—
MSINDO-D3H+	1.5	1.1	1.4	1.5	5.7	9.5	6.6
PM7	0.8	0.7	1.8	0.5	6.5	16	—

MAD / kcal·mol⁻¹

relative error: ■ < 12.5% < ■ < 25% < ■ < 37.5% < ■

Figure 5.8.: Mean absolute deviations (MAD) in kcal/mol for various low-cost methods on all investigated benchmark sets. The color code highlights the error relative to the mean binding energy of the corresponding test set.

mented in generally available program packages. The proposed methods were tested in comparison to other semiempirical approaches (DFTB3-D3, OM2-D3, PM6-D3H+, PM6-D3H4X, and PM7) regarding their performance for non-covalent interactions in various systems.

A summary of our investigations is shown in Figure 5.8 as mean absolute deviations in kcal/mol. In order to easily judge the results, we give the errors relative to the mean binding energy of the corresponding test set in the color code. To put these values into perspective, we also added values for a commonly used dispersion corrected density functional (PBE-D3) in the estimated basis set limit ('CBS').

For the small organic systems (S22, S66x8) most methods perform well with MAD values close or below 1 kcal/mol. Note, however, that such values here correspond to typical relative errors of about 20-30 % while an MAD of <0.5 kcal/mol is required to bring the typical error down to a more acceptable 5-10% range. The halogenated systems are problematic for all minimal basis set methods, which probably can only be reduced by special correction schemes.

The larger molecular complexes (L7, S30L) are more challenging, but most of the methods give very reasonable results in particular for neutral complexes. HF-3c and DFTB3-D3 perform excellent with relative errors below about 10%. Though PM7 was presented as successor of PM6, we cannot recommend it for large non-covalently bound systems with relative errors close to 50%.

The strategy of separately adjusting the short-range electronic potentials and the non-covalent corrections (as conducted in HF-3c and DFTB3-D3) seems to be superior to an over-fitting prone determination of all parameters on a huge reference set.

The X23 set of organic solids is most sensitive to the correct treatment of long-range interactions. All technically feasible and converging computations give good results with relative errors for sublimation energies ranging from 10 to 20%. The mean absolute deviation increases in the order PBE-D3, HF-3c, DFTB3-D3, MSINDO-D3H+.

Regarding the overall picture, we loose a factor of about two in accuracy for the semiempirical methods compared to *first principle* DFT results. At the same time, a computational speed-up of two orders of magnitude is achieved. While we have analyzed the interaction energies in huge data sets, future benchmarks should also examine also structures (molecular complexes as well as unit cells) provided by the low-cost methods in more detail. We have recently shown that the rotational constants of medium-sized molecules are well described by dispersion-corrected DFT and dispersion-corrected semiempirical methods.¹²⁷ PM6 and PM7 geometries were also investigated for small NCI dimers²⁷³. This should be extended to larger and periodic systems because geometry optimization represents a main area of application of low-cost methods.

In conclusion, the here presented methods are ideally suited for large scale applications as either pre-screening tools or in a multi-level approach by combining *ab-initio* electronic energies with semiempirically derived thermostistical corrections. In any case, due to the increased accuracy and meanwhile relative broad range of applicability, the future for low-cost quantum chemical methods for investigating non-covalent interactions in all their aspects seems bright.

Part IV.

Application to Organic and Organometallic Crystals

In part II and III electronic structure methods for noncovalent interactions in general and for organic solids in particular have been presented. Part IV of this thesis focuses on explicit applications of these methods.

It has been shown that DFT-D2 cannot properly describe the relative gas phase conformations and the crystal structure of ethyl acetate. In Chapter 6, we demonstrate that with modern dispersion corrections like the D3 model, the results can be significantly improved. In particular, the relative energies of the gas phase conformers are very close to the reference and the crystal geometry agrees reasonably well with the experimental structure. Interestingly, ethyl acetate does not crystallize in the *gauche* form (nearly isoenergetic with the most stable gas phase conformer), but in the *trans* conformer. The stability of the *trans* form is probably increased due to the increased number of hydrogen bonds compared to the other conformers. This study shows that DFT in combination with modern dispersion corrections can accurately describe both the gas and the solid phase of organic molecules.

The subsequent study on tribenzotriquinacene (TBTQ) and its methyl derivative (Me-TBTQ) has been conducted in close collaboration with experimental groups (Chapter 7). A refined X-ray geometry of TBTQ has been measured with subtle changes in the space group symmetry compared to earlier measurements. Specifically, a slight rotation between the stacked molecules could be detected which is not present in Me-TBTQ. DFT-D3 calculations could confirm this structure and assign the tilting to the three-dimensional packing dominated by intermolecular London dispersion interactions.

A most pronounced CPE was found for a zirconium complex presented in Chapter 8. The X-ray geometry of two complexes (differing only in substitution of a tert-butyl group with a trimethylsilyl group) was found to deviate significantly. Due to the coordination to

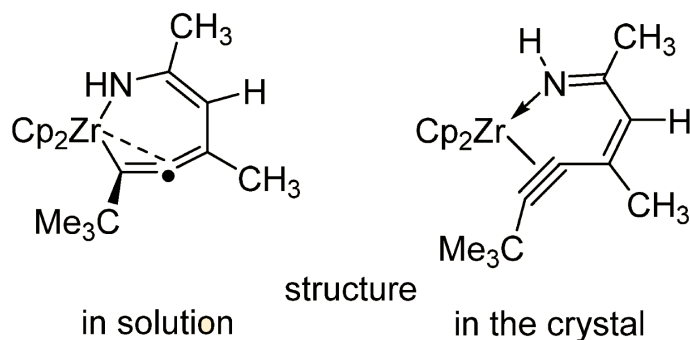


Figure 5.9.: Lewis structure of the seven-membered ring zirconium compound with differing bond character in the gas and the solid state.

the metal atom, a formally forbidden (η^2 -allenyl)/enamido-Zr to (η^2 -alkyne)/ κN -imine-Zr

5. Application to Organic and Organometallic Crystals

complex isomerization (from the gas to the solid state) is feasible (compare with structure in Figure 5.9). Our theoretical analysis revealed that the gas phase structure of both compounds is identical (allene-type), but with a slightly different barrier of the isomerization to the structure with triple-bond character (η^2 alkyne). In the crystal, intermolecular London dispersion forces stabilize this structure which is unfavored in the gas phase. Because of the metal center, this formally forbidden isomerization is feasible and consistently explains the measured structures. The hypothesis of a CPE induced bond isomerization was then confirmed by solid and liquid phase NMR measurements. The correct assignment of the NMR chemical shifts (and anisotropies) was guided by corresponding DFT calculations.

6. A Dispersion-Corrected Density Functional Theory Case Study on Ethyl Acetate: Conformers, Dimer, and Molecular Crystal

Jan Gerit Brandenburg* and Stefan Grimme*

Keywords: Density Functional Theory, Dispersion Corrections, Molecular Conformation, and Crystal Structure Prediction

Received 16th of August 2013, Published online 24th of September 2013

Reprinted (adapted) with permission from

J. G. Brandenburg and S. Grimme, *Theor. Chem. Acc.* **2013**, *132*, 1399.

— Copyright ©2013, Springer-Verlag Berlin Heidelberg. DOI 10.1007/s00214-013-1399-8

Own manuscript contribution

- Computations of gas phase conformers, dimer and the solid state
- Interpretation of results
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract We present a dispersion corrected Density Functional Theory case study on recently reported apparently difficult systems (Boese *et al.*, ChemPhysChem **14**, 799 (2013)). The relative stability of the *trans*, *gauche*, and *cis* conformers of ethyl acetate, the dissociation energy of the (*trans-trans*) dimer, and the structure and electronic lattice energy of the corresponding molecular crystal is calculated. We utilize the generalized gradient approximation density functionals PBE and BLYP, the hybrid functional B3LYP, and the double-hybrid functional B2PLYP. It is shown that all semilocal density functionals must be corrected for missing long-range electron correlation, a.k.a. London dispersion interaction. The performance of the *ab initio* dispersion correction DFT-D3 is excellent and significantly improves the results compared to the uncorrected functionals and compared to the older more empirical DFT-D2 correction. The three-body dispersion contribution to the lattice energy is 7%, while its impact on the crystal geometry and the conformer energies is negligible. A nonlocal correction approach termed DFT-NL is also tested and shows good performance comparable to the DFT-D3 results. Overall, it is shown that dispersion corrected Density Functional Theory can accurately describe the properties of ethyl acetate in various states ranging from single molecule conformers to the infinite periodic molecular crystal.

6.1. Introduction

Because of the very good computational cost over accuracy ratio, Density Functional Theory (DFT) has been the 'work horse' for many applications in chemistry and physics and is still an active research field of general interest^{73–76}. In many covalently bound (periodic and non-periodic) systems, DFT provides a very good compromise between accuracy and computational cost. However, common generalized gradient approximated (GGA) functionals are not capable of describing long-range electron correlation^{23–25,77}. This London dispersion term can be empirically defined as the attractive part of the van der Waals-type interaction between atoms and molecules that are not directly bonded to each other. For the physically correct description of molecular crystals, dispersion interactions are crucial^{20,78}. Several well-established methods for the inclusion of dispersion interactions into DFT are available. For an overview and reviews of the different approaches, see e.g. Refs.^{32,33,79,80,82}. The computation of reliable sublimation energies, geometries and relative energies of molecular crystals and their polymorphic forms is of utmost importance. Especially for crystal structure predictions accurate and fast methods are essential^{37,41,72}. The theoretical evaluation of mainly organic molecular crystals and their polymorphs is an active research field^{7,30,40,45,80,83,84}. While the total dispersion energy is smaller in

molecular systems, its relative impact can be large and even change the energetic order of different conformers. It was shown for a huge benchmark set that density functionals augmented with a modern dispersion correction perform exceptionally well for general main group thermochemistry, kinetics, and noncovalent interactions⁸⁸. For a general method it is important to describe the properties of molecules in various states equally well.

In a recent article by Boese, Kirchner, Echeverria and Boese²⁷⁴, the crystal structure of ethyl acetate was presented, high-level wavefunction correlation methods (W1 and W2) were used to calculate the molecular stability, and some dispersion corrected density functionals were tested against the experimental crystal geometry and the reference energies, respectively. Although the authors pointed out the importance of an accurate theoretical description of weak molecular interactions for these properties, their dispersion-corrected DFT results were only of medium quality and the reader could get the impression as if the use of dispersion corrections to DFT is not justified in particular for the description of the crystal structure. Because these topics are currently intensively investigated and are of general interest, we re-investigated these apparently difficult systems with state-of-the-art dispersion corrected DFT methods. As will be shown in the following, one problem in this study was the use of an out-dated (seven year old), rather empirical atom-pairwise dispersion correction termed DFT-D2²⁷⁵ in the literature. Although one can argue that DFT-D2 is still widely used (also because of its implementation in popular program packages like Gaussian or VASP), we think that methodological progress in quantum chemistry should be reflected in current applications. A more refined *ab initio* version dubbed DFT-D3^{26,87} was already developed in 2010 and is freely available as a computer code²⁷⁶ for general use in molecular as well as periodic calculations. It incorporates non-empirical, atom pairwise-specific, chemical environment-dependent dispersion coefficients, a physically more sound damping function according to Becke and Johnson¹¹¹, and a non-additive Axilrod-Teller-Muto three-body dispersion term (E_{ABC} , see^{26,102}).

Conclusions based on DFT-D2 results can be misleading and as shown herein, the reported errors are mostly corrected at the DFT-D3 level. We will discuss the following sources of error in DFT-D type methods: (1) wrong dispersion coefficients, (2) neglect of many-body dispersion terms, (3) density-independent dispersion coefficients, (4) fitting of DFT-D2 on a benchmark set of too small molecules, and (5) shortcomings of the underlying density functional. Because the implications of these statements are far-reaching and as already mentioned, the DFT based computation of molecular crystals and their polymorphs is an active field of research, the present re-investigation seems mandatory. In the following we show that the reported deviations in Ref.²⁷⁴ can mainly be attributed to the use of the older dispersion correction D2. By comparing D2, D3, and D3+ E_{ABC} results we

will confirm that point (1) is the major issue but that no significant error arises from point (2). Regarding point (3) we have recently shown that density based dispersion corrections, e.g. the nonlocal VV10 functional³⁰ (dubbed DFT-NL) give no significant improvement compared to DFT-D3 for small as well as large organic complexes^{96,98}, which is confirmed for the present system. The DFT-NL results are slightly better than the DFT-D3 results, but of comparable quality. This is expected because the VV10 dispersion coefficients are not better than those from D3 and hence the general opinion that density-based schemes are uniformly better than atom pairwise ones has (at least for insulators as considered here) no basis.

The long-range behavior of the London dispersion energy (the more short-ranged part is also called medium-range correlation energy²⁷⁷), which is important in large systems and crystals, is determined by the C_6 coefficients. The damping function merely affects the short/medium range part by interpolating between the correction term and the density functional. The D3 fitting procedure which employs a benchmark set of small to medium sized non-covalently bound molecules does not affect the long-range part and is accurate also for large complexes as recently demonstrated for supramolecular systems⁹⁸. Furthermore, a scaling of the long-range part as proposed by different authors^{83,84} changes the physically correct asymptotic behavior. Therefore, the point (4) is not relevant for D3 but is somewhat problematic for D2 and other methods which include inaccurate or scaled C_6 coefficients. Beside point (1), the density functionals have the a strong impact on the results with the expected trends between the GGA, the hybrid, and the double-hybrid functionals.

6.2. Computational Details

The molecular calculations are carried out with the TURBOMOLE²²⁵ program package. We utilize large Gaussian single-particle basis sets, namely Ahlrichs' doubly polarized triple-zeta basis set def2-TZVPP¹³⁵ for the GGA functionals and for the hybrid functionals, while we use the larger def2-QZVP basis¹³⁴ for the double hybrid functional. The basis sets are large enough that errors arising from the basis set superposition error can be excluded. The RI-approximation is used for the Coulomb integrals, which speeds up the computation significantly without any significant loss of accuracy²⁷⁸. The numerical quadrature grid m4 (m5 for the double hybrid) is employed for the integration of the exchange-correlation contribution.

The periodic calculations are performed with the VASP suite of programs^{279,280} utilizing a projector-augmented plane wave basis set^{122,123} with an energy cutoff of 1000 eV, a Γ -

centered k -mesh of $(2 \times 2 \times 2)$ points, and a tight geometry optimization convergence criterion of 0.005 eV/Å. The starting geometry is the experimental crystal structure²⁷⁴.

We utilize the GGA functionals PBE^{121,180} and BLYP^{166,167}, the hybrid functional B3LYP^{138,139}, and in addition the double-hybrid B2PLYP²⁸¹, which includes nonlocal virtual orbital dependent correlation effects. To account for London dispersion interactions, we investigate different schemes. We use our semiempirical dispersion correction DFT-D2, our standard *ab initio* dispersion correction DFT-D3^{26,87} in the Becke-Johnson damping scheme with and without the non-additive Axilrod-Teller-Muto three-body dispersion term (E_{ABC}), and the nonlocal VV10 density functional kernel of Vydrov and van Voorhis³⁰ (based on earlier work by Langreth and Lundquist⁹¹). Note that the E_{ABC} term used is based on a common, functional independent short-range damping and is numerically identical for the different density functionals employed in this study.

6.3. Results and Discussion

Table 6.1.: Energy difference of ethyl acetate conformers relative to the *trans* form $\Delta E_{gauche/cis}$ and dissociation energy D_e of the centro-symmetric *trans-trans* dimer. All values in kJ/mol.

Method	ΔE_{gauche}	ΔE_{cis}	D_e
W1 (reference value) ^[a]	0.33	34.30	18.36
PBE	0.81	31.20	13.58
PBE-D2	-0.38	28.27	19.84
PBE-D3	-0.09	30.20	18.83
PBE-NL ^[b]	-0.13	30.93	19.68
BLYP	1.40	32.11	8.86
BLYP-D2	-0.43	27.60	20.16
BLYP-D3	-0.30	30.25	19.02
BLYP-NL ^[b]	-0.14	31.65	18.49
B3LYP	1.63	33.71	11.59
B3LYP-D2	-0.13	29.63	20.16
B3LYP-D3	0.11	32.14	19.02
B3LYP-D3+ E_{ABC} ^[b]	0.14	32.15	19.07
B3LYP-NL ^[b]	0.30	33.34	19.64
B2PLYP ^[c]	1.10	34.20	13.81
B2PLYP-D2 ^[c]	0.22	32.06	18.51
B2PLYP-D3(zero) ^[c]	0.49	33.91	18.82
B2PLYP-D3 ^[c]	0.34	33.46	17.83

^a Values from Ref. ²⁷⁴. ^b Singlepoint energies on the DFT-D3 structures. ^c Singlepoint energies on the B3LYP-D3 structures.

First, we compute the relative stability of the *cis*, *trans*, and *gauche*-conformers of ethyl acetate. Such conformational changes involving small energy differences are also of importance for the formation of crystal polymorphs and hence a detailed investigation is worthwhile. We perform unconstrained geometry optimization on the different theoretical

6. DFT-D Study on Ethyl Acetate: Conformers, Dimer, and Molecular Crystal

levels. The final relative energies on these optimized structures are presented in Table 6.1. All uncorrected functionals predict the correct energetic order ($trans < gauche < cis$), but with quite large deviations from the reference, especially for the small $trans$ - $gauche$ energy difference ΔE_{gauche} with a reference value of 0.33 kJ/mol. On the DFT-D2 level the energy difference ΔE_{gauche} is reduced, but the $gauche$ conformer is wrongly predicted as the most stable conformer. Optimization on the DFT-D3 level correctly increases the stability of the $trans$ conformer. The higher accuracy of the D3 compared to the D2 correction is apparent. This is independent of the explicit form of the damping function. The deviation of the D3 scheme with the original zero-damping to the here used Becke-Johnson-damping is small as shown for the B2PLYP functional. However, the Becke-Johnson damping has been shown to be more robust and is recommended as the default choice. Furthermore, this property seems to be sensitive to the choice of the density functional as demonstrated by the wrong PBE-D3 and BLYP-D3 (semilocal GGA) energetic order, the correct order of the more sophisticated hybrid B3LYP-D3, and the improvement by the generally more accurate double-hybrid B2PLYP-D3 which yields a value of 0.34 kJ/mol very close to the reference value. For the $trans$ - cis difference the D3 correction improves the B3LYP-D2 result by 2.8 kJ/mol although we note an fortuitously good result of dispersion-devoid B3LYP. The wrong energy of the $gauche$ conformer on the same level strongly indicates that one cannot rely on this error cancellation. The best value is again computed with B2PLYP-D3.

Single-point energy calculations with DFT-NL on the DFT-D3 structures have only a small effect, but (with the exception of PBE-NL) uniformly improve the results compared to plain functionals. Especially the results of B3LYP-NL are exceptionally good with deviations below 0.03 kJ/mol and 0.96 kJ/mol (corresponding to relative errors of 9 % and 3 %) for the $gauche$ - $trans$ and the cis - $trans$ stability, respectively. Because the damping function of the three-body term in the D3 method is independent of the used functional, we only add the term for the B3LYP functional. Inclusion of the three-body dispersion energy corrects in the right direction, but the effect is only minor and negligible compared to the inherent errors of the density functional. The results for the conformational energies are systematically improved when a) the dispersion correction is changed from D2 to D3 (and slightly improved by NL) and b) the underlying density functional is improved which is in agreement with the conclusions from extensive thermochemical benchmark studies⁸⁸.

The dimerisation energy of a $trans$ - $trans$ -dimer is substantially underestimated by all plain functionals. However, it is overestimated by the D2 corrected functionals, e.g. by 1.7 kJ/mol (9 %) on the B3LYP-D2 level (8 % for PBE-D2). The error is reduced significantly to 0.6 kJ/mol (4 %) for B3LYP-D3 (3% for PBE-D3). Deviations from accurate reference interaction energies on the order of 5 % are considered as small and have been

observed for various types of non-covalent interaction between organic molecules¹⁷⁹. The D3 dispersion corrected double-hybrid performs best with errors smaller than 3 %. Again, DFT-D3 significantly improves the already good DFT-D2 results (as noted before²⁶) and DFT-NL performs even slightly better (with the exception of PBE-NL). We investigate the effect of the three-body dispersion term on the dimerisation by calculating single-point energies on the optimized structures. Its influence is co-operative, but very small (about 0.2 % of the dissociation energy).

In a previous publication²⁷⁴, the crystal structure for ethyl acetate was calculated on the B3LYP and PBE level, with and without D2 dispersion correction, and compared to experimental X-ray structure. It crystallizes in the *trans* conformation. According to these data, B3LYP-D2 and PBE-D2 significantly underestimate the cell volume by 17 % and 10 %, respectively. We optimize the unit cell and all internal coordinates without any symmetry constraints on the PBE-D3 level and obtain a small relative unit cell error of 2 % (Table 2). The unconstrained optimization is important to get the true minimum of the corresponding theoretical level. Keeping in mind that the thermal volume expansions can be on the order of 3 %, our theoretical zero Kelvin result seems to be excellent. A

Table 6.2.: Cell parameters (in Å and °), cell volume Ω (in Å³), and sublimation energy E_{sub} (in kJ/mol) compared to experimental X-ray data (at 175 K).

Method	a	b	c	α	β	γ	Ω	E_{sub}
Exptl. ^[b]	7.17	7.66	10.9	70.7	83.6	65.8	515	36 – 63 ^[a]
PBE	7.75	8.11	11.29	70.6	86.4	70.3	629	18
PBE-D2	6.82	7.51	10.76	70.9	84.3	64.0	467	61
PBE-D3	7.10	7.65	10.86	70.3	84.5	65.8	506	57
PBE-D3 + E_{ABC}	—	—	10.86	—	—	—	—	53

^a Examples for sublimation energies of organic crystals with molecules of similar size (1,4-dioxane and trioxane). ^b Values from Ref.²⁷⁴.

comparison of the raw PBE and PBE-D2 equilibrium values can be misleading, which is demonstrated by the calculated potential energy curves plotted in Figure 6.1, where the dependence of the cohesive energy on the vertical stacking distance c is shown. The cell parameters a and b are fixed to the experimental values and for each point we perform a geometry optimization with a fixed unit cell. Because of a small attractive electrostatic component to the interaction (and potentially weak hydrogen bonding), even the uncorrected PBE curve exhibits a very shallow minimum. However, the potential is unrealistic and the optimal cell parameter c is too large by approximately 8 % compared to the experimental value. The binding is so weak that even at the low measurement temperature of 175 K no net bonding is expected for raw PBE. The shallow potential also explains the necessity of very tight structure optimization criteria, which should be employed in general full molecular crystal optimizations.

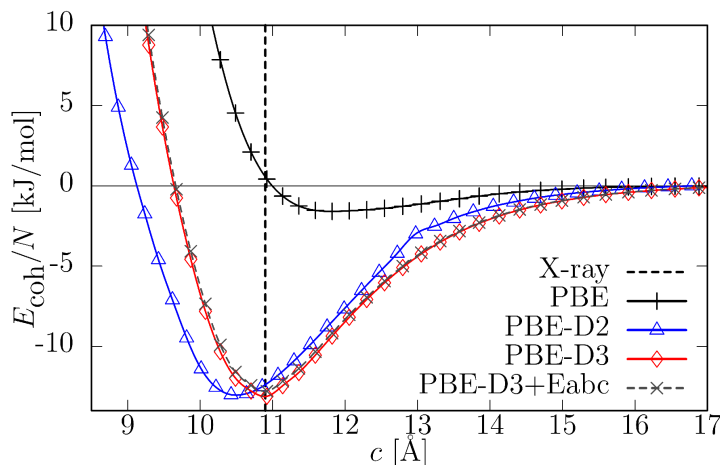


Figure 6.1.: Performance of PBE-D: Dependence of the cohesive crystal energy E_{coh} per molecule on the vertical cell parameter c . The vertical line denotes the experimental value (at $T = 175$ K). The results refer to the PBE functional with different dispersion corrections. The three-body dispersion energies E_{ABC} are calculated on the PBE-D3 structures. The cell parameters a and b are fixed to the experimental values. The asymptotic energy limit $c \rightarrow \infty$ corresponds to the internal interaction of one ethyl acetate layer.

The potential on the PBE-D2 level shows a much more pronounced minimum for the cell parameter c which is approximately 5% too small but is clearly more reasonable than the one obtained by plain PBE. For PBE-D3 we find an almost perfect agreement between theory and X-ray measurement. The calculated minimum at $T = 0$ K is by only 0.5% too short, which is rather accurate when taking the expected typical thermal length expansion of up to 1 – 2% into account, i.e., the theoretical data always should be slightly too small. The calculated sublimation energy of 57 kJ/mol (i.e., neglecting all phonon contributions) is typical for organic compounds of this size, e.g. 36 kJ/mol for 1,4-dioxane²⁸² or 63 kJ/mol for trioxane⁴⁰. Additionally, we investigated the impact of the three-body dispersion energy. We calculate single-point three-body energies E_{ABC} on the PBE-D3 structures. As can be seen in Figure 1, only the minimum of the potential is slightly raised. The computed sublimation energy is lowered by approximately 4 kJ/mol while the equilibrium distance is not affected. According to our experience⁶⁶ the three-body dispersion term is often negligible for the crystal geometry but should be included for energetic properties. This clarifies that the neglect of many body-dispersion terms are unproblematic in the D2 and D3 schemes (point (2) mentioned in the introduction and raised in Ref.²⁷⁴). This effect is significantly smaller than found in recent studies by another group^{29,125} employing a general many-body dispersion scheme.

Due to the significantly larger computational costs, only B3LYP single-point energies

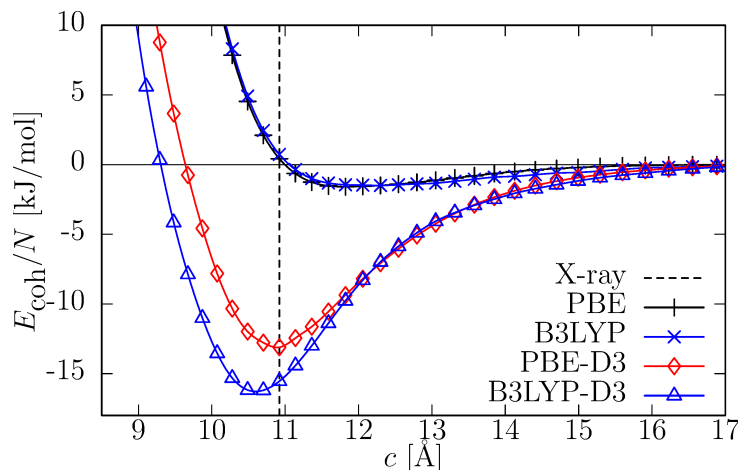


Figure 6.2.: Performance of B3LYP-D: Dependence of the cohesive crystal energy E_{coh} per molecule on the vertical cell parameter c . The vertical line denotes the experimental value (at $T = 175$ K). The plot compares the PBE functional with the hybrid functional B3LYP. The B3LYP energies are calculated on the corresponding PBE optimized structures. The cell parameters a and b are fixed to the experimental values. The asymptotic energy limit $c \rightarrow \infty$ corresponds to the internal interaction of one ethyl acetate layer.

could be calculated on the optimized PBE-D3 crystal structures. A comparison between B3LYP and PBE is shown in Figure 6.2. Somewhat surprisingly, the PBE and B3LYP potentials are nearly identical (B3LYP is usually more repulsive than PBE for non-covalent interactions). This leads to a slightly more attractive B3LYP-D3 potential compared to that of PBE-D3. Compared to the experimental value, the B3LYP-D3 cell parameter c is too small by approximately 2.5 % which we consider not as a very good result but a still reasonably small deviation. However, the value is larger (better) than with PBE-D2 and therefore, we expect the fully relaxed cell volume also to be larger than the PBE-D2 volume of 467 \AA^3 , i.e., better than with plain B3LYP. We emphasize again that one should consider in such cases the whole potential energy curve to evaluate a theoretical method and that computed (zero temperature, equilibrium) cell volumina are more reasonable when on the lower side of the experimental value. Hence, also for B3LYP the dispersion corrected functional performs better (and in fact much better when the unreliable raw B3LYP curve is considered) than the uncorrected one.

6.4. Summary

It is confirmed that the semiempirical D2 dispersion correction is overly attractive for ethyl acetate as noted previously for other organic crystals^{83,84}. The D2 dispersion coef-

6. DFT-D Study on Ethyl Acetate: Conformers, Dimer, and Molecular Crystal

ficients are not very accurate, do not depend on the chemical environment and are scaled depending on the density functional which changes the asymptotic behavior incorrectly (this has been extensively discussed already, see Refs.^{26,32}). These issues are corrected in DFT-D3 which yields good results for all investigated molecular as well as crystal properties in full agreement with results from recent tests by other groups^{33,104}. All D3 results are much better than those from uncorrected density functionals, which substantiates our opinion that dispersion corrections should be the default in all DFT treatments. We have furthermore shown that the three-body dispersion term has negligible effects on the crystal structure as well as the dimerisation energy while the sublimation energy is lowered significantly by approximately 7%. For recent work on the importance of many-body dispersion effects in dense systems see Refs.^{29,200,201} Moreover, we would like to point out that molecular crystals can involve very flexible degrees of freedom with concomitantly shallow potential energy surfaces so that structure optimizations may require very stringent numerical thresholds and cut-offs in order to obtain converged conclusive results.

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State: A Combined X-Ray and Theoretical Study

Jan Gerit Brandenburg*, Stefan Grimme*, Peter G. Jones[†], Georgios Markopoulos, Henning Hopf[‡], Michał K. Cyrański[§], and Dietmar Kuck[¶]

Keywords: Bowl-Shaped Molecules, Tribenzotriquinacenes, C_{3v} -Symmetrical Molecules, Crystal Structure Prediction, Unidirectional Stacking, Density Functional Theory, Dispersion Correction

Received 26th of February 2013, Published online 13th of June 2013

Reprinted (adapted) with permission from

J. G. Brandenburg, S. Grimme, P. G. Jones, G. Markopoulos, H. Hopf, M. Cyrański, and D. Kuck *Chem. Eur. J.* **2013**, *19*, 9930–9938.

Copyright ©2013 WILEY-VCH, Weinheim. DOI 10.1002/chem.201300761

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

[†]Institut für Anorganische und Analytische Chemie, Technische Universität Braunschweig, Hagenring 30, 38106 Braunschweig, Germany

[‡]Institut für Organische Chemie, Technische Universität Braunschweig, Hagenring 30, 38106 Braunschweig, Germany

[§]Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

[¶]Department of Chemistry, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

7. *Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State*

Own manuscript contribution

- Computations of crystal structures
- Interpretation of computational results
- Writing the theory part of the manuscript

Abstract A combined X-ray diffraction and theoretical study of the solid-state molecular and crystal structure of tribenzotriquinacene (TBTQ, **2**) and its centro-methyl derivative **3** is presented. The molecular structure of the parent hydrocarbon displays C_{3v} symmetry and bears the three indane wings in mutually orthogonal orientations, similar to the methyl derivative **3** studied earlier. Also similarly to the latter, the bowl-shaped molecules of **2** form infinite molecular stacks with perfectly axial, face-to-back (convex-concave) packing and with parallel and unidirectional orientation of the stacks. The experimentally determined intra-stack molecular distance is 4.75 Å for **2** and 5.95 Å for **3**. Whereas the molecules of **2** show a slight alternating rotation (ca. $\pm 6^\circ$) about the common axis of each stack, **3** shows perfect translation symmetry within the stacks. We use dispersion corrected density functional theory to compute the crystal structures of the tribenzotriquinacenes **2** and **3**. The London dispersion correction is crucial for an accurate description of the crystallization of both analyzed systems and the calculated results agree excellently with the experimental measurements. We also obtain reasonable sublimation energies for both compounds. Additionally, geometries and dimerization energies of oligomeric stacks of **2** were computed and show smooth convergence to the properties of the infinite polymeric stack.

7.1. Introduction

Triquinacene (**1**), first synthesized by Woodward and his associates in 1964²⁸³ has always remained prominent as one of the most inspiring non-natural polycyclic hydrocarbons.²⁸⁴ Its C_{3v} -symmetrical, convex-concave structure, bearing three formally isolated C=C double bonds and three bis-allylic bridgehead positions, has been of particular interest because of the early, yet unsuccessful idea that **1** could give rise to dodecahedrane, the largest Platonic hydrocarbon, in a single dimerization step.^{285,286} Triquinacene has also played a pivotal role in studies on neutral homoaromaticity²⁸⁷ and as a starting material for the fully unsaturated analogue, acepentalene.^{288,289} The fact that the triquinacene framework contains three mutually orthogonal axes (Figure 7.1), each of which crosses the central carbon atom and bisects one of the double bonds, has not been explicitly commented on, although an X-ray crystal structure of compound **1** was published in 1976.²⁹⁰ This particular feature only became obvious when 12d-methyltribenzotriquinacene (**3**)^{291,292} and, some years later, the parent tribenzotriquinacene (“TBTQ”, **2**)^{292–296} were studied by using X-ray diffraction.^{297,298} Hydrocarbons **2** and **3** constitute logical extensions of the structural motif of compound **1** and, because of their three aromatic units, they are much more readily accessible, stable, and chemically versatile than the non-aromatic parent

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

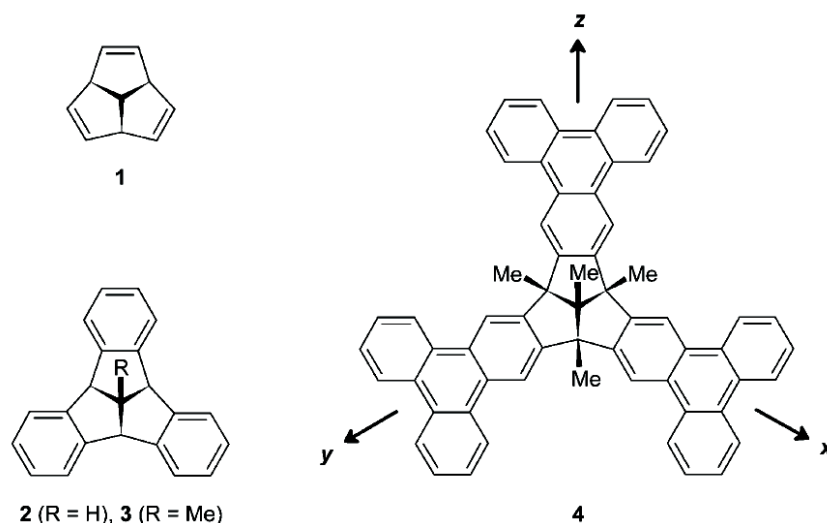


Figure 7.1.: Triquinacene, **1**, the tribenzotriquinacenes **2** and **3** and an extended TBQT derivative, **4**.

hydrocarbon.^{297,299–301} Facile multiple functionalization of compound **3** has been demonstrated, which enables various extensions of the TBQT framework at its indane wings, such as the construction of the polycondensed analogue tris(triphenyleno)triquinacene (**4**),³⁰² among others.^{303–307} Moreover, the single-wing extension and optical resolution of TBQT derivatives have recently been demonstrated.^{308–311} Besides the in-space orthogonal orientation of its three indane wings, X-ray crystal structure analysis of the methyl derivative **3** revealed stacking of the convex “face” of one TBQT bowl with the concave “back” of the next one, thus generating a rhombohedral lattice of parallel, unidirectional, and effectively infinite molecular stacks.²⁹⁷ The same stacking motif was later found for the parent hydrocarbon (**2**)²⁹⁸ and a more-extended TBQT-based tris-cyclophane.³¹² Convex–concave, unidirectional stacking is known for the C_{3v} -symmetrical sumanene^{313,314} and several corannulene derivatives, such as circumtrindene (C_{3v}),³¹⁵ hemibuckminsterfullerene (C_3),³¹⁶ and cyclopenta[bc]corannulene (C_s),³¹⁷ but the TBQT bowl of compound **3** was found to be unique because of its perfectly eclipsed stacking of the C_{3v} -symmetrical bowls.²⁹⁷

Herein, we report a detailed combined study of the solidstate structures of homologues **2** and **3**, based on a recent, more refined X-ray crystal-structure analysis and on theoretical calculations by using dispersion-corrected density functional theory. We show that the unique crystal structures of these TBQT hydrocarbons can be calculated in detail by including our latest, atom-pairwise dispersion correction with Becke and Johnson damping, DFT-D3(BJ).^{26,87} The dispersion correction is crucial for an accurate description of the crystal structures of both derivatives **2** and **3**. As will be shown, the agreement of

the calculated results with the experimental data is excellent. In the case of the parent compound, we show that geometries and dimerization energies of oligomeric stacks of compound **2** can be extrapolated to the infinite single-stack polymer.

7.2. Results and Discussion

7.2.1. Experimental Studies on the Solid-State Structures of Tribenzotriquinacenes **2** and **3**

Previous Observations

In our previous studies on tribenzotriquinacenes, it became obvious that the tendency of the parent hydrocarbon (**2**) to crystallize was much more pronounced than that of its 12d-methyl derivative (**3**). Whereas the higher homologue (**3**) forms large colorless

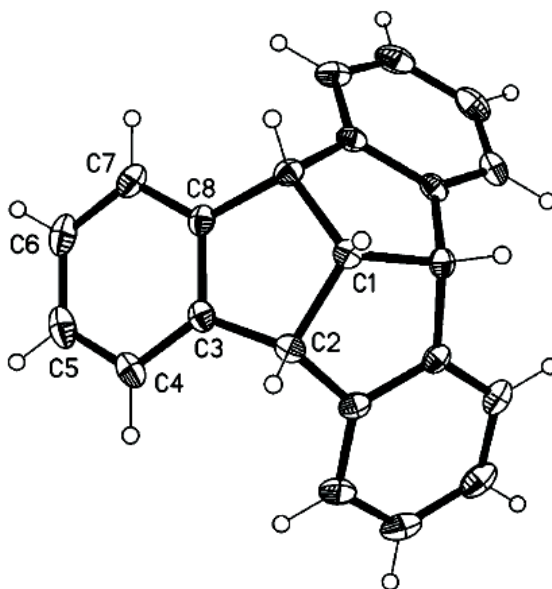


Figure 7.2.: Single-crystal molecular structure of tribenzotriquinacene (**2**); thermal ellipsoids are set at 50% probability. Only the asymmetric unit is numbered because of the crystallographic symmetry; for atom numbering, see Table 7.1.

crystals with m.p. $243\text{--}245\text{ }^{\circ}\text{C}$,^{291,292} the lower homologue (**2**) starts to crystallize from hot solvents, such as toluene or xylenes, thus forming extremely thin and long needles that melt at about $390\text{ }^{\circ}\text{C}$,^{293–296} that is, almost $150\text{ }^{\circ}\text{C}$ higher than compound **3**. With the X-ray crystal structure of compound **3** in hand,²⁹⁷ it was natural to assume that the solid-state aggregation of the lower homologue (**2**) would display a similar (but much tighter) crystal packing because of the absence of the apical methyl group. Our early

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

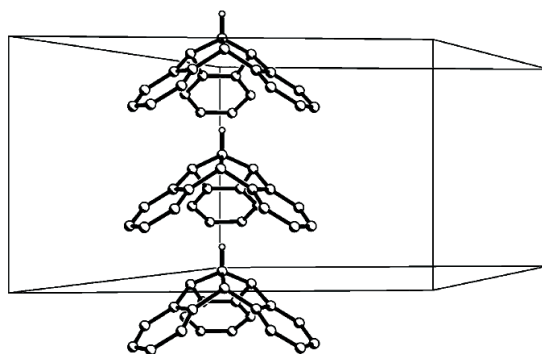


Figure 7.3.: Molecular stacking in tribenzotriquinacene (**2**); there is a slight alternating rotation of adjacent molecules within each stack (by 4 ± 68 , Figure 7)

X-ray structure analysis confirmed this evaluation.²⁹⁸ The distance between the central carbon atoms (C12d) of two adjacent molecules, all of which were located on the stacking axis, was found to decrease by about 1.25\AA without any change in the other geometrical parameters within each stack. The experimental study reported herein largely confirms this picture but reveals additional interesting details.

X-ray Structural Analyses of Tribenzotriquinacenes **2** and **3**:

Our first X-ray structure determination of the parent tribenzotriquinacene (**2**) was performed on thin (cross-section: $0.05 \times 0.05\text{ mm}^2$) and weakly diffracting needles that were obtained by the slow evaporation from toluene. That investigation did not detect the weak reflections with odd values of l and, thus, gave a halved c axis and the space group $R3m$. In our new determination, we used a larger and more strongly diffracting crystal that was obtained from the same solvent and we found the true space group to be $R3c$. The solid-state molecular structure of compound **2**, as obtained from our recent measurements, is shown in Figure 7.4. The central atom, C12d (X-ray numbering: C1), lies along the threefold axis, so that the formal molecular symmetry is exactly threefold, but the effective non-crystallographic symmetry is C_{3v} , with a root mean square deviation (RMSD) of only 0.001\AA . As found earlier for the 12d-methyl homologue (**3**, see below),²⁹⁷ the molecules are stacked parallel to the polar c axis (Figure 7.3). However, in the case of the parent hydrocarbon (**2**), the operator is no longer a simple c axis translation, but a c glide plane. Most notably, neighboring molecules in the stack are alternately rotated relative to each other by about $+6^\circ$ and -6° . As a result, only every other molecule is related by translational symmetry. The average stacking distance is $c/2 = 4.748\text{\AA}$.

Because the previous X-ray structure of 12d-methyltribenzotriquinacene (**3**), as determined at room temperature, had also been reported to crystallize in the space group

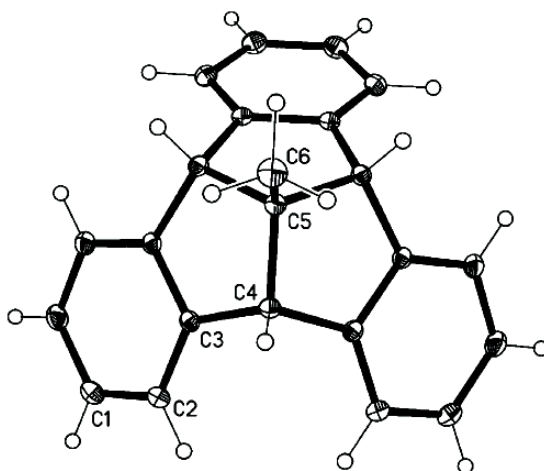


Figure 7.4.: Single-crystal molecular structure of 12d-methyltribenzotriquinacene (**3**); thermal ellipsoids are set at 50% probability. Only the asymmetric unit is numbered because of the crystallographic symmetry; for atom numbering, see Table 7.1.

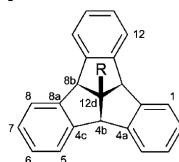
$R3m$,²⁹⁷ we repeated this structure determination at low temperatures, again suspecting the presence of a c glide plane. However, we could not detect any larger cells and, thus, the space group for compound **3** was confirmed as $R3m$. We will use the new results for compound **3** herein because the lowtemperature measurements should be marginally more reliable.

The solid-state molecular structure of compound **3** is shown in Figure 4. The central atoms C12d (C5) and C^{Me} (C6) lie along the threefold axis, which also contains a mirror plane; thus, the molecular symmetry is exactly $3m$ (C_{3v}). The molecules are stacked parallel to the c axis (Figure 7.5). Because the axis is polar, all of the stacks point in the same direction. The operator between adjacent molecules is a simple c axis translation with a repeat distance of 5.951 Å. The stacking can be interpreted in terms of C-H^{Me}... π interactions, with a H^{Me}... π distance of 2.68 Å and a H^{Me} angle of 155°.

Selected bond lengths and angles in tribenzotriquinacenes **2** and **3** are collected in Table 7.1. As expected, the presence of the centro-methyl group in compound **3** does not cause any significant change in the molecular geometry. In both cases, the carbon atoms of each of the three indane wings are coplanar to within a mean deviation of 0.01 Å. The dihedral angles within each pair of indane wings are 62.2° for compound **2** and 62.6° for compound **3**. The three axes from the central carbon atom to the midpoints of the C4c–C8a bonds are almost perpendicular to each other (angles 87.7° for compound **2** and 87.2° for compound **3**), as noted earlier for the methyl derivative (**3**).²⁹⁷ However, as has already been commented upon,²⁹⁹ the distance between the central carbon atoms (C12d)

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

Table 7.1.: Selected bond lengths [Å] and angles [°] in tribenzotriquinacenes **2** and **3**.



2 (R = H), **3** (R = Me)

Bond length ^a	2	3	Bond angle	2	3
C12d–C4b	1.5593(15)	1.5658(17)	C4b–C12d–C8b	107.30(9)	106.61(10)
			C4c–C4b–C12d	104.70(9)	105.12(9)
C4b–C4c	1.5172(16)	1.5142(13)	C4a–C4b–C12d	104.72(9)	
C4b–C4a	1.5140(16)		C4b–C4c–C8a	111.52(10)	111.52(7)
			C4c–C8a–C8b	111.72(10)	
C4c–C8a	1.3949(19)	1.400(2)	C4a–C4b–C4c	113.47(10)	113.53(11)
C4c–C5	1.3956(17)	1.3961(14)	C4b–C4c–C5	128.30(12)	128.00(10)
C5–C6	1.392(2)	1.3956(15)	C8–C8a–C8b	127.76(12)	
C6–C7	1.391(2)	1.396(2)	C4c–C5–C6	119.11(12)	119.04(10)
C7–C8	1.3908(19)		C7–C8–C8a	119.19(12)	
C8–C8a	1.3958(18)		C5–C6–C7	120.74(12)	120.53(6)
			C6–C7–C8	120.29(12)	
			C8a–C4c–C5	120.16(11)	120.42(6)
C12d–C ^{Me}	–	1.524(3)	C4c–C8a–C8	120.50(11)	

^a Libration corrections³¹⁸ of +0.001 Å and +0.002 Å to all C–C bond lengths were calculated for compounds **2** and **3**, respectively.

of adjacent molecules of compound **2** decreases to 4.75 Å, as compared to a distance of 5.95 Å between the central carbon atoms of adjacent methyl-substituted TBTQ molecules.

7.2.2. Theoretical Studies on the Solid-State Structures of Tribenzotriquinacenes **2** and **3**

From an intuitive point of view, the unidirectional stacking of tribenzotriquinacenes **2** and **3** in the solid state appears to be a consequence of their highly rigid, C_{3v} -symmetrical carbon framework and the particular spatial orientation of the aromatic units of the three mutually orthogonal indane wings. However, the question arises to what extent the strictly “translational” stacking, as found earlier in the case of compound **3**, would be preferred over some alternating–albeit unidirectional–stacking of the molecules. As mentioned above, the slight back-and-forth rotation of the molecules within the stacks of parent compound **2** was a surprising detail that was only found upon refined X-ray analysis. Moreover, to what extent would the different distances between molecules of compounds **2** and **3** be reflected in the stability of the stacks? In turn, would the stability within the stacks have some consequence for the propensity of the stacks to orient themselves unidirectionally? To address these and other questions, detailed theoretical studies, which would allow us to simulate the intermolecular interactions between adjacent

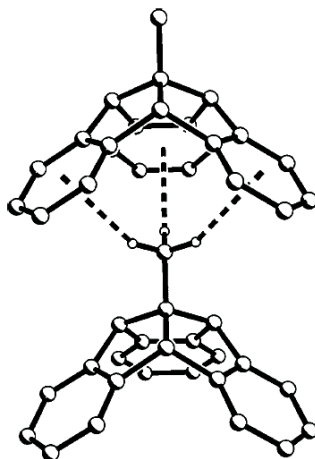


Figure 7.5.: Molecular stacking in 12-methyltribenzotriquinacene (**3**); C–H^{Me}... π interactions are represented as dashed lines. There is no mutual rotation of the molecules (cf. Figure 7.3 and Figure 7.8).

TBTQ molecules, were highly desirable. From a theoretical point of view, the usefulness of our pairwise dispersion interaction for density functional theory was to be tested for the strongly bowlshaped tribenzotriquinacenes.

To this end, we used dispersion-corrected density functional theory to calculate the crystal structures and ab initio sublimation energies. To take long-range electron correlation into account, we included the atom-pairwise dispersion correction D3(BJ).^{26,87,101} Thus, we decomposed the total energy into a density functional part and an additive dispersion part, that is, $E_{DFT-D} = E_{DFT} + E_{disp}$. Computational details that are relevant for the periodic calculations for predicting the (Me-)TBTQ crystal structure are presented and the geometries of the optimized TBTQ and Me-TBQT crystals are discussed. An investigation of the convergence of the properties of TBTQ oligomers to the infinite polymeric stacks is reported. Therein, we specify the computational details for the non-periodic oligomer calculations and present an extrapolation to the infinite stack.

Structural Optimization–Computational Details

The periodic electronic structure calculations were carried out using the Vienna Ab-initio Simulation Package (VASP).^{119,164} We began the geometry optimization with the experimental X-ray structure and six molecules per unit cell. A full optimization, including all atom positions and cell parameters, was carried out. To achieve a good balance between the computational efficiency and computational cost, the generalized gradient-approximated functional of Perdew, Burke, and Ernzerhof (PBE)^{121,180} was used, in combination with a projector-augmented plane-wave basis set (PAW).^{122,122} The energy cutoff

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

for the plane wave set was set at 1000 eV to achieve energy convergence. The Brillouin zone was sampled by a Γ -centered mesh of $(2 \times 2 \times 4)$ k -points. All monomer energies were calculated at the Γ point in a large unit cell of size $20 \times 20 \times 10 \text{ \AA}^3$. The energy convergence conditions for the solution of the selfconsistent Kohn–Sham equations (SCF) was $\Delta E < 10^{-7}$ eV. We used the VASP-internal quasi-Newton optimization algorithm in Cartesian coordinates. The structures were optimized until the forces on all atoms and cell parameters were smaller than 0.01 eV/\AA . Long-range electron correlation was taken into account by adding the above-mentioned D3 dispersion correction.^{26,87} The D3 corrections were applied with the locally developed program dftd3. In the following section, DFT-D3 always corresponds to the above specifications PBE-D3(BJ)/PAW. The crystals analyzed herein contain 228 atoms per unit cell for tribenzotriquinacene (**2**) and 246 atoms per unit cell for the methyl derivative **3**.

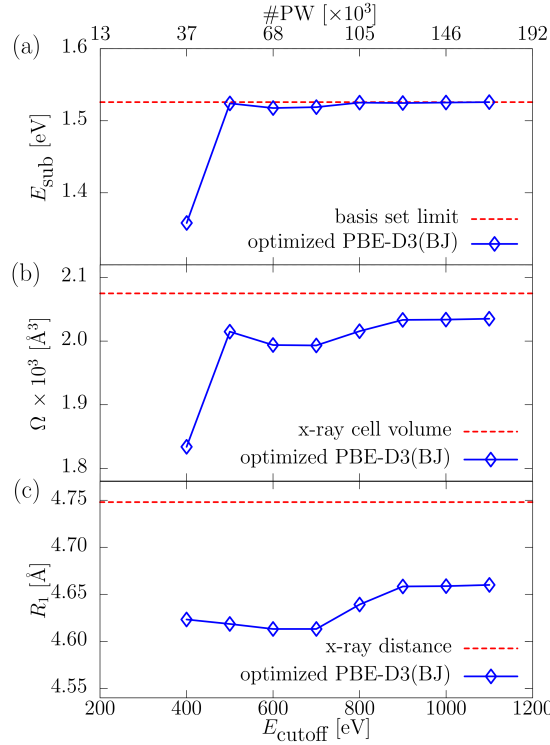


Figure 7.6.: Dependence of the computed properties on the size of the basis set: a) Sublimation energy, E_{sub} ; b) cell volume, Ω ; and c) stacking distance, R_1 . For each point, a full optimization with a fixed size of the basis set was performed. The size of the basis set was characterized by the energy cutoff, E_{cutoff} , which corresponded to a certain number of plane waves, #PW.

The dependence of the sublimation energy on the basis set is of particular interest. As evident from Figure 7.6a, a cutoff of 1000 eV is close to the complete basis-set limit. We investigated whether the structural parameters, in particular the cell volume (Ω) and the

distance of stacked molecules (R_1), could be reproduced with a smaller basis set. However, such a reproduction was found not to be the case, as shown by the volume and distance dependence on the basis set in Figure 7.6b, c. The cutoff of 400 eV, as recommended for VASP calculations,¹²³ is certainly not sufficient for obtaining reliable geometries for the molecular crystals. The balance between electrostatic repulsion and dispersive attraction has a major impact on the cell parameters. Therefore, the cell volume and the stacking distance are very sensitive towards the basis size. The dispersion correction is independent of the basis set, because it is only a function of the atomic coordinates. With a larger basis set, the calculated cell volume approaches the X-ray value (within 1.4%). Our results imply that the repulsion is underestimated by the DFT-PBE contribution for a small basis set.

Optimized Crystal Structure of Tribenzotriquinacene (**2**)

Figure 7.7 shows the theoretically determined crystal structure of tribenzotriquinacene (**2**). A projection onto the ab plane, equivalent to a view along the molecular axes, is presented.

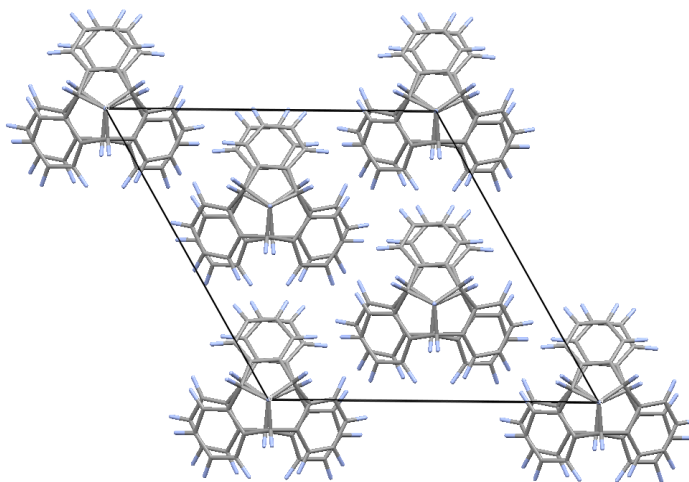


Figure 7.7.: Projection of the theoretically determined crystal structure of tribenzotriquinacene (**2**) onto the ab plane.

We start with a comparison of the theoretically and experimentally determined bond lengths, bond angles, and torsion angles of a molecule of compound **2** that is embedded in a crystal (Table 7.2, left). As expected for state-of-the-art DFT methods, our calculations are accurate: The bond lengths deviate by less than 0.006 Å, the angles deviate by less than 0.2°, and the torsion angles deviate by less than 0.6°. More interesting and challenging is the arrangement of neighboring molecules (Table 7.2, right). The shape of

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

Table 7.2.: Comparison of the structural parameters of tribenzotriquinacene (**2**) as obtained by X-ray diffraction and DFT-D3 calculations. Left: Bond lengths, angles, and torsion angles of a single molecule. Right: Cell parameters and significant intermolecular distances and angles. All distances are in [Å] and all angles in [°].

Bond length	DFT-D3	$\Delta_{theo-exp}$	Cell parameter	DFT-D3	$\Delta_{theo-exp}$
C12d-C4b	1.559	0.000	<i>a</i>	15.92	-0.04
C4b-C4c	1.515	-0.002	<i>b</i>	15.92	-0.04
C4c-C5	1.396	0.000	<i>c</i>	9.32	-0.17
C5-C6	1.398	0.006	α	90.00	0.00
C12d-C ^{Me}	–	–	β	90.00	0.00
			γ	120.00	0.00
Bond angle			Sign. properties		
C4c-C5-C6	119.3	0.2	cell volume, Ω^a	2046.29	-28.69
C4b-C4c-C5	128.4	0.1	R_1^b	4.66	-0.09
C4c-C4b-C12d	104.7	0.0	R_2^c	9.32	0.01
C4b-C12d-C ^{Me}	–	–	θ^d	9.8	3.6
Torsion angle			θ_1^e	58.6	0.1
C4b-C4c-C5-C6	179.6	0.6	θ_2^e	62.7	-0.1
C12d-C4b-C4c-C5	179.5	0.2	θ_3^e	58.6	0.1

^a Volume of one unit cell, which includes six molecules. ^b Average distance between the central carbon atoms (C12d) in two vertically stacked molecules. ^c Average distance between the central carbon atom (C12d) of two molecules in neighboring stacks. ^d Average torsion angle between two vertically stacked molecules. ^e Average angles between neighboring molecules in the *ab* plane.

the cell is reproduced very well: Cell parameter *c* (i.e., the axis length along the stacking direction) has the largest deviation (0.17 Å). This result can be explained by a slight over-binding, which is presumably produced by a dispersion attraction that is too strong or by the neglect of thermodynamic contributions. A small underestimation of the cell parameter lies well within the typical extent of thermal expansion. To analyze this result in detail, we compared further structural parameters. The cell volume, Ω , was very sensitive towards the above-mentioned attraction/repulsion balance. The observed deviation of only 1.4% is considered to be accurate. In the calculations, the R_1 distance between two vertically stacked molecules was underestimated by 0.09 Å. The torsion angle between stacked molecules, Θ , was overestimated by 3.6°, but it was qualitatively correct. The torsion has to be carefully interpreted because the corresponding potential is very flat. The arrangement of the molecules in the *ab* plane, including the angles and distances between neighboring molecules, is very accurate; the small deviation from the close-packed structure (in both theory and experiment) can be explained by the small differences in the length of the *c* axis and the alternating mutual rotation between neighboring molecules.

To judge the overall agreement between the experimental and theoretical structures, we calculated the mean absolute deviation for all of the atomic positions. The deviation, excluding hydrogen atoms, is RMSD=0.11 Å. As already analyzed, the distance between

two vertically stacked molecules is systematically underestimated by 0.09 Å. From the RMSD value, we can conclude that this difference is the dominant deviation between the experimental and theoretical structures.

To demonstrate the importance of the London dispersion correction, we calculated the sublimation energy (without phonon contribution) with and without the dispersion correction. The results are presented in Table 7.3. Because of the complete neglect of

Table 7.3.: Sublimation energy of a crystal of tribenzotriquinacene (**2**), as determined from the difference between the monomer and crystal energy per molecule with and without dispersion correction. Phonon contributions are ignored [kcal·mol⁻¹].

	DFT	DFT-D3(BJ)
E(monomer)	-5933.6	-5970.2
E(crystal)/n	-5930.9	-6005.3
ΔE	-2.6	35.2

phonon modes, the values cannot be directly compared to the experimental data. For molecular complexes, typical phonon contributions to dimerization energies are 10–20.⁴⁰ Typical sublimation energies for organic crystals of this size are 30–40 kcal·mol⁻¹, for example, 33.6 kcal·mol⁻¹ for picene.^{319,320} We note that, without the dispersion interaction, the sublimation energy even has the wrong sign and no net binding is observed. The inclusion of dispersion by using our pairwise correction gives a reasonable value of about 35 kcal·mol⁻¹.

Optimized Crystal Structure of 12d-Methyltribenzotriquinacene (**3**)

To further analyze the stacking effects of the tribenzotriquinacene bowls, the intermolecular distance and torsion in the crystals of the 12d-methyl derivative (**3**) were investigated by using theory and compared to the experimental data. The calculated crystal structure of hydrocarbon **3** is shown in Figure 7.8.

A comparison of the theoretical and experimental bond lengths, bond angles, and torsion angles of a single molecule of compound **3** (Table 7.4, left) shows no significant difference with respect to the parent compound. The molecular structure is only very slightly perturbed by the presence of the methyl group. The agreement between the theoretical and experimental lengths and angles is excellent. Larger differences between the theoretical and experimental structures can be seen in the cell parameters (Table 7.4, right). The first effect of the methyl group is as expected: It pushes the molecules of compound **3** farther apart by a distance that approximately corresponds to the bond length $R(\text{C12d}-\text{C}^{Me})$, without changing the overall lattice symmetry of the unit cell. The system crystallizes trigonally in space group $R\bar{3}m$, which is reproduced by theory. The

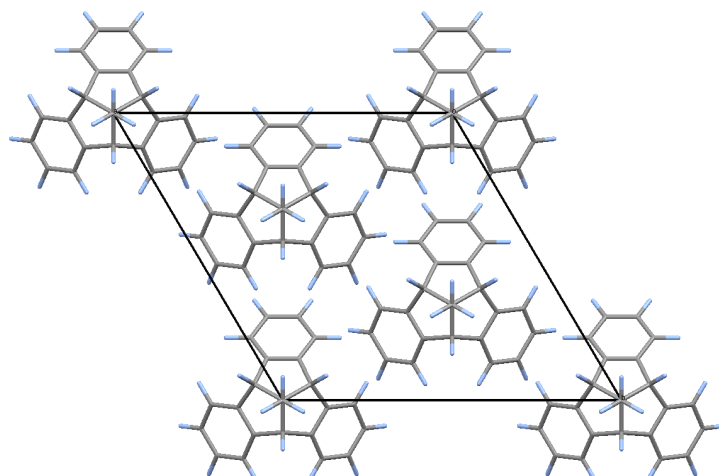


Figure 7.8.: Projection of the theoretically determined crystal structure of 12d-methyltribenzotriquinacene (**3**) onto the ab plane.

cell volume deviates by 1.5%. In the case of compound **3**, the slight over-binding is uniformly partitioned in all directions, in contrast to the crystal of the parent compound (**2**). The intermolecular distances along the c direction and in the ab plane as computed for compound **3** are too small. Whereas the first result is better than the calculation for the crystal of compound **2**, the latter result is slightly worse. This result can be intuitively explained by inspecting the molecular geometries. The molecule of the parent tribenzotriquinacene (**2**) has a curved but approximately 2D geometry. Because the angular dependence of the London dispersion interactions is minor and atom-pairwise to the first approximation, the dispersive attraction dominates perpendicularly to this plane. The additional methyl group perturbs this 2D structure and enlarges the in-plane dispersion attraction, as observable (in part) in the smaller average nearest-neighbor distance in the ab plane, R_2 . Whereas the attraction in the c direction is still present and, because of the additional atoms, the dispersion is potentially even stronger, the larger average distance between stacked molecules, R_1 , cancels this effect. The molecules stack without torsion. Presumably, the longer c distance and shorter a and b distances influence this result. Once again, we should not overestimate this effect. However, it is remarkable that, as shown later, the dimer geometry of compound **2** is perfectly C_{3v} symmetrical; therefore, the torsion is associated with the 3D arrangement.

To further demonstrate the importance of an accurate description of dispersion interactions in molecular crystals, we calculated the dependence of the total energy on cell parameter c . The potential-energy scan, as shown in Figure 7.9, shows a plot of the total energy per molecule of compound **3** as a function of cell parameter c . With fixed

Table 7.4.: Comparison of the structural parameters of 12d-methyltribenzotriquinacene (**3**) as obtained by X-ray diffraction and DFT-D3 calculations. Left: Bond lengths, angles, and torsion angles of a single molecule. Right: Cell parameters and significant intermolecular distances and angles. All distances are in [Å] and all angles in [°].

Bond length	DFT-D3	$\Delta_{theo-exp}$	Cell parameter	DFT-D3	$\Delta_{theo-exp}$
C12d-C4b	1.556	0.000	a	14.90	-0.06
C4b-C4c	1.512	-0.002	b	14.90	-0.06
C4c-C5	1.396	0.001	c	11.82	-0.08
C5-C6	1.398	0.002	α	90.00	0.00
C12d-C ^{Me}	1.517	-0.007	β	90.00	0.00
			γ	120.00	0.00
Bond angle			Sign. properties		
C4c-C5-C6	119.3	0.2	cell volume, Ω^a	2272.09	-33.99
C4b-C4c-C5	128.1	0.1	R_1^b	5.91	-0.04
C4c-C4b-C12d	105.2	0.0	R_2^c	8.82	0.04
C4b-C12d-C ^{Me}	112.3	0.1	θ^d	0.0	0.0
Torsion angle			θ_1^e	57.6	0.0
C4b-C4c-C5-C6	179.8	0.7	θ_2^e	64.8	0.0
C12d-C4b-C4c-C5	179.5	-0.3	θ_3^e	57.6	0.0
C ^{Me} -C12d-C4b-C4c	119.7	-0.3			

^{a-e} as defined in Table 7.2

cell constants, we optimized the geometry at each point, according to the computational details described previously. By using the pure PBE functional, the energy decreases monotonously with increasing cell parameter c . As a consequence, the molecular crystal is not bound within this level of theory. Only the inclusion of dispersion interactions qualitatively affords the potential and quantitatively affords the correct form, with a minimum that is very close to the experimental X-ray value. A sublimation energy of 29 kcal·mol⁻¹ for compound **3** is smaller than the sublimation energy of 35 kcal·mol⁻¹ for compound **2**. This result agrees with the experimental observation that the parent compound (**2**) crystallizes much more readily than the 12d-methyl derivative (**3**), even from hot solvents.

7.2.3. Oligomer-to-Polymer Convergence of Stacked Tribenzotriquinacene **2**

In this section, we will discuss the convergence behavior of dimerization energies and stack distances of axially stacked tribenzotriquinacene oligomers towards the properties of an infinite 1D TBTQ stack. For this purpose, we investigated the molecular distances, R_1 , and dimeization energies, E_{dim} , of the parent tribenzotriquinacene (**2**) with an increasing number of molecules. We use the terms “oligomer” and “polymer” as convenient descriptors for aggregates of non-covalently bound molecules.

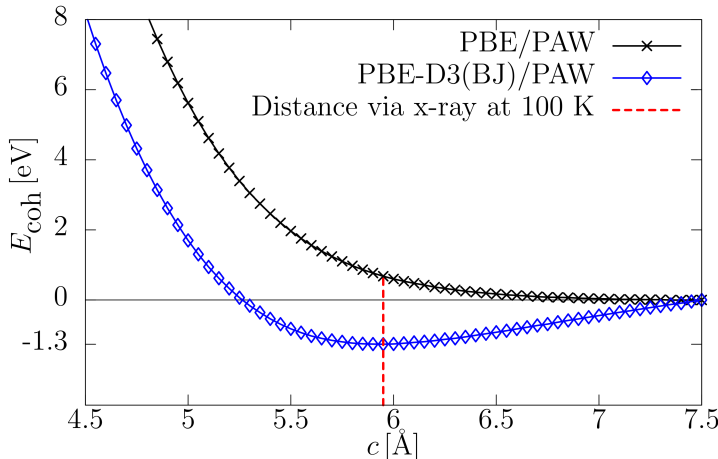


Figure 7.9.: Potential-energy surface, that is, the dependence of the cohesive energy per molecule of 12d-methyltribenzotriquinacene (**3**) on the vertical cell parameter, c , with and without dispersion interactions.

Computational Details

The non-periodic electronicstructure calculations were carried out by using the DFT program TURBOMOLE.²²⁵ For transferability, we used the PBE functional^{121,180} in combination with the atom-centered def2-QZVP(-f,-g) basis set.¹³⁴ The geometries were optimized at generalized internal coordinates.³²¹ A dispersion correction was added by using our program dftd3.^{26,87} The infinite 1D polymer was again optimized by using the VASP program package with analogous computational details to those given previously. We modeled the 1D polymer by using a large unit cell of dimensions $40 \times 40 \text{ Å}^2$ in the ab plane. The Brillouin zone was sampled by a Γ -centered mesh of $(1 \times 1 \times 2)$ k -points. With both basis sets, we are close to the complete basis-set limit and the results should be comparable.

The oligomers were optimized by using the relax code of TURBOMOLE without any symmetry constraints. In contrast to the crystal structure of the parent tribenzotriquinacene (**2**), no torsion occurred between the molecules. Because it is unlikely that one would break this symmetry by increasing the number of molecules, we optimized the polymer with one molecule per unit cell. We determined the optimized molecular distance in the periodic system through a potential scan with a fixed molecular distance, but full optimization of all other degrees of freedom of the molecule. By relaxing all of the atomic positions with a fixed unit cell, we obtained a smooth potential-energy surface with a minimum at $R_1 = 4.47 \text{ Å}$.

Extrapolation to the Infinite Polymer

The values of the molecular distances, R_1 , and the dimerization energies per pair, E_{dim} , as a function of the number of molecules are summarized in Table 7.5. Both the molecular

Table 7.5.: Molecular distance, R_1 , dimerization energy per pair, E_{dim} , and dipole moment, μ , as functions of the number of molecules, n (see text).

	R_1 [Å]	E_{dim} [kcal·mol ⁻¹]	μ [Debye]
dimer ^a	4.549	-15.6	0.8
trimer ^a	4.510	-15.8	1.0
tetramer ^a	4.495	-15.9	1.1
1D polymer ^b	4.470	-16.2	1.4
3D crystal ^b	4.660	-35.2	1.1

distances and the dimerization energies per pair were found to decrease monotonously with an increasing number of molecules, n . As shown by the increasing dipole moments, electronic induction probably causes the stronger intermolecular interaction. We fitted a power law, $f(n)$, to both data sets and enforced the correct limits

$$f(n) = \frac{1}{1 + \lambda n^\zeta} + \eta. \quad (7.1)$$

The fit parameters (λ, ζ, η) that were used to determine the dependence of distance and dimerization energy were (3.45, 1.76, 4.47) and (0.17, 1.81, -16.23), respectively. Critical exponents close to 2 are reasonable. The results as obtained from extrapolating the molecular TURBOMOLE calculations and the periodic VASP calculations agree almost perfectly, as shown in Figure 7.10. A fit procedure with three points is certainly a crude approximation, but both trends are correct. In one dimension, it seems to be possible to extrapolate from the oligomer properties (distances and dimerization energies per pair) to the corresponding polymer properties. However, a direct continuation into three dimensions is not possible. For instance, the molecular distances decrease in the 1D model, whereas the corresponding distance in the crystal is significantly larger than the dimer distance. Hence, one should always be careful in making predictions of the crystal properties from molecular calculations.

7.3. Conclusions

The solid-state molecular and crystal structures of parent tribenzotriquinacene (**2**), a bowl-shaped, C_{3v} -symmetrical, and rigid polycyclic hydrocarbon, were investigated by using X-ray diffraction and compared to the corresponding structures of its centro-methyl-

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

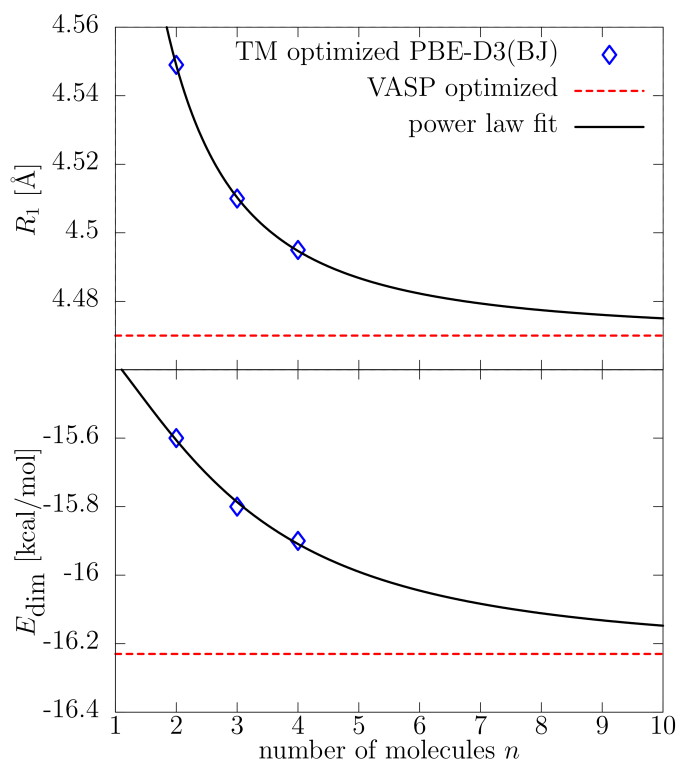


Figure 7.10.: a) Molecule distances, R_1 , as calculated by using TURBOMOLE for the dimer, trimer, and tetramer and by using VASP for the polymeric stack of tribenzotriquinacene (**2**). b) Corresponding dimerization energies per pair, E_{dim} . All of the calculations were performed at the PBE-D3(BJ) level of theory. The fitted power laws illustrate the trends.

substituted homologue, 12d-methyltribenzotriquinacene (**3**), as studied earlier.²⁹⁷ Theoretical studies, which were performed by using dispersion-corrected density functional theory (DFT-D3), were found to reproduce the crystal structures of tribenzotriquinacenes **2** and **3** with excellent agreement, not only in terms of the molecular structures but also in the details of the molecular packing. In particular, the intermolecular distances within the strictly coaxial molecular stacks (X-ray data: 4.75 and 5.95 Å for compounds **2** and **3**, respectively) were reliably calculated (DFT-D3: 4.66 and 5.91 Å, respectively). Moreover, the slight alternating rotations of the molecules of compound **2** within the stacks and the absence of this effect within the stacks of compound **3** could be reproduced by theory. The negative sublimation energy of both systems, as calculated in the framework of the pure density functional, which would indicate no net bonding in the solid state, shows the importance of an accurate treatment of the dispersion interaction. The excellent agreement between the calculated and experimental results suggests that our pairwise dispersion correction is capable of describing the relevant physics at a sound level. Furthermore, we show that the extrapolation of the dimerization energy and intermolecular

distance of oligomeric stacks of compound **2** to the infinite 1D polymer is possible and straightforward. The permanent solid-state dipole moment of tribenzotriquinacene **2** and its suitable TBTQ derivatives may be of interest with respect to the pyroelectric effect and electron-transport properties of similar compounds.^{322–325}

7.4. Experimental Section

X-ray Structure Determination

Crystallographic data for compounds **2** and **3** are presented in Table 7.6. *Data collection and reduction:* Crystals were mounted in inert oil onto glass fibers and transferred into the cold gas stream of an Oxford Diffraction Xcalibur E diffractometer. Measurements were performed at 100 K by using monochromated $MoK\alpha$ radiation ($\lambda = 0.71073 \text{ \AA}$). No absorption corrections were applied. *Structure refinement:* The structures were refined anisotropically against F^2 (SHELXL-97).³²⁶ The methyl hydrogen atom of compound **3** was freely refined; all other hydrogen atoms were included using a riding model. For both structures, the anomalous dispersion was insignificant and, thus, the absolute structure could not be determined. Therefore, Friedel opposite reflections were merged and the Flack parameter was meaningless.

7. Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State

Table 7.6.: Crystallographic data for tribenzotriquinacenes **2** and **3**.

Compound	2	3
formula	C ₂₂ H ₁₆	C ₂₃ H ₁₈
M_r	278.33	294.37
color/shape	colorless prisms	colorless tablets
crystal size [mm ³]	0.3 × 0.15 × 0.1	0.4 × 0.4 × 0.2
crystal system	trigonal	trigonal
space group	<i>R</i> 3 <i>c</i>	<i>R</i> 3 <i>m</i>
a,b [Å]	15.8850(6)	14.9573(3)
c [Å]	9.4953(4)	5.9513(2)
V [Å ³]	2074.98	1153.04
Z	6	3
ρ_{calcd} [Mgm ⁻³]	1.346	1.272
μ [mm ⁻¹]	0.08	0.07
$F(000)$	888	468
$2\theta_{\text{max}}$	60	62
total reflns	25268	22910
unique reflns	681	447
R_{int}	0.027	0.021
parameters	67	43
wR (F ² , all reflns)	0.082	0.079
$R(F > 4\sigma(F))$	0.031	0.030
S	1.05	1.11
max. Δ/ρ [eÅ ⁻³]	0.29	0.31

CCDC-907602 (**2**) and CCDC-907601 (**3**) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

8. Crystal Packing Induced Carbon-Carbon Double-Triple Bond Isomerization in a Zirconocene Complex

Jan Gerit Brandenburg[†], Georg Bender[§], Jinjun Ren[‡], Andreas Hansen[†], Stefan Grimme[†], Hellmut Eckert[‡], Constantin G. Daniliuc[§], Gerald Kehr[§], and Gerhard Erker[§]

Section: Organometallic and Organometalloidal Compounds Received 27th of June 2013, Published online 15th of September 2014

Reprinted (adapted) with permission from

J. G. Brandenburg, G. Bender, J. Ren, A. Hansen, S. Grimme, H. Eckert, C. G. Daniliuc, G. Kehr, and G. Erker *Organometallics*, **2014**, *33*, 5358–5364.

— Copyright ©2014 American Chemical Society. DOI 10.1021/om500678p

Own manuscript contribution

- Computations of crystal structures, gas phase potentials, and NMR tensors
- Interpretation of computational results
- Writing the theory part and main discussion of the manuscript

[†]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstraße 4, 53115 Bonn, Germany

[‡]Institut für Physikalische Chemie, Westfälische Wilhelms-Universität Münster, Corrensstraße 28/30, 48149 Münster, Germany

[§]Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Corrensstraße 40, 48149 Münster, Germany

Abstract We present a combined theoretical and experimental analysis of the carbon–carbon bond character in two prototypical zirconocene complexes. The two cyclic seven-membered ring zirconium compounds **2a** and **2b** differ by the substitution of a tert-butyl by a trimethylsilyl group. Due to the coordination of the π -system to the metal atom, a formally forbidden (η^2 -allenyl)/enamido-Zr to (η^2 -alkyne)/ κ Nimine- Zr complex isomerization is feasible. State-of-the-art dispersioncorrected density functional theory (DFT-D3) is used in both the solid and condensed phase to examine and quantify the experimental structures (X-ray diffraction) and ^{13}C NMR magnetic shielding. The complementary investigations demonstrate the importance of nonlocal London dispersion interactions. Both X-ray structures agree excellently with the results of the solid-state DFT-D3 calculations. Interestingly, **2b** exhibits a mixed allene–alkyne form in the solid state, while its gas phase structure has a strong allene character. The substitution leading to **2a** prevents this isomerization in the solid state by the intramolecular stabilization of the allene structure. NMR solid and liquid phase measurements confirm the theoretically proposed transition. By combining the experimental and theoretical information, the rather unusual triple/single to double/double-bond transition is attributed to an intermolecular London dispersion induced crystal packing effect.

8.1. Introduction

We had recently shown that the five-membered zirconacycloallenoid complex^{327–330} **1a** reacts with a variety of unsaturated reagents to form seven-membered metallacycles that contain an (η^2 -allenyl)zirconocene subunit.³³¹ The reaction of **1a** with acetonitrile is a typical example. It gives the cyclic (allenyl)- zirconocene complex **2a** in good yield. Complex **2a** shows the typical structural features of this class of compounds, which will briefly be reviewed in the main text of this account. We have now reacted the tert-butyl-substituted zirconacycloallenoid **1b** with acetonitrile and obtained the seven-membered metallacyclic product **2b**. Formally compound **2b** is analogous to **2a**, but it features amazing structural differences in detail in the solid state, indicating a pronounced distortion from an allenyl zirconocene toward an alkyne zirconocene bonding situation. This study is aimed at arriving at a theoretical explanation for this rather unusual phenomenon.

The two isomers are specifically influenced by crystal packing opposed to solvent interactions. The theoretical prediction of these crystal packing effects (CPE) is rather challenging due to the various electronic contributions. Long-range interactions of electrostatic, induction, exchange, and London dispersion type have to be treated accurately. Related to the calculation of CPE is the analysis of conformational polymorphism^{6,7} and

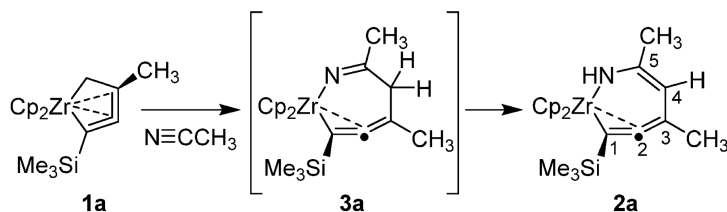
the ambitious aim of organic crystal structure prediction, which is a highly active research field.^{37,41,72,229,230} State-of-the-art dispersioncorrected density functional theory (DFT-D3) methods for molecular as well as for the solid-state electronic and geometric structure are applied to quantify and analyze the experimental X-ray and NMR data.

8.2. Results and Discussion

8.2.1. Preparation and Structural Characterization of the Compounds **2**

We had previously synthesized the five-membered zirconacycloallenoids **1a** (R=SiMe₃) and **1b** (R=CMe₃) by in situ generation of zirconocene by the Negishi/ Takahashi method³³² in the presence of the conjugated enynes R–C≡C–C(Me)=CH₂, respectively.^{333–335} Complex **1a** had been treated with acetonitrile (see Scheme 1). The nitrile reagent had inserted

Scheme 1



into the Zr–C(sp³) bond of **1a** to probably generate the seven-membered metallacycle **3a**, which underwent a rapid subsequent Zr-imido/-enamido tautomerization reaction^{336,337} under the applied reaction conditions to eventually give the product **2a**.³³¹

Table 8.1.: Comparison of Selected Structural Data of the Complexes **2a** (R = SiMe₃) and **2b** (R = CMe₃) in the Crystal^a

compound	2a (R = SiMe ₃) ^b	2b (R = CMe ₃) ^c
Zr–C1	2.257(2)	2.220(3)
Zr–C2	2.471(2)	2.186(3)
Zr–N	2.166(2)	2.292(3)
C1–C2	1.284(3)	1.304(4)
C2–C3	1.342(3)	1.437(5)
C3–C4	1.455(3)	1.343(5)
C4–C5	1.352(3)	1.443(5)
C5–N	1.363(3)	1.285(4)
Zr–N–C5	131.2(1)	133.8(2)
C1–C2–C3	173.0(2)	148.8(3)

^aBond lengths in Å, angles in deg. ^bFrom ref³³¹. ^cThis study.

8. Crystal Packing Induced Bond Isomerization

The X-ray crystal structure analysis of **2a** showed typical (η^2 - allenyl)Zr characteristics. The Zr–C1 bond is short (see Table 8.1), and the Zr–C2 bond is markedly longer. The Zr-allenyl unit has a close to linear framework, and the substituent planes at C1 and C3 are not too far away from orthogonality. The C1–C2 bond is very short, and the adjacent C2–C3 bond is in the typical C=C double-bond range. The C3–C4 distance indicated a typical C(sp²)–C(sp²) single bond, and C4–C5 again is in the C=C double-bond range. The N–C5 bond is long, as expected for a typical Zr-enamido system. The previously reported NMR spectra³³¹ are in accord with this metallacyclic (σ -allenyl)Zr(IV) complex structure.

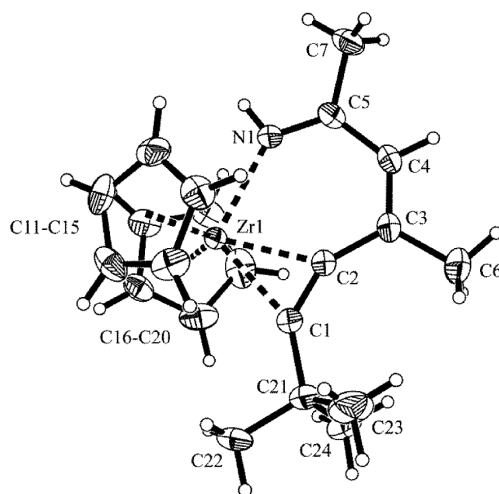
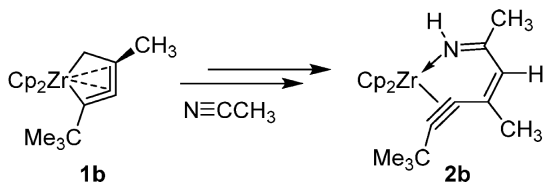


Figure 8.1.: Molecular structure of complex **2b** in the crystal (thermal ellipsoids are shown with 30% probability).

We prepared complex **2b** by treatment of the five-membered zirconacycloallenoid **1b** with acetonitrile in benzene solution at rt. Blue crystals of compound **2b** were obtained in 31% yield after crystallization from pentane at -30°C. Complex **2b** was characterized by X-ray diffraction. This showed that the compound had the same formal overall atom connectivity as **2a** but that it differed substantially in the bonding details (see Table 8.1). Overall, compound **2b** features bond characteristics close to a metallacyclic (η^2 -alkyne)zirconium(II) complex.^{338,339} It shows a markedly different sequence of single and double bonds of its carbon/nitrogen perimeter as compared to **2a**. Complex **2b** in the crystal shows a pair of almost equidistant Zr–C1/C2 bonds, quite different from the situation encountered for **2a**. The C1–C2 bond in **2b** is short, but then the adjacent C2–C3 linkage is long; it is in the C–C σ -bond range. The C3–C4 bond in **2b** is in the C=C double-bond range, and the C4–C5 bond is again long. Eventually, the C5–N bond in complex **2b** represents an imino C=N double bond (quite contrary to the situation en-

countered for **2a** in the crystal). This bonding situation of **2b** in the crystal is markedly different from the (η^2 -allenyl)/enamido-Zr type encountered for **2a**. The structure of **2b** found in the crystal might schematically be described by a (η^2 -alkyne)/ κN -imine-Zr coordination type (see Scheme 2).

Scheme 2



8.2.2. Theoretical Studies

In order to analyze the bond character (η^2 -allenyl vs η^2 -alkyne) of the cyclic seven-ring zirconium compounds **2**, we have investigated the electronic structure with dispersion-corrected density functional theory methods. To distinguish between intra- and inter-molecular interactions, we considered both molecules in the gas phase and in the solid state.

Computational Details

We optimized all structures at the PBE-D3/600 eV level with the VASP 5.3 simulation package.^{119,164} The PBE density functional¹²¹ was employed with a large projector-augmented plane-wave basis set (PAW)^{122,123} with an energy cutoff of 600 eV. The k -space was sampled with a Γ -centered grid of $2 \times 2 \times 1$ k -points generated via the Monkhorst-Pack scheme. The single molecules were evaluated at the Γ -point in a large unit cell with a minimum distance of 16 Å to neighboring images. Long-range electron correlation effects were treated with the atom-pairwise London dispersion correction D3 in the Becke-Johnson damping scheme.^{26,87} This method combination provides reliable results for both the gas phase and the solid state, as shown in a number of publications by us^{38,41,98,201} and other groups.^{104,120} In this context we have applied DFT-D3 successfully to various molecular crystals^{57,65} and were able to describe complicated CPE accurately also for organometallic structures³⁴⁰ by full solid-state (cell) optimizations. Additionally, we have confirmed the molecular structure by repeating the calculations with the Gaussian basis set based program ORCA 2.9.³⁴¹ Atomic orbital basis sets of def2-TZVP and def2-QZVP quality with the corresponding effective smallcore potential (Zr def2-ECP) are used.^{135,256,257} At this level of theory, errors in the optimized structures arising from

8. Crystal Packing Induced Bond Isomerization

basis set incompleteness can be neglected. The calculation of chemical shifts requires an accurate description of the core electron density. Therefore, we preferred to use Slater basis sets as implemented in the program ADF 2012.^{342,343} The PBE0 hybrid functional together with the TZ2P all-electron basis set was employed.^{170,344} Scalar relativistic effects as well as explicit spin-orbit coupling were treated within the ZORA approximation.³⁴⁵ The NMR shielding tensors are calculated using gaugeincluding atomic orbitals relative to a standard TMS reference.^{346,347} Due to the huge computational demands of such calculations, only GGA density functionals like PBE can be applied in these large-scale solid-state optimizations. Note that the solid-state structures of **2a** and **2b** consist of 400 and 200 atoms per unit cell, respectively. These semilocal functionals yield a systematic error for the parent allene-propyne isomerization of about 3 kcal/mol,³⁴⁸ i.e., incorrectly favor allene over propyne. This defect can be alleviated by adding nonlocal Fock exchange in the functional. We have checked this problem for the here-investigated case by single-point computations at the higher PWPB95-D3 double-hybrid level (inclusion of 50% Fock exchange and 27% SOS-MP2 correlation gives the correct parent isomerization energy within 1 kcal/mol). However, at least qualitatively the same findings as discussed below are obtained, which rules out methodological issues for the theoretically observed effects.

Results

The different geometries, namely, the X-ray crystal structures, the computed crystal structures, and the computed gas phase structures of compounds **2a** and **2b**, are presented in Table 8.2. We compare the cell parameters (cell lengths and angles) and the allenoid dihedral angle θ between the C3-CH₃ and the C1-SiMe₃ (**2a**) and C1-CMe₃ (**2b**) vectors, respectively. Because of the two orthogonal π -systems, this angle is typically close to 90° for allene. For a propyne subunit with one single and one triple bond, the dihedral angle θ is rather close to 0°. The C-C bond lengths R_1 and R_2 are used to additionally characterize the bond character (compare with Figure 8.3). The optimized crystal structures are shown in Figure 8.2.

The computed geometries agree very well with the X-ray structures. The symmetry groups (monoclinic and orthorhombic, respectively) are reproduced correctly, and all unit cell lengths deviate less than 1%. The shrinking of the unit cell volume by 4.4% and 4.2%, respectively, is within typical thermal expansions of organic crystals of approximately 3%.^{40,211} One has to consider some remaining artificial Pulay stress arising from the finite basis set. This can account for an underestimation of the cell volume of approximately 1% to 2%.⁶⁶ For compound **2a**, the intramolecular structure is very similar for the gas

Table 8.2.: Geometries of the Experimental and Theoretical Crystal Structure of Compounds **2a** and **2b** Compared to the Corresponding Gas Phase Geometries^a

	crystal		gas phase
	X-ray	DFT	DFT
2a (monoclinic)			
<i>a</i>	11.99	11.85	
<i>b</i>	9.83	9.67	
<i>c</i>	33.03	32.54	
$R_1(\text{C1-C2})$	1.285	1.298	1.298
$R_2(\text{C2-C3})$	1.342	1.350	1.358
θ	69.4	66.1	61.9
2b (orthorhombic)			
<i>a</i>	8.05	7.91	
<i>b</i>	8.79	8.70	
<i>c</i>	27.53	26.92	
β	92.9	93.1	
$R_1(\text{C1-C2})$	1.304	1.323	1.307
$R_2(\text{C2-C3})$	1.437	1.422	1.379
θ	5.6	8.3	47.3

^aCalculations were done consistently at the PBE-D3/600 eV level without any symmetry constraints. The dihedral angle θ (^tBu-C1-C3-Me) and the bond distances $R_1(\text{C1-C2})$ and $R_2(\text{C2-C3})$ are given.

Bond lengths in Å, angles in deg.

phase and the solid state. The dihedral angle of 66° and 69°, respectively, indicates a strong allene character of the -MeC=C=C(Zr)-SiMe₃ system. In contrast, the gas phase and the crystal structure of compound **2b** differ significantly. The gas phase structure exhibits again the allene-type bonds, while the solid state corresponds to an (η^2 -alkyne) Zr system. Hence, we have investigated the θ -dependent potential in detail. In Figure 8.3, the potential energy surface of compound **2b** in the gas phase with respect to the dihedral angle θ is shown. Every point corresponds to a θ -constrained optimization at the PBE-D3/ def2-TZVP level. We can identify a very flat minimum around 50°, which is in agreement with the unconstrained optimization. A free optimization at the PBE-D3/def2-QZVP level confirms the results ($R_1 = 1.306$, $R_2 = 1.371$, $\theta = 52.1^\circ$). The energy gain due to the propyne-allene transition is approximately 4.5 and 2.5 kcal/mol for **2a** and **2b**, respectively. This is corrected to 4.0 and 1.6 kcal/mol at the higher PWPB95-D3 level of theory. In both cases, the propyne form does not correspond to a minimum energy structure in the gas phase. Qualitatively, **2a**, which does not “isomerize” in the solid according to theory and experiment, shows the higher energy difference between the two forms. This larger energy gap makes the discussed transition in the solid unlikely. In **2b**, however, the smaller difference of 1.6 kcal/mol can be compensated by stronger intermolecular interactions in the solid. We attribute the preferred propyne character in the solid state to a CPE. This hypothesis was investigated by complementary theoretical

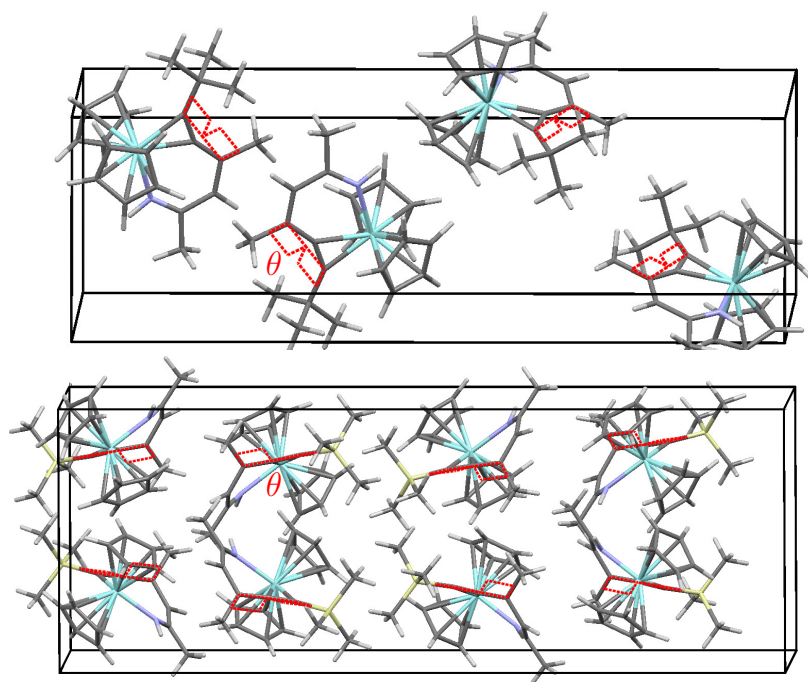


Figure 8.2.: PBE-D3/600 eV-optimized crystal geometry of **2b** (upper part) and **2a** (lower part). A comparison with the X-ray structure is given in Table 8.2

and experimental means. We also investigated the electronic substituent effect of the SiMe_3 vs the CMe_3 group, which is qualitatively opposite of that in **2a** and **2b**. Here, the allene form is stabilized by tert-butyl, while in the parent it is destabilized (also by about 2 kcal/mol), and hence, electronic substituent effects cannot explain the observations.

We have further analyzed the different behavior of **2a** and **2b**. Two smaller model systems were investigated in which the SiMe_3 and CMe_3 groups are replaced by SiH_3 (**2c**) and CH_3 (**2d**), respectively, and the rest of the structure was kept the same. The corresponding potential energy curves are included in Figure 8.3 (dashed crosses). Apparently, these groups have a significant impact on the rotation barrier. However, the effect is opposite for the two compounds: while for **2a** introducing “steric strain” increases the torsion barrier, it slightly diminishes for **2b**. Hence we can only conclude that the difference between **2a** and **2b** is related to the size of the SiMe_3 /tert-butyl substituents but in a more complicated way is also due to electronic effects arising from the silicon atom.

Long-range London dispersion effects are the dominant intermolecular forces in such molecular crystals. The substitution of the CMe_3 by a SiMe_3 group also changes these interactions. For example the C_6 dispersion coefficients of the 4-fold coordinated silicon and carbon atoms in the two groups are 150 and 18 au, respectively. In a computational “experiment” we switched off the D3 dispersion correction in the crystal optimization of

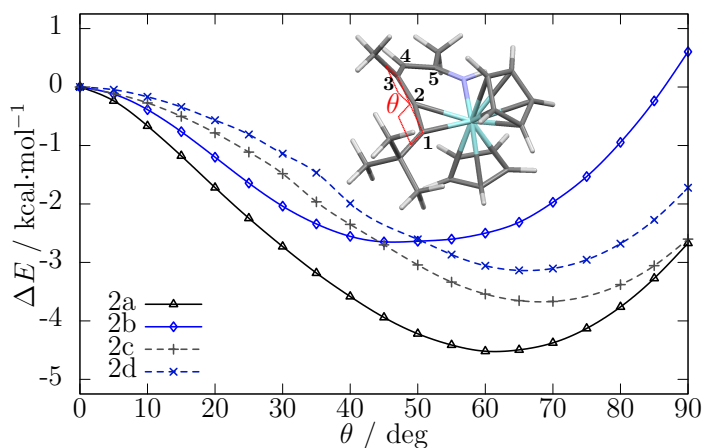


Figure 8.3.: Potential energy surface of compounds **2a** to **2d** in the gas phase with respect to the dihedral angle θ . A value of zero corresponds to a propyne form. At a fixed value θ , all other degrees of freedom are optimized at the PBE-D3/def2-TZVP level of theory. The inset shows the equilibrium geometry ($\theta = 47^\circ$).

2b with fixed cell volume. Interestingly, the molecular geometry in the solid then relaxes to the gas phase structure. The reverse procedure of optimizing the gas phase structure without dispersion correction does not change the result. Therefore, the CPE can be clearly assigned to the intermolecular London dispersion interaction.

In order to verify the theoretically proposed differences between gas and solid-state structure, we measured the ^{13}C NMR spectra of compounds **2a** and **2b** in the solid state and in solution, and we calculated the ^{13}C NMR chemical shifts of the two compounds under these different conditions and compared them. Assignments of the ^{13}C NMR resonances in the solid state were based on short contact time cross-polarization experiments: As the rates of the ^{13}C NMR signal buildup depend on the strength of the ^1H – ^{13}C dipolar interactions, the five unsaturated C atoms C1–C5 fall into three distinct groups: while the signal of the protonated carbon atom C4 is already fully developed after a very short contact time of 200 μs , the quaternary atoms C1 and C2, which are rather remote from protons, yield no signal at all under such conditions, whereas the methyl-substituted atoms C3 and C5 show intermediate behavior. Furthermore, the signal of the nitrogen-bonded C5 atom is identified by an asymmetric line shape, which arises from the dipolar coupling of ^{13}C to the ^{14}N ($I = 1$) nucleus, which is subject to a strong quadrupolar interaction. The assignment of the ^{13}C NMR signals is further aided by comparing the DFT-calculated chemical shift anisotropies and asymmetry parameters with experimental values. The latter were extracted from the spinning sideband intensity profiles of MAS NMR spectra obtained at low (3.0 kHz) rotation speeds. The results are presented in Tables 8.3 and 8.4. The DFT calculations are at the PBE0/ZORA/TZV2P level. The

8. Crystal Packing Induced Bond Isomerization

Table 8.3.: Measured and Calculated ^{13}C NMR Chemical Shifts (in ppm) for Compounds **2a** and **2b** in Both the Gas and Condensed Phase

	solid		liquid/gas		$\Delta_{\text{solid-liquid/gas}}$	
	expt	DFT ^c	expt ^a	DFT ^c	expt	DFT
compound 2a						
C1	146.6	163.5	137.1	154.7	9.5	8.8
C2	167.4	182.2	157.6	171.0	9.8	11.1
C3	118.7	133.2	115.3	127.6	3.4	5.6
C4	93.9	98.8	94.7	105.6	-0.8	-6.8
C5	154.9	168.9	155.8	168.0	-0.9	0.8
Me3	21.2	25.6	21.0	24.6	0.2	1.0
Me5	25.5	28.3	25.1	30.7	0.4	-2.4
TMS	1.6	3.0	1.8	4.4	-0.2	-0.6
compound 2b						
C1	217.4	231.1	180.9	205.6	36.5	25.5
C2	148.0	154.7	133.6 ^b	158.0	14.4	-3.2
C3	142.6	151.3	149.5 ^b	148.7	-6.9	2.6
C4	122.2	125.3	103.5	117.0	18.7	8.3
C5	174.5	181.2	162.8	175.2	11.7	6.0
Me3	25.8	28.9	22.9	27.0	2.9	1.9
Me5	40.6	34.1	27.0	32.8	13.6	1.3
^t Bu ^b	31.0	47.9	39.9	48.9	8.9	-1.0
^t Bu	33.2	36.0	32.8	36.6	0.4	-0.6

^aC₆D₆, 299 K. ^bTentatively assigned. ^cCalculations at the PBE0/ZORA/TZV2P level. Molecular structures were taken from PBE-D3/600 eV optimizations.

corresponding structures are taken from the PBE-D3/600 eV optimizations. The numbering is according to Figure 8.1. The isotropic chemical shifts are given compared to a standard TMS reference (δ -scale). The results reveal that accurate NMR calculations for the investigated zirconium complexes are indeed challenging. Table 8.3 indicates that with very few exceptions the DFT-calculated isotropic chemical shift values of the unsaturated carbon backbone atoms C1–C5 are consistently higher by $\sim 10 \pm 5$ ppm than those observed experimentally. These discrepancies are observed both in the solid state (comparison between columns 1 and 2) and in the liquid state (comparison between columns 3 and 4) and may be related to a referencing problem. Columns 5 and 6 compare the difference $\Delta_{\text{solid/liquid}}$ of the experimental isotropic chemical shift values and the difference $\Delta_{\text{solid/gas}}$ between the DFT-calculated values in the solid state and the gas phase optimized molecule (column 6). For compound **2a**, these differences are found to be small for most of the carbon atoms. Moderately sized effects are observed only for C1 and C2, the ones expected to be most affected by a change in the propyne- vs allene-type bonding character. Altogether, Table 8.3 suggests, however, that the structures of **2a** are rather similar in the liquid and the solid state (in accordance with the findings discussed above) and that the relatively slight change in bonding character is well described by the DFT

calculations. In contrast, for compound **2b** the values $\Delta_{solid/liquid}$ and $\Delta_{solid/gas}$ are rather large for most of the C atoms considered, indicating a significant structural difference between the liquid and the solid states. Besides C1 and C2, the carbon atoms C4 and C5 and the tert-butyl group show particularly strong deviations. The mean absolute deviation between the experimental ^{13}C shifts of both phases is 14 and 3 ppm for compounds **2b** and **2a**, respectively. Comparison of $\Delta_{solid/liquid}$ and $\Delta_{solid/gas}$ indicates that the agreement between experiment and theory for **2b** is not as good as for **2a**. This could be attributed to neglected solvation effects, which are expected to change the very sensitive structure of **2b** due to the flatness of the corresponding potential (see Figure 8.3).

Table 8.4.: Calculated ^{13}C NMR Chemical Shift Anisotropies (in ppm) for Compounds **2a** and **2b** in Both the Gas and the Solid State and Measured Values in the Solid State

	DFT(gas) ^b		DFT(solid) ^b		experimental	
	$\Delta\sigma$	η_σ	$\Delta\sigma$	η_σ	$\Delta\sigma \pm 5$	$\eta_\sigma \pm 0.3$
compound 2a						
C1	-96.9	0.71	-104.5	0.80	-96	0.7
C2	77.2	0.56	-71.4	0.81	-80	0.7
C3	-89.3	0.89	109.0	0.87	93	0.8
C4	78.7	0.30	74.9	0.06	74	0.1
C5	-103.2	0.14	-101.4	0.30	-98	0.5
Me3	-21.1	0.55	-17.9	0.51	-15.6	<i>a</i>
Me5	-30.5	0.63	-28.7	0.62	-30.3	<i>a</i>
TMS	11.3	0.74	10.4	0.80	-17.7	<i>a</i>
compound 2b						
C1	-121.3	0.90	-156.0	0.80	-147	0.8
C2	68.2	0.93	-89.9	0.49	-109	0.4
C3	-114.3	0.73	-117.9	0.87	-104	0.8
C4	82.4	0.32	87.6	0.35	81	0.5
C5	-116.3	0.29	-123.2	0.40	-123	0.5
Me3	-22.1	0.71	-21.1	0.76	-24.7	<i>a</i>
Me5	-32.9	0.65	-33.8	0.70	31.3	<i>a</i>
^t Bu ^b	17.2	0.93	-7.5	0.75	-27.5	<i>a</i>
^t Bu	-24.6	0.50	-24.0	0.51	-14.2	<i>a</i>

^aNot determined; experimental error too large. ^bCalculations at the PBE0/ZORA/TZ2P level.

Molecular structures were taken from the PBE-D3/600 eV optimizations.

The above conclusions are further bolstered by an analysis of the chemical shielding anisotropies $\Delta\sigma$ and asymmetry parameters η_σ of the individual C atoms present in both compounds. First of all, Table 8.4 indicates that for both **2a** and **2b** the numerical values measured in the solid state are generally found in excellent agreement with those calculated from the solid-state structures. In fact, the numerical agreement is much better than for the isotropic chemical shift values, presumably because the theoretical determination of the anisotropy is not affected by problems related to referencing. Interesting

8. Crystal Packing Induced Bond Isomerization

additional insights arise from the data comparison between the solid-state values of $\Delta\sigma$ and η_σ (experimental or DFT) and the corresponding values calculated for the gas phase molecule (no experimental liquid-state NMR data are available as the magnetic shielding anisotropy is averaged out by rapid molecular tumbling in solution). For compound **2a**, there is generally excellent agreement between the corresponding values. In contrast, for compound **2b** the experimental $\Delta\sigma$ values measured for the carbon atoms C1 and C2, the ones primarily implicated in the allene–propyne isomerism, deviate substantially from those calculated in the gas phase, which are more consistent with those calculated (and measured) for compound **2a**. Thus, the data shown in Table 8.4 also support the conclusion that the solid-state structure of **2b** is substantially different from that of the isolated molecule, suggesting that the propyne-like structure is a consequence of the molecular packing in the crystal structure.

8.3. Discussion and Conclusion

Crystal packing effects on stacking distances between inter- and intramolecular π -systems are not uncommon.^{349–352} However, a molecular crystal packing effect modifying a covalent bond character as it was found in complex **2b** is rather unusual, but was indeed encountered. At a first superficial sight one might have suspected to have a case of bond-stretch isomerism,^{353–356} but that is definitely not the case. This would have required the presence of two local minimum structures. However, our compounds are structurally clearly single minimum systems; the structures may only vary with the external conditions. The $-\text{SiMe}_3$ -substituted compound **2a** behaves conventionally. It features an $(\eta^2\text{-alleny})/\kappa N\text{-enamido-Zr}$ type metallacyclic structure in the crystal, and it is of the same structural type in solution and in the gas phase (DFT-D3). In strong contrast, the observed preferred structure of the otherwise closely related $-\text{CMe}_3$ -substituted compound **2b** is markedly dependent on the aggregation state: in solution and in the gas phase it is probably characterized as being of a distorted $(\eta^2\text{-alleny})/\kappa N\text{-enamido-Zr}$ type (although there are a few minor differences compared to **2a**), but in the crystal compound **2b** adopts a fundamentally different structure. To a first approximation, it can be described as derived from a $(\eta^2\text{-alkyne})/\kappa N\text{-imine-Zr(II)}$ type, although this might be a bit of a too schematic description. Nevertheless, the strong structural dependence of compound **2b** (in contrast to **2a**) on crystal forces is quite unusual and noteworthy. It once again emphasizes the chemical importance of formally “weak” noncovalent interactions in large and complex systems.

8.4. Experimental Section

8.4.1. Preparation of Complex 2b

At -78°C *n*-butylmagnesium chloride solution (0.34 mL, 2 M diethyl ether solution, 0.68 mmol, 2 equiv) was added to a solution of dichlorobis(η^5 -cyclopentadienyl)- zirconium (100 mg, 0.34 mmol) and 2,5,5-trimethyl-1-hexen-3-yne (42 mg, 0.34 mmol) in THF (5 mL). After removal of the dry ice bath, the mixture was warmed to rt and stirred for 1 h. Subsequently the yellow solution was heated at 60°C for an additional 1 h. All volatiles were removed in vacuo, the residue was suspended in *n*-pentane (3×5 mL), and the insoluble material was filtered off through Celite. The orange-red filtrate was concentrated in vacuo, and the oily residue was dissolved in benzene (5 mL). Then acetonitrile (40 mg, 0.97 mmol, 2.9 equiv) was added to the solution, which was stirred overnight at rt. Then all volatiles were removed in vacuo, the obtained residue was suspended in *n*-pentane (2×10 mL), and the suspension was filtered through a short pad of Celite. Crystallization from *n*-pentane at -30°C gave the complex 2b as blue crystals (40 mg, 31%), which were suitable for X-ray crystal structure analysis. MS-ESI-EM: calcd for $[\text{C}_{21}\text{H}_{27}\text{NZr} + \text{H}]^{+}$ 384.1263 g/mol, found 384.1247 g/mol. Anal. Calcd for $\text{C}_{21}\text{H}_{27}\text{NZr}$: C 65.57, H 7.07, N 3.64. Found: C 64.88, H 7.08, N 3.48. ^1H NMR (500 MHz, 299 K, C_6D_6): δ 5.52 (s, 10H, Cp), 4.98 (s, 1H, 4-H), 4.83 (br, 1H, NH), 1.94 (d, $^4J_{\text{HH}} = 0.8$ Hz, 3H, 3-Me), 1.71 (s, 3H, 5-Me), 1.37 (s, 9H, ^tBu). ^{13}C NMR (126 MHz, 299 K, C_6D_6): δ 180.9 (C-1), 162.8 (C-5), 149.5 (C-3), 133.6 (C-2), 106.8 ($^1J_{\text{CH}} \approx 170.7$ Hz, Cp), 103.5 ($^1J_{\text{CH}} \approx 155.6$ Hz, C-4), 39.9 (^1tBu), 32.8 (^tBu), 27.0 (Me-5), 22.9 (Me-3).

8.4.2. Solid-State NMR Studies

^{13}C cross-polarization solid-state NMR experiments were conducted on a Bruker DSX-400 MHz solidstate NMR spectrometer operating at a resonance frequency of 100.67 MHz. Spectra were recorded on samples within 4 mm rotors, spun at 10.0 kHz. The Hartmann-Hahn matching condition was set at a nutation frequency of 38.5 kHz. Typical contact times of 5.0 ms were used, and the ^1H radio frequency amplitude was subjected to a linear ramp corresponding to nutation frequencies from 58.8 to 29.4 kHz during that period. Typically 1024-3072 scans were acquired under TPPM-15 decoupling, using a relaxation delay of 5 s. To differentiate between different ^1H - ^{13}C dipolar interaction strengths, experiments using shorter contact times (down to 200 μs) were conducted. Shielding tensor components were extracted from the spinning sideband intensity profiles of spectra obtained at a rotation speed of 3.0 kHz, using the DMFIT software. Data are

8. *Crystal Packing Induced Bond Isomerization*

reported according to the convention

$$\Delta\sigma = \sigma_{zz} - \sigma_{iso}; \quad \eta = \frac{|\sigma_{xx} - \sigma_{yy}|}{|\sigma_{zz} - \sigma_{iso}|} \quad (8.1)$$

Part V.

Summary and Conclusions

In this thesis, various electronic structure methods for the description of noncovalent interactions in non-metallic systems have been developed and evaluated. The main focus was the description of solids in the context of organic crystal structure prediction. A number of low-cost quantum chemical methods were established. Specifically, small basis set DFT-D3-gCP/'small', HF-3c, and DFTB3-D3 show good accuracies at only medium computational cost (decreasing in the order given above). The methods have been examined on several test sets and compared to competing methods. The results for the two organic solid benchmark sets, X23 and ICE10, are summarized in Figure 8.4 for selected method combinations.^{41,42,57} The performance of dispersion corrected hybrid functionals

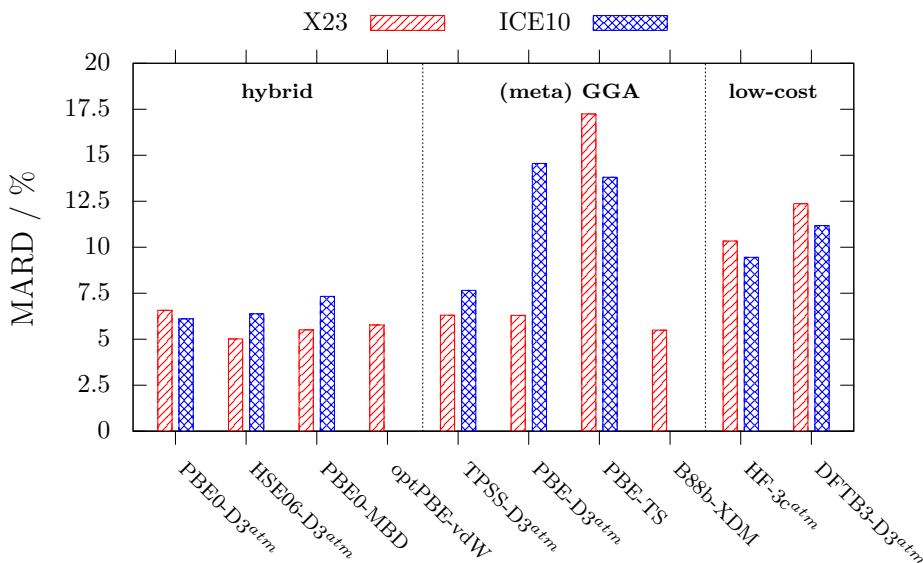


Figure 8.4.: Mean absolute relative deviations (MARD) of various dispersion corrected DFAs and alternative low-cost methods on the lattice energy benchmark sets X23 and ICE10. PBE0-MBD, optPBE-vdW, PBE-TS, and B88b-XDM values have been taken from Ref. 39,40.

is very good with mean absolute relative deviations (MARD) for lattice energies of 5–7%, which corresponds to absolute lattice energies close to (or within) the chemical accuracy of 1 kcal/mol. Note that the optPBE-vdW DFA has no exact exchange contribution but has larger computational costs compared to a GGA due to the nonlocal integral kernel. The PBE functional cannot be generally recommended. While PBE-D3 performs reasonable for vdW-bonded systems and for systems with weak hydrogen bonds, it strongly overestimates the lattice energies of the ice polymorphs. The Tkatchenko-Scheffler (TS) scheme is not recommended. The dispersion coefficients seem to be too large which leads to significantly overbound systems. This is compensated by larger many-body disper-

8. Final Summary and Conclusions

sion contributions in the many-body dispersion (MBD) correction with good performance comparable to D3^{atm}. We recommend the meta GGA functional TPSS-D3^{atm} for a generally good performance at reasonable computational costs. For large scale polymorph screening or frequency calculations, we recommend the low-cost alternatives HF-3c and DFTB3-D3. The MARDs on both test sets are slightly below and slightly above 10%, respectively.^{62,64} It has been shown that the Axilrod-Teller-Muto three-body dispersion contributes approximately 2–7% to the lattice energy and overall improves the results. This agrees with extensive studies by Tkatchenko and coworkers where, however, the MBD contribution is estimated to be significantly larger.

Additionally, it has been shown that the unit cell geometries of dispersion corrected DFAs deviate only by 1–3% from the references. The two main contributions arise from the intramolecular structure and the intermolecular bonding. Concerning the size of the molecules, the presented methodologies can be judged by their performance on a test set of rotational constants.¹²⁷ Specifically, the gas phase molecules at the TPSS-D3 level are too large by 1–2%. This is consistent with too large unit cell volumes in the solid state indicating that no additional systematic error is introduced. A direct comparisons to X-ray geometries have to be judged carefully because the free optimizations are conducted on the purely electronic energy surface. A possible way to exclude zero-point and thermal effects from the crystal geometry was presented for the ICE10 set of ten ice polymorphs.⁴² The rather costly potential energy scan with calculation of phonon spectra at each point, can be conducted ideally with the developed low-cost methods. Especially HF-3c can be recommended, while the tight binding model DFTB3 yields larger errors for dense systems.

The reliable description of organic solids was demonstrated in a number of case studies. In many computational studies, the molecular geometry is calculated as gas phase structure. This can lead to qualitatively wrong conclusions as demonstrated by the analyzed CPEs in the last part of this thesis. With the presented methods, all subtle details of the crystal packing could be calculated accurately. The molecular packing in the solid state can alter the molecular conformer, their relative orientation, the spin state, and even the bonding character. It was shown that the molecular conformer can change between the gas phase and the solid state for ethyl acetate.⁶⁵ A state-of-the-art treatment of London dispersion forces is crucial for accurate results. The deviation with respect to the W1 reference follows the Jacobs’ ladder classification of DFAs. A similar CPE alters the conformer of oxalic acid (part of the X23 set) during crystallization. When calculating the corresponding sublimation energy, the wrong gas phase conformer can significantly deteriorate the results.³⁵⁷ The next CPE introduced a subtle rotation between stacked tribenzotriquinacene molecules.⁶⁶ Interestingly, its centro-methyl derivative does

not has this feature and it could be assigned to the denser molecular packing enlarging the intermolecular London dispersion interactions. One dimensional packed molecules maintain the C_{3v} symmetry, the tilting only occurs in the three-dimensional crystal. A study not included in the thesis analyzed the spin state of iron spin crossover compounds.³⁵⁸ The spin state crucially depends on the molecular packing in the solid state. Although the theoretical investigations cannot predict the absolute spin state in advance (due to a strong dependence on the amount of exact Fock-exchange), (1) a certain theoretical model reproduces the experimental findings (B3LYP-D3) and (2) strong evidences were found that the stacking with strong noncovalent interactions stabilizes the low-spin state.

The last presented CPE study demonstrated that the combination of complementary information from X-ray crystallography, NMR chemical shift measurements, and theoretical investigations can be ideally used to quantitatively describe complex systems with unusual behavior.⁶⁷ Two prototypical zirconocene complexes have been analyzed, which differ by the substitution of a tert-butyl by a trimethylsilyl group. Due to the coordination to the metal center, a formally forbidden (η^2 -allenyl)/enamido-Zr to (η^2 -alkyne)/ κ Nimine-Zr bond isomerisation is feasible. The DFT-D3 solid and gas phase optimizations were guided by the measured X-ray structures. The bond isomerization due to the intermolecular London dispersion interaction in the solid state was proposed based on the theoretical investigation. Subsequently, isotropic and anisotropic ^{13}C NMR magnetic shieldings confirmed the proposed CPE. The assignment of the NMR chemical shifts had to be guided by corresponding calculations at the PBE0/TZVPP hybrid level with inclusions of scalar relativistic and spin-orbit coupling effects.

In the context of CSP, one can imagine a multilevel approach conducting a cascade of energy rankings with classical force fields, semiempirical MO methods, and dispersion corrected DFAs. In future work, some additional issues should be investigated. First, the DFT-D3 (hybrid) energy on GGA geometries seem to be insufficient to distinguish between very close lying polymorphs. A compound method based on a hybrid functional evaluated in a medium sized basis set in combination with a dispersion and a counterpoise correction could possibly improve the crystal geometries further. For the final lattice energy, a double-hybrid functional should yield consistently better results. If the corresponding DFA correlation energy is partially replace by correlation energy in the random-phase approximation, the computational demands should be feasible. The three-body dispersion within the D3 scheme is only the third order of the many-body expansion. Using an approach similar to the MBD scheme, but with the more accurate D3 dispersion coefficients could improve the dispersion contribution further. Second, the general applicability of typical empirical force fields used in the CSP task is limited. The quantum mechanically derived force field (QMDF¹⁹) shows promising accuracy for molecular

8. *Final Summary and Conclusions*

systems and should be extended to periodic boundaries. However, the current monopole treatment of the electrostatics could be insufficient for molecular crystals with complex hydrogen bonding networks. It should be replaced by higher multipole expansions. Once implemented (and re-parametrized), this possibly leads to a better screening of crystal polymorphs and subsequently to a more accurate treatment of the remaining structures in a certain energy window. Additionally, harmonic and anharmonic contributions to free energies could be routinely computed, which is prohibitive with the more expensive DFT methods.

The development and benchmarking of the various electronic structure approaches hopefully helps computational chemists and physicists to choose the best method for their desired application. Perhaps even new developments have been triggered by the extensive work on cost-effective schemes for molecular crystals. The answer to the previously raised question “Are crystal structures predictable?” has to be revised from Gavezottis “No” in 1994 to a “Maybe” in 2015. The future for low-cost quantum chemical methods to investigate noncovalent interactions in organic solids seems bright and they show great potential to be used in crystal structure predictions.

Bibliography

- [1] Li, Y.; Sonar, P.; Murphy, L.; Hong, W. *Energy Environ. Sci.* **2013**, *6*, 1684–1710.
- [2] Gélinas, S.; Rao, A.; Kumar, A.; Smith, S. L.; Chin, A. W.; Clark, J.; van der Poll, T. S.; Bazan, G. C.; Friend, R. H. *Science* **2014**, *343*, 512–516.
- [3] Park, N.-G. *J. Phys. Chem. Lett.* **2013**, *4*, 2423–2429.
- [4] Beckett, S. T. *Science of Chocolate*; RSC Paperbacks; The Royal Society of Chemistry, 2000.
- [5] Yu, L. *J. Am. Chem. Soc.* **2000**, *122*, 508.
- [6] Cruz-Cabeza, A. J.; Bernstein, J. *Chem. Rev.* **2014**, *114*, 2170–2191.
- [7] Nanda, K.; Beran, G. *J. Chem. Phys.* **2012**, *138*, 174106.
- [8] Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J. *Pharm. Res.* **2001**, *18*, 859–866.
- [9] Wilmer, C. E.; Snurr, R. Q. *Top Curr Chem* **2014**, *345*, 257–289.
- [10] Price, S. L. *Chem. Soc. Rev.* **2014**, *43*, 2098–2111.
- [11] Day, G. M. *Crystallogr. Rev.* **2011**, *17*, 3–52.
- [12] Gavezzotti, A. *Acc. Chem. Res.* **1994**, *27*, 309–314.
- [13] Vasileiadis, M.; Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. *Acta Cryst. B* **2012**, *68*, 677–685.
- [14] Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. *Phys. Rev. B* **2010**, *82*, 094116.
- [15] Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. *Comput. Phys. Commun.* **2012**, *183*, 2063–2070.
- [16] Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.

Bibliography

- [17] Neumann, M. A. *J. Phys. Chem. B* . **2008**, *112*, 9810–9829.
- [18] Day, G. M. et al. *Acta Cryst. B* **2005**, *61*, 511–527.
- [19] Grimme, S. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514.
- [20] Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1997.
- [21] Hohenstein, E. G.; Sherrill, C. D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.
- [22] Jansen, G. *WIREs Comput. Mol. Sci.* **2014**, *4*, 127–144.
- [23] Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.
- [24] Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315–1325.
- [25] Allen, M.; Tozer, D. J. *J. Chem. Phys.* **2002**, *117*, 11113.
- [26] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- [27] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.
- [28] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.
- [29] Tkatchenko, A.; DiStasio., R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.
- [30] Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.
- [31] Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023–5063.
- [32] Grimme, S. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.
- [33] Burns, L. A.; Vazquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *134*, 084107.
- [34] Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavez-zotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Cryst. B* **2000**, *56*, 697–714.
- [35] Motherwell, W. D. S. et al. *Acta Cryst. B* **2002**, *58*, 647–661.
- [36] Day, G. M. et al. *Acta Cryst. B* **2009**, *65*, 107–125.

- [37] Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. *Angew. Chem. Int. Ed.* **2008**, *47*, 2427–2430.
- [38] Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
- [39] Reilly, A. M.; Tkatchenko, A. *J. Chem. Phys.* **2013**, *139*, 024705.
- [40] Otero-de-la-Roza, A.; Johnson, E. R. *J. Chem. Phys.* **2012**, *137*, 054103.
- [41] Brandenburg, J. G.; Grimme, S. *Top Curr Chem* **2014**, *345*, 1–23.
- [42] Brandenburg, J. G.; Maas, T.; Grimme, S. *J. Chem. Phys.* **2015**, in press.
- [43] Bardwell, D. A. et al. *Acta Cryst. B* **2011**, *67*, 535–551.
- [44] Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013**, *139*, 134101.
- [45] Wen, S.; Nanda, K.; Huang, Y.; Beran, G. J. O. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578–7590.
- [46] Hirata, S.; Gilliard, K.; He, X.; Li, J.; Sode, O. *Acc. Chem. Res.* **2014**, *47*, 2721–2730.
- [47] Bygrave, P. J.; Allan, N. L.; Manby, F. R. *J. Chem. Phys.* **2012**, *137*, 164102.
- [48] Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K.-L. *Science* **2014**, *345*, 640–643.
- [49] Fang, T.; Li, W.; Gu, F.; Li, S. *J. Chem. Theory Comput.* **2015**, *11*, 91–98.
- [50] Macher, M.; Klimeš, J.; Franchini, C.; Kresse, G. *J. Chem. Phys.* **2014**, *140*, 084502.
- [51] Gillan, M. J.; Alfè, D.; Bygrave, P. J.; Taylor, C. R.; Manby, F. R. *J. Chem. Phys.* **2013**, *139*, 114101.
- [52] Ambrosetti, A.; Alfè, D.; DiStasio, R. A.; Tkatchenko, A. *J. Phys. Chem. Lett.* **2014**, *5*, 849–855.
- [53] Korth, M.; Lüchow, A.; Grimme, S. *J. Phys. Chem. A* **2008**, *112*, 2104–2109.
- [54] Al-Hamdani, Y. S.; Alfè, D.; von Lilienfeld, O. A.; Michaelides, A. *J. Chem. Phys.* **2014**, *141*, 18C530.

Bibliography

- [55] Benali, A.; Shulenburger, L.; Romero, N. A.; Kim, J.; von Lilienfeld, O. A. *J. Chem. Theory Comput.* **2014**, *10*, 3417–3422.
- [56] Kruse, H.; Grimme, S. *J. Chem. Phys.* **2012**, *136*, 154101.
- [57] Brandenburg, J. G.; Alessio, M.; Civalleri, B.; Peintinger, M. F.; Bredow, T.; Grimme, S. *J. Phys. Chem. A* **2013**, *117*, 9282–9292.
- [58] Thiel, W. *WIREs Comput Mol Sci* **2014**, *4*, 145–157.
- [59] Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- [60] Gaus, M.; Cui, Q.; Elstner, M. *WIREs Comput. Mol. Sci.* **2014**, *4*, 49–61.
- [61] Seifert, G.; Joswig, J.-O. *WIREs Comput. Mol. Sci.* **2012**, *2*, 456–465.
- [62] Brandenburg, J. G.; Grimme, S. *J. Phys. Chem. Lett.* **2014**, *5*, 1785–1789.
- [63] Sure, R.; Grimme, S. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- [64] Brandenburg, J. G.; Hochheim, M.; Bredow, T.; Grimme, S. *J. Phys. Chem. Lett.* **2014**, *5*, 4275–4284.
- [65] Brandenburg, J. G.; Grimme, S. *Theor. Chem. Acc.* **2013**, *132*, 1399.
- [66] Brandenburg, J. G.; Grimme, S.; Jones, P. G.; Markopoulos, G.; Hopf, H.; Cyran-ski, M. K.; Kuck, D. *Chem. Eur. J.* **2013**, *19*, 9930–9938.
- [67] Brandenburg, J. G.; Bender, G.; Ren, J.; Hansen, A.; Grimme, S.; Eckert, H.; Daniliuc, C. G.; Kehr, G.; Erker, G. *Organometallics* **2014**, *33*, 5358–5364.
- [68] Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188–5192.
- [69] Oganov, A. R. *Modern Methods of Crystal Structure Prediction*; Wiley-VCH: Berlin, 2010.
- [70] Oganov, A. R.; Glass, C. W. *J. Chem. Phys.* **2006**, *124*, 244704.
- [71] Sanderson, K. *Nature* **2007**, *450*, 771–771.
- [72] Woodley, S. M.; Catlow, R. *Nature Materials* **2008**, *7*, 937–964.
- [73] Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.

- [74] Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: New York, 2001.
- [75] Dreizler, J.; Gross, E. K. U. *Density Functional Theory, An Approach to the Quantum Many-Body Problem*; Springer: Berlin, 1990.
- [76] Paverati, R.; Truhlar, D. G. *Phil. Trans. R. Soc. A* **2014**, *372*.
- [77] Pérez-Jordá, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134–137.
- [78] Kaplan, I. G. *Intermolecular Interactions*; J. Wiley & Sons: Chichester, 2006.
- [79] Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. *Org. Biomol. Chem.* **2007**, *5*, 741–758.
- [80] Klimes, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.
- [81] Gräfenstein, J.; Cremer, D. *J. Chem. Phys.* **2009**, *130*, 124105.
- [82] Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* **2009**, *22*, 1127–1135.
- [83] Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. *CrystEngComm* **2008**, *10*, 405–410.
- [84] Jacobsen, H.; Cavallo, L. *ChemPhysChem* **2012**, *13*, 562–569.
- [85] Eshuis, H.; Bates, J. E.; Furche, F. *Theor. Chem. Acc.* **2012**, *131*, 1084.
- [86] Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
- [87] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [88] Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.
- [89] Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 224106.
- [90] Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- [91] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.
- [92] Lee, K.; Murray, E. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.

Bibliography

- [93] Hujo, W.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 3866–3871.
- [94] Grimme, S.; Hujo, W.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4875–4883.
- [95] Hujo, W.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13942–13950.
- [96] Hujo, W.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 308–315.
- [97] Casimir, H. B. G.; Polder, D. *Phys. Rev.* **1948**, *73*, 360–372.
- [98] Grimme, S. *Chem. Eur. J.* **2012**, *18*, 9955–9964.
- [99] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.
- [100] Starkschall, G.; Gordon, R. G. *J. Chem. Phys.* **1972**, *56*, 2801.
- [101] Moellmann, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8500–8504.
- [102] Axilrod, B. M.; Teller, E. *J. Chem. Phys.* **1943**, *11*, 299–300.
- [103] Ehrlich, S.; Moellmann, J.; Grimme, S. *Acc. Chem. Res.* **2013**, *46*, 916–926.
- [104] Reckien, W.; Janetzko, F.; Peintinger, M. F.; Bredow, T. *J. Comput. Chem.* **2012**, *33*, 2023–2031.
- [105] Cohen, J. S.; Pack, R. T. *J. Chem. Phys.* **1974**, *61*, 2372–2382.
- [106] Gianturco, F. A.; Paesani, F.; Laranjeira, M. F.; Vassilenko, V.; Cunha, M. A. *J. Chem. Phys.* **1999**, *110*, 7832.
- [107] Ahlrichs, R.; Penco, R.; Scoles, G. *Chem. Phys.* **1977**, *19*, 119–130.
- [108] Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys.* **2001**, *115*, 8748.
- [109] Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515–524.
- [110] Sato, T.; Nakai, H. *J. Chem Phys.* **2009**, *131*, 224104.
- [111] Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2006**, *124*, 174104.
- [112] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007**, *127*, 124108.
- [113] Kim, H.-J.; Tkatchenko, A.; Cho, J.-H.; Scheffler, M. *Phys. Rev. B* **2012**, *85*, 041403.

- [114] Ruiz, V. G.; Liu, W.; Zojer, E.; Scheffler, M.; Tkatchenko, A. *Phys. Rev. Lett.* **2012**, *108*, 146103.
- [115] Johnson, E. R.; Otero-de-la-Roza, A. *J. Chem. Theory Comput.* **2012**, *8*, 5124–5131.
- [116] Otero-de-la-Roza, A.; Johnson, E. R.; Contreras-García, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12165–1217.
- [117] Reilly, A. M.; Tkatchenko, A. *J. Phys. Chem. Lett.* **2013**, *4*, 1028–1033.
- [118] Chickos, J. S. *Netsu Sokutei* **2003**, *30*, 116.
- [119] Kresse, G.; Furthmüller, J. *J. Comp. Mat. Sci.* **1996**, *6*, 15.
- [120] Bucko, T.; Hafner, J.; Lebegue, S.; Angyan, J. G. *J. Phys. Chem. A* **2010**, *114*, 11814–11824.
- [121] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868, erratum *Phys. Rev. Lett.* **78**, 1396 (1997).
- [122] Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.
- [123] Kresse, G.; Joubert, J. *Phys. Rev. B* **1999**, *59*, 1758.
- [124] Grimme, S. ANCOPT: Approximate Normal Coordinate Rational Function Optimization Program. Universität Bonn, 2013.
- [125] Otero-de-la-Roza, A.; Johnson, E. R. *J. Chem. Phys.* **2013**, *138*, 054103.
- [126] Johnson, E. R. 2013; private communication.
- [127] Grimme, S.; Steinmetz, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.
- [128] Hammer, B.; Hansen, L. B.; Norskov, J. K. *Phys. Rev. B* **1999**, *59*, 7413–7421.
- [129] Kannemann, F. O.; Becke, A. D. *J. Chem. Theory Comput.* **2010**, *6*, 1081–1088.
- [130] Otero-de-la-Roza, A.; Johnson, E. R. *J. Chem. Phys.* **2012**, *136*, 174109.
- [131] Zhang, Y.; Yang, W. *J. Chem. Phys.* **1998**, *109*, 2604–2608.
- [132] Gritsenko, O.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J. *J. Phys. Chem. A* **2000**, *104*, 8558–8565.
- [133] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

Bibliography

- [134] Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753–12762.
- [135] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [136] Dovesi, R.; Orlando, R.; Civalleri, B.; Roetti, C.; Saunders, V. R.; Zicovich-Wilson, C. M. *Z. Kristallogr.* **2005**, *220*, 571–573.
- [137] Dovesi, R.; Saunders, V. R.; Roetti, C.; Orlando, R.; Zicovich-Wilson, C. M.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N. M.; Bush, I. J.; D’Arco, P.; Llunell, M. CRYSTAL09 User’s Manual (University of Torino, Torino, 2009). 2009.
- [138] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [139] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [140] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- [141] Ruzsinszky, A.; Perdew, J. P. *Comput. Theor. Chem.* **2011**, *963*, 2–6.
- [142] Bartels-Rausch, T. et al. *Rev. Mod. Phys.* **2012**, *84*, 885–944.
- [143] Palmer, D. A.; Fernández-Prini, R.; Harvey, A. H. *Aqueous Systems at Elevated Temperatures and Pressure*; Academic Press: London, 2004.
- [144] Ben, M. D.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2013**, *9*, 2654.
- [145] Ben, M. D.; Schoenherr, M.; Hutter, J.; VandeVondele, J. *J. Phys. Chem. Lett.* **2013**, *4*, 3753.
- [146] Carrasco, J.; Hodgson, A.; Michaelides, A. *Nature Materials* **2012**, *11*, 667.
- [147] Chutia, S.; Rossi, M.; Blum, V. *J. Phys. Chem. B* **2012**, *116*, 14788.
- [148] DiStasio, R. A.; Santra, B.; Li, Z.; Wu, X.; Car, R. *J. Chem. Phys.* **2014**, *141*, 084502.
- [149] Santra, B.; Klimeš, J.; Tkatchenko, A.; Alfè, D.; Slater, B.; Michaelides, A.; Car, R.; Scheffler, M. *J. Chem. Phys.* **2013**, *139*, 154702.
- [150] Kambara, O.; Takahashi, K.; Hayashi, M.; Kuo, J.-L. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11484–11490.
- [151] Takamuku, T.; Saisho, K.; Nozawa, S.; Yamaguchi, T. *J. Mol. Liq.* **2005**, *119*, 122–146.

- [152] Hobbs, P. V. *Ice Physics*; Oxford University Press: New York, 1974.
- [153] Vega, C.; McBride, C.; Sanz, E.; Abascal, J. L. F. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1450–1456.
- [154] Fortes, A. D.; Wood, I. G.; Alfredsson, M.; Vocadlo, L.; Knight, K. S. *J. Appl. Cryst.* **2005**, *38*, 612–618.
- [155] Lobban, C.; Finney, J. L.; Kuhs, W. F. *J. Chem. Phys.* **2002**, *117*, 3928–3934.
- [156] Whalley, E. *J. Chem. Phys.* **1984**, *81*, 4087–4092.
- [157] Kamb, B. *Science* **1965**, *150*, 205–209.
- [158] Yoshimura, Y.; Stewart, S. T.; Somayazulu, M.; Mao, H.-k.; Hemley, R. J. *J. Chem. Phys.* **2006**, *124*, 024502.
- [159] Jorgensen, J. D.; Beyerlein, R. A.; Watanabe, N.; Worlton, T. G. *J. Chem. Phys.* **1984**, *81*, 3211–3214.
- [160] Londono, J. D.; Kuhs, W. F.; Finney, J. L. *J. Chem. Phys.* **1993**, *98*, 4878–4888.
- [161] La Placa, S. J.; Hamilton, W. C.; Kamb, B.; Prakash, A. *J. Chem. Phys.* **1973**, *58*, 567–580.
- [162] Salzmann, C. G.; Radaelli, P. G.; Hallbrucker, A.; Mayer, E.; Finney, J. L. *Science* **2006**, *311*, 1758–1761.
- [163] Salzmann, C. G.; Radaelli, P. G.; Mayer, E.; Finney, J. L. *Phys. Rev. Lett.* **2009**, *103*, 105701.
- [164] Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169.
- [165] Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890–890.
- [166] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [167] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- [168] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- [169] Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
- [170] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

Bibliography

- [171] Muto, Y. *Proc. Phys. Math. Soc. Jpn.* **1944**, *17*, 629.
- [172] Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; D’Arco, P.; Noël, Y.; Causà, M.; Rérat, M.; Kirtman, B. *Int. J. Quantum Chem.* **2014**, *114*, 1287–1317.
- [173] CRYSTAL14: R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N.M. Harrison, I.J. Bush, P. D’Arco, M. Llunel, M. Caus.
- [174] Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678–5684.
- [175] Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
- [176] Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.
- [177] Santra, B.; Klimeš, J. c. v.; Alfè, D.; Tkatchenko, A.; Slater, B.; Michaelides, A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2011**, *107*, 185701.
- [178] Murray, E. D.; Galli, G. *Phys. Rev. Lett.* **2012**, *108*, 105502.
- [179] Goerigk, L.; Kruse, H.; Grimme, S. *Chem. Phys. Chem* **2011**, *12*, 3421–3433.
- [180] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1998**, *80*, 891–891.
- [181] Moellmann, J.; Grimme, S. *J. Phys. Chem. C* **2014**, *118*, 7615–7621.
- [182] Del Ben, M.; Schönherr, M.; Hutter, J.; VandeVondele, J. *J. Phys. Chem. Lett.* **2013**, *4*, 3753–3759.
- [183] Forster-Tonigold, K.; Groß, A. *J. Chem. Phys.* **2014**, *141*, 064501.
- [184] Kruse, H.; Goerigk, L.; Grimme, S. *J. Org. Chem.* **2012**, *77*, 10824–10834.
- [185] Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.
- [186] Anacker, T.; Friedrich, J. *J. Comput. Chem.* **2014**, *35*, 634–643.
- [187] Lippert, G.; Hutter, J.; Parrinello, M. *Mol. Phys.* **1997**, *92*, 477–488.
- [188] Lippert, G.; Hutter, J.; Parrinello, M. *Theo. Chem. Acc.* **1999**, *103*, 124–140.
- [189] Pisani, C.; Dovesi, R.; Roetti, C. *Hartree-Fock Ab Initio Treatment of Crystalline Solids Lecture Notes in Chemistr Series*; Springer Verlag: Berlin, 1988; Vol. 48.

- [190] Dovesi, R.; Civalieri, B.; Orlando, R.; Roetti, C.; Saunders, V. *Ab Initio Quantum Simulation in Solid State Chemistry*; Wiley, 2005; Vol. 21.
- [191] Gutowski, M.; Chalasiński, G. *J. Chem. Phys.* **1993**, *98*, 5540.
- [192] Ferrero, M.; Civalieri, B.; Rerat, M.; Orlando, R.; Dovesi, R. *J. Chem. Phys.* **2009**, *132*, 214704.
- [193] Maschio, L.; Usvyat, D.; Schütz, M.; Civalieri, B. *J. Chem. Phys.* **2010**, *132*, 134706.
- [194] Maschio, L.; Usvyat, D.; Civalieri, B. *CrystEngComm* **2010**, *12*, 2429–2435.
- [195] Peintinger, M. F.; Oliveira, D. V.; Bredow, T. *J. Comput. Chem.* **2013**, *34*, 451–459.
- [196] Josaa, D.; Rodríguez-Oteroa, J.; Cabaleiro-Lagob, E. M.; neiroa, M. R.-P. *Chem. Phys. Lett.* **2013**, *557*, 170–175.
- [197] Chan, B.; Ball, G. E. *J. Chem. Theory Comput.* **2013**, *9*, 2199–2208.
- [198] Moellmann, J.; Ehrlich, S.; Tonner, R.; Grimme, S. *J. Phys.: Condens. Matter* **2012**, *24*, 424206.
- [199] J. C. Sancho-García and J. Aragó and E. Ortí and Y. Olivier, *J. Chem. Phys.* **2013**, *138*, 204304.
- [200] von Lilienfeld, O. A.; Tkatchenko, A. *J. Chem. Phys.* **2010**, *132*, 234109.
- [201] Risthaus, T.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 1580–1591.
- [202] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- [203] Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5235.
- [204] Tatewaki, H.; Huzinaga, S. *J. Comput. Chem.* **1980**, *1*, 205–228.
- [205] Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *Journal of Chemical Information and Modeling* **2007**, *47*, 1045–1052.
- [206] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2262.
- [207] Eckert, F.; Pulay, P.; Werner, H.-J. *J. Comput. Chem.* **1997**, *18*, 1473–1483.
- [208] Thakkar, A. J.; Koga, T.; Saito, M.; Hoffmeyer, R. E. *Int. J. Quantum Chem. Symp.* **1993**, *48*, 343.

Bibliography

- [209] Swaminathan, S.; Craven, B. M.; McMullan, R. K. *Acta Crystallogr. Sec. B* **1984**, *40*, 300–306.
- [210] Tkatchenko, A.; von Lilienfeld, O. A. *Phys. Rev. B* **2008**, *78*, 045116.
- [211] Boese, R.; Weiß, H.-C.; Bläser, D. *Angew. Chem.* **1999**, *111*, 1042–1045.
- [212] Hyla-Kryspin, I.; Haufe, G.; Grimme, S. *Chem. Eur. J.* **2004**, *10*, 3411–3422.
- [213] Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
- [214] Zhao, Y.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2813–2818.
- [215] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
- [216] Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145–4186.
- [217] Wu, J.; Pisula, W.; Müllen, K. *Chem. Rev.* **2007**, *107*, 718–747.
- [218] Girifalco, L. A.; Lad, R. A. *J. Chem. Phys.* **1956**, *25*, 693–698.
- [219] Trickey, S. B.; Müller-Plathe, F.; Diercksen, G. H. F. *Phys. Rev. B* **1992**, *45*, 4460–4468.
- [220] Benedict, L. X.; Chopra, N. G.; Cohen, M. L.; Zettl, A.; Louie, S. G.; Crespi, V. H. *Chem. Phys. Lett.* **1998**, *286*, 490–496.
- [221] Zacharia, R.; Ulbricht, H.; Hertel, T. *Phys. Rev. B* **2004**, *69*, 155406–155413.
- [222] Grimme, S.; Mück-Lichtenfeld, C.; Antony, J. *J. Phys. Chem. C* **2007**, *111*, 11199.
- [223] Dobson, J. F.; White, A.; Rubio, A. *Phys. Rev. Lett.* **2006**, *96*, 073201.
- [224] Spanu, L.; Sorella, S.; Galli, G. *Phys. Rev. Lett.* **2009**, *103*, 196401.
- [225] TURBOMOLE 6.4: R. Ahlrichs, M. K. Armbruster, M. Bär, H.-P. Baron, R. Bauernschmitt, N. Crawford, P. Deglmann, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, C. Hättig, A. Hellweg, H. Horn, C. Huber, U. Huniar, M. Kattannek, C. Kölmel, M. Kollwitz, K. May, P. Nava, C. Ochsenfeld, H. Öhm, H. Patzelt, D. Rappoport, O. Rubner, A. Schäfer, U. Schneider, M. Sierka, O. Treutler, B. Unterreiner, M. von Arnim, F. Weigend, P. Weis and H. Weiss. Universität Karlsruhe 2012. See also: <http://www.turbomole.com>.
- [226] Henderson, M. A. *Surface Science Reports* **2002**, *46*, 1–308.

- [227] Pangher, N.; Schmalz, A.; Haase, J. *Chem. Phys. Lett.* **1994**, *221*, 189–193.
- [228] Chichagov, A. V.; Varlamov, D. A.; Dilanyan, R. A.; Dokina, T. N.; Drozhzhina, N. A.; Samokhvalova, O. L.; Ushakovskaya, T. V. *Crystallography Reports* **2001**, *46*, 876–879.
- [229] Beran, G. J. O.; Wen, S.; Nand, K.; Huang, Y.; Heit, Y. *Top Curr Chem* **2014**, *345*, 59–93.
- [230] Pantelides, C. C.; Adjiman, C. S.; Kazantsev, A. V. *Top Curr Chem* 25–58.
- [231] Sure, R.; Antony, J.; Grimme, S. *J. Phys. Chem. B* **2014**, *118*, 3431–3440.
- [232] Beran, G. J. O.; Nanda, K. *J. Phys. Chem. Lett.* **2010**, *1*, 3480–3487.
- [233] Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.
- [234] Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149–5155.
- [235] van Gisbergen, S. J. A.; Snijders, J. G.; Baerends, E. J. *J. Chem. Phys.* **1995**, *103*, 9347–9354.
- [236] Stewart, J. J. P. *J. Mol. Mod.* **2007**, *13*, 1173.
- [237] Jurečka, P.; Šponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- [238] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.
- [239] Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.
- [240] Korth, M. *ChemPhysChem* **2011**, *12*, 3131–3142.
- [241] Korth, M. *J. Chem. Theory Comput.* **2010**, *6*, 3808–3816.
- [242] Grimme, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 6306–6312.
- [243] Eckert, F.; Pulay, P.; Werner, H.-J. *J. Comput. Chem.* **1997**, *18*, 1473–1483.
- [244] Steed, J. W.; Atwood, J. L. *Supramolecular Chemistry, 2nd Edition*; Wiley-VCH: Chichester, 2009.

Bibliography

- [245] Wang, W.; Domini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243.
- [246] Raha, K.; Merz, K. M. *J. Med. Chem.* **2005**, *48*, 4558–4575.
- [247] Bredow, T.; Jug, K. *Theor. Chem. Acc.* **2005**, *113*, 1–14.
- [248] Hennemann, M.; Clark, T. *J. Mol. Mod.* **2014**, *20*, 2331.
- [249] Dannenberg, J. *J. Mol. Struct. THEOCHEM* **1997**, *401*, 279–286.
- [250] Csonka, G. I.; Ángyán, J. G. *J. Mol. Struct. THEOCHEM* **1997**, *393*, 31–38.
- [251] Winget, P.; Selçuki, C.; Horn, A. H. C.; Martin, B.; Clark, T. *Theor. Chem. Acc.* **2003**, *110*, 254–266.
- [252] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141–151.
- [253] Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495–506.
- [254] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- [255] Ahlswede, B.; Jug, K. *J. Comput. Chem.* **1999**, *20*, 563–572.
- [256] Igel-Mann, G.; Stoll, H.; Preuss, H. *Mol. Phys.* **1988**, *65*, 1321.
- [257] Andrae, D.; Haeussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123–141.
- [258] Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, 129–135.
- [259] Ridley, J.; Zerner, M. *Theoret. Chim. Acta.* **1973**, *32*, 111–124.
- [260] Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907–4917.
- [261] Jug, K.; Geudtner, G.; Homann, T. *J. Comp. Chem.* **2000**, *21*, 974–987.
- [262] Scholten, M. Semiempirical Methods with Orthogonalization Corrections: The OM3 Method. PhD Thesis, Heinrich-Heine-Universität Düsseldorf, Germany, 2003.
- [263] Stewart, J. J. P. *J. Mol. Mod.* **2013**, *19*, 1–32.
- [264] Bredow, T.; Geudtner, G.; Jug, K. *J. Comput. Chem.* **2001**, *22*, 89–101.
- [265] DiStasio, R. A.; Gobre, V. V.; Tkatchenko, A. *J. Phys.: Condens. Matter* **2014**, *26*, 213202.

- [266] Dobson, J. F. *Int. J. Quant. Chem.* **2014**, *114*, 1157–1161.
- [267] Kennedy, M. R.; McDonald, A. R.; DePrince, A. E.; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *140*, 121104.
- [268] Hostaš, J.; Řezáč, J.; Hobza, P. *Chem. Phys. Lett.* **2013**, *568*, 161–166.
- [269] MOPAC2012, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, [HTTP://OpenMOPAC.net](http://OpenMOPAC.net) (2012).
- [270] Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.
- [271] Kozuch, S.; Martin, J. M. L. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.
- [272] Mu, X.; Wang, Q.; Wang, L.-P.; Fried, S. D.; Piquemal, J.-P.; Dalby, K. N.; Ren, P. *J. Phys. Chem. B* **2014**, *118*, 6456–6465.
- [273] Li, A.; Muddana, H. S.; Gilson, M. K. *J. Chem. Theory Comput.* **2014**, *10*, 1563–1575.
- [274] Boese, D.; Kirchner, M.; Echeverria, G. A.; Boese, R. *ChemPhysChem* **2013**, *14*, 799–804.
- [275] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- [276] See <http://www.thch.uni-bonn.de/tc/dftd3>.
- [277] Grimme, S. *Angew. Chem. Int. Ed.* **2006**, *45*, 4460.
- [278] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.
- [279] Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, *47*, 558–561.
- [280] G. Kresse and J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169–11186.
- [281] Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
- [282] Stull, D. R. *Ind. Eng. Chem.* **1947**, *39*, 517–540.
- [283] Woodward, R. B.; Fukunaga, T. *J. Am. Chem. Soc.* **1964**, *86*, 3162–3164.
- [284] Hopf, H. *Classics in Hydrocarbon Chemistry: Syntheses, Concepts, Perspectives*; Wiley-VCH: Weinheim, 2000.

Bibliography

- [285] Bertz, S. H.; Kourouklis, G. A.; Jayaraman, A.; Lannoye, G.; Cook, J. M. *Can. J. Chem.* **1993**, *71*, 352–357.
- [286] Carceller, E.; García, M. L.; Moyano, A.; Pericàs, M. A.; Serratosa, F. *Tetrahedron* **1986**, *42*, 1831–1839.
- [287] Verevkin, S. P.; Beckhaus, H. D.; Rüchardt, C.; Haag, R.; Kozhushkov, S. I.; Zy-wietz, T. K.; de Meijere, A.; Jiao, H.; von Ragué Schleyer, P. *J. Am. Chem. Soc.* **1998**, *120*, 11130–11135.
- [288] Haag, R.; Schröder, D.; Zy-wietz, T.; Jiao, H.; Schwarz, H.; von Ragué Schleyer, P.; de Meijere, A. *Angew. Chem. Int. Ed.* **1996**, *108*, 1217–1319.
- [289] Haag, R.; Ohlhorst, B.; Noltemeyer, M.; Fleischer, R.; Stahlke, D.; Schuster, A.; Kuck, D.; de Meijere, A. *J. Am. Chem. Soc.* **1995**, *117*, 10474–10485.
- [290] Stevens, E. D.; Kramer, J. D.; Paquette, L. A. *J. Org. Chem.* **1976**, *41*, 2266–2269.
- [291] Kuck, D. *Angew. Chem. Int. Ed.* **1984**, *23*, 508–509.
- [292] Kuck, D.; Lindenthal, T.; Schuster, A. *Chem. Ber.* **1992**, *125*, 1449–1460.
- [293] Kuck, D.; Schuster, A.; Ohlhorst, B.; Sinnwell, V.; de Meijere, A. *Angew. Chem. Int. Ed.* **1989**, *28*, 595–597.
- [294] Kuck, D.; Neumann, E.; Schuster, A. *Chem. Ber.* **1994**, *127*, 151–164.
- [295] Markopoulos, G.; Henneicke, L.; Shen, J.; Okamoto, Y.; Jones, P. G.; Hopf, H. *Angew. Chem. Int. Ed.* **2012**, *51*, 12884–12887.
- [296] Saravanakumar, R.; Markopoulos, G.; Bahrin, L. G.; Jones, P. G.; Hopf, H. *Synlett* **2013**, *24*, 453–456.
- [297] Kuck, D.; Schuster, A.; Krause, R. A.; Tellenbröker, J.; Exner, C. P.; nad H. Bögge, M. P.; Müller, A. *Tetrahedron* **2001**, *57*, 3587–3613.
- [298] Cyranski, M.; Kuck, D. **2004**, The first X-ray structural analysis of the parent hydrocarbon **2** was not published.
- [299] Kuck, D. *Chem. Rev.* **2006**, *106*, 4885–4925.
- [300] Kuck, D. *Pure Appl. Chem.* **2006**, *78*, 749–775.

- [301] Zhang, T. X.; Zhou, L.; Cao, X. P.; Kuck, D. *Chin. J. Org. Chem.* **2007**, *27*, 946–957.
- [302] Tellenbröcker, J.; Kuck, D. *Angew. Chem. Int. Ed.* **1999**, *38*, 919–922.
- [303] Bredenkötter, B.; Henne, S.; Volkmer, D. *Chem. Eur. J.* **2007**, *13*, 9931–9938.
- [304] Georghiou, P. E.; Dawe, L.; Tran, H. A.; Stübe, J.; Neumann, B.; Stammeler, H. G.; Kuck, D. *J. Org. Chem.* **2008**, *73*, 9040–9047.
- [305] Wang, T.; Li, Z. Y.; Xie, A. L.; Ya, X. J.; Cao, X. P.; Kuck, D. *J. Org. Chem.* **2011**, *76*, 3231–3238.
- [306] Henne, S.; Bredenkötter, B.; Baghi, A. A. D.; Schmid, R.; Volkmer, D. *Dalton Trans.* **2012**, *41*, 5995–6002.
- [307] Langhals, H.; Rauscher, M.; Strübe, J.; Kuck, D. *J. Org. Chem.* **2008**, *73*, 1113–1116.
- [308] Niu, W. X.; Yang, E. Q.; Shi, Z. F.; Cao, X. P.; Kuck, D. *J. Org. Chem.* **2012**, *77*, 1422–1434.
- [309] Wang, T.; Hou, Q. Q.; Teng, Q. F.; Yao, X. J.; Niu, W. X.; Cao, X. P.; Kuck, D. *Chem. Eur. J.* **2010**, *16*, 12412–12424.
- [310] Niu, W. X.; Wang, T.; Hou, Q. Q.; Li, Z. Y.; Cao, X. P.; Kuck, D. *J. Org. Chem.* **2010**, *75*, 6704–6707.
- [311] Wang, T.; Zhang, Y. F.; Hou, Q. Q.; Xu, W. R.; Cao, X. P.; Chow, H. F.; Kuck, D. *J. Org. Chem.* **2013**, *78*, 1062–1069.
- [312] Zhou, L.; Zhang, T. X.; Li, B. R.; Cao, X. P.; Kuck, D. *J. Org. Chem.* **2007**, *72*, 6382–6389.
- [313] Sakurai, H.; Daiko, T.; Sakane, H.; Amaya, T.; Hirao, T. *Am. Chem. Soc.* **2005**, *127*, 11580–11581.
- [314] Hirao, T.; Amaya, T. *Fragments of Fullerenes and Carbon Nanotubes*; Wiley-VCH: Hoboken, 2012.
- [315] Forkey, D. M.; Attar, S.; Noll, B. C.; Koerner, R.; Olmstead, M. M.; Balch, A. L. *J. Am. Chem. Soc.* **1997**, *119*, 5766–5767.

Bibliography

- [316] Petrukhina, M. A.; Andreini, K. W.; Peng, L.; Scott, L. T. *Angew. Chem. Int. Ed.* **2004**, *43*, 5477–5481.
- [317] Sygula, A.; Folsom, H. E.; Sygula, R.; Abdourazak, A. H.; Marcinow, Z.; Fronczek, F. R.; Rabideau, P. W. *J. Chem. Soc. Chem. Commun.* **1994**, 2571–2572.
- [318] Schomaker, V.; Trueblood, K. N. *Acta Cryst. B* **1968**, *24*, 63–76.
- [319] Stephenson, R. M.; Malanowski, S. *Handbook of the Thermodynamics of Organic Compounds*; Elsevier: New York, 1987.
- [320] Neumann, M. A.; Perrin, M. A. *J. Phys. Chem. B* **2005**, *109*, 15531–15541.
- [321] von Arnim, M.; Ahlrichs, R. *J. Chem. Phys.* **1999**, *111*, 9183–9190.
- [322] Krebs, F. C.; Larsen, P. S.; Larsen, J.; Jacobsen, C. S.; Boutton, C.; Thorup, N. *J. Am. Chem. Soc.* **1997**, *119*, 1208–1216.
- [323] Madsen, G. K. H.; Krebs, F. C.; Lebech, B.; Larsen, F. K. *Chem. Eur. J.* **2000**, *6*, 1797–1804;.
- [324] Curtin, D. Y.; Paul, I. C. *Chem. Rev.* **1981**, *81*, 525–541.
- [325] Amaya, T.; Seki, S.; Moriuchi, T.; Nakamoto, K.; Nakata, T.; Sakane, H.; Saeki, A.; Tagawa, S.; Hirao, T. *J. Am. Chem. Soc.* **2009**, *131*, 408–409.
- [326] Sheldrick, G. M. *Acta Crystallogr. B* **2008**, *64*, 112–122.
- [327] Ugolotti, J.; Dierker, G.; Kehr, G.; Fröhlich, R.; Grimme, S.; Erker, G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2622.
- [328] Ugolotti, J.; Kehr, G.; Fröhlich, R.; Grimme, S.; Erker, G. *J. Am. Chem. Soc.* **2009**, *131*, 1996.
- [329] Suzuki, N.; Hashizume, D.; Koshino, H.; Chihara, T. *Angew. Chem. Int. Ed.* **2008**, *47*, 5198.
- [330] Suzuki, N.; Shimura, T.; Sakaguchi, Y.; Masuyama, Y. *Pure Appl. Chem.* **2011**, *83*, 1781.
- [331] Bender, G.; Kehr, G.; Fröhlich, R.; Petersen, J. L.; Erker, G. *Chem. Sci.* **2012**, *3*, 3534.
- [332] Negishi, E.; Cederbaum, F. E.; Takahashi, T. *Tetrahedron Lett.* **1986**, *27*, 2829.

- [333] Bender, G.; Kehr, G.; Daniliuc, C. G.; Dao, Q. M.; Ehrlich, S.; Grimme, S.; Erker, G. *Chem. Commun.* **2012**, 48, 11085.
- [334] Bender, G.; Kehr, G.; Daniliuc, C. G.; Wibbeling, B.; Erker, G. *Dalton Trans.* **2013**, 42, 14673.
- [335] Bender, G.; Wiegand, T.; Eckert, H.; R., F.; Daniliuc, C. G.; Mück-Lichtenfeld, C.; Ndambuki, S.; Ziegler, T.; Kehr, G.; Erker, G. *Angew. Chem. Int. Ed.* **2012**, 51, 8846.
- [336] Erker, G.; Pfaff, R.; Kowalski, D.; Würthwein, E.-U.; Krüger, C.; Goddard, R. *J. Org. Chem.* **1993**, 58, 6771.
- [337] Erker, G.; Pfaff, R. *Organometallics* **1993**, 12, 1921.
- [338] Rosenthal, U.; Ohff, A.; Baumann, W.; Tillack, A.; Görls, H.; Burlakov, V. V.; Shur, V. B. *Z. Anorg. Allg. Chem.* **1995**, 621, 77.
- [339] Bach, M. A.; Beweries, T.; Burlakov, V. V.; Arndt, P.; Baumann, W.; Spannenberg, A.; Rosenthal, U. *Organometallics* **2007**, 26, 4592.
- [340] Moellmann, J.; Grimme, S. *Organometallics* **2013**, 37, 3784–3787.
- [341] Neese, F. *WIREs Comput. Mol. Sci.* **2012**, 2, 73.
- [342] Fonseca Guerra, C.; Snijders, J. C.; te Velde, G.; Baerends, E. J. *Theor. Chem. Acc.* **1998**, 99, 391.
- [343] te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, 22, 931.
- [344] E., V.; Baerends, E. J. *J. Comput. Chem.* **2003**, 24, 1142.
- [345] Van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1993**, 99, 4597.
- [346] Schreckenbach, G.; Ziegler, T. *J. Phys. Chem.* **1995**, 99, 606.
- [347] Wolff, S. K.; Ziegler, T.; van Lenthe, E.; Baerends, E. J. *J. Chem. Phys.* **1999**, 110, 7689.
- [348] Woodcock, H. L.; Schaefer, H. F.; Schreiner, P. R. *J. Phys. Chem. A* **2002**, 106, 11923.

Bibliography

- [349] Bradley, J. P.; Velaga, S. P.; Antzutkin, O. N.; Brown, S. P. *Cryst. Growth Des.* **2011**, *11*, 3463.
- [350] Pascal, R. A.; Wang, C. M.; Wang, G. C.; Koplitz, L. V. *Cryst. Growth Des.* **2012**, *12*, 4367.
- [351] Seth, P.; Bauza, A.; Frontera, A.; Massera, C.; Gamez, P.; Ghosh, A. *CrystEngComm* **2013**, *15*, 3031.
- [352] Zhang, L.; Fonari, A.; Zhang, Y.; Zhao, G.; Coropceanu, V.; Hu, W.; Parkin, S.; Bredas, J.-L.; Briseno, A. L. *Chem. Eur. J.* **2013**, *19*, 17907.
- [353] Jean, Y.; Lledos, A.; Burdett, J. K.; Hoffmann, R. *J. Am. Chem. Soc.* **1988**, *110*, 4506.
- [354] Parkin, G. *Acc. Chem. Res.* **1992**, *25*, 455.
- [355] Labinger, J. A. *C. R. Chim.* **2002**, *5*, 235.
- [356] Rodriguez, A.; Olsen, R. A.; Ghaderi, N.; Scheschkewitz, D.; Tham, F. S.; Mueller, L. J.; Bertrand, G. *Angew. Chem. Int. Ed.* **2004**, *43*, 4880.
- [357] Carter, D. J.; Rohl, A. L. *J. Chem. Theory Comput.* **2014**, *10*, 3423–3437.
- [358] Schweinfurth, D.; Demeshko, S.; Hohloch, S.; Steinmetz, M.; Brandenburg, J. G.; Dechert, S.; Meyer, F.; Grimme, S.; Sarkar, B. *Inorg. Chem.* **2014**, *53*, 8203–8212.
- [359] Hansen, A.; Grimme, S. To be published.

Part VI.

Appendix

A. Supporting Information to Chapter 2

Appendix A contains:

- Explicit k -point grid
- Lattice energies
- Unit cell parameters
- Unit cell volumes and statistics

K-Point Grid

Table A.1.: Number of k -points in each direction for all 10 polymorphs. The k -grid is constructed via a Γ -centered Monhorst-Pack generation.

No.	polymorph	$\#k_x$	$\#k_y$	$\#k_z$
1	Ih	5	5	5
2	II	5	5	5
3	III	6	6	6
4	VI	5	5	5
5	VII	5	5	5
6	VIII	8	8	6
7	IX	6	6	6
8	XII	4	5	4
9	XIV	5	5	10
10	XV	6	6	7

Interaction Energies

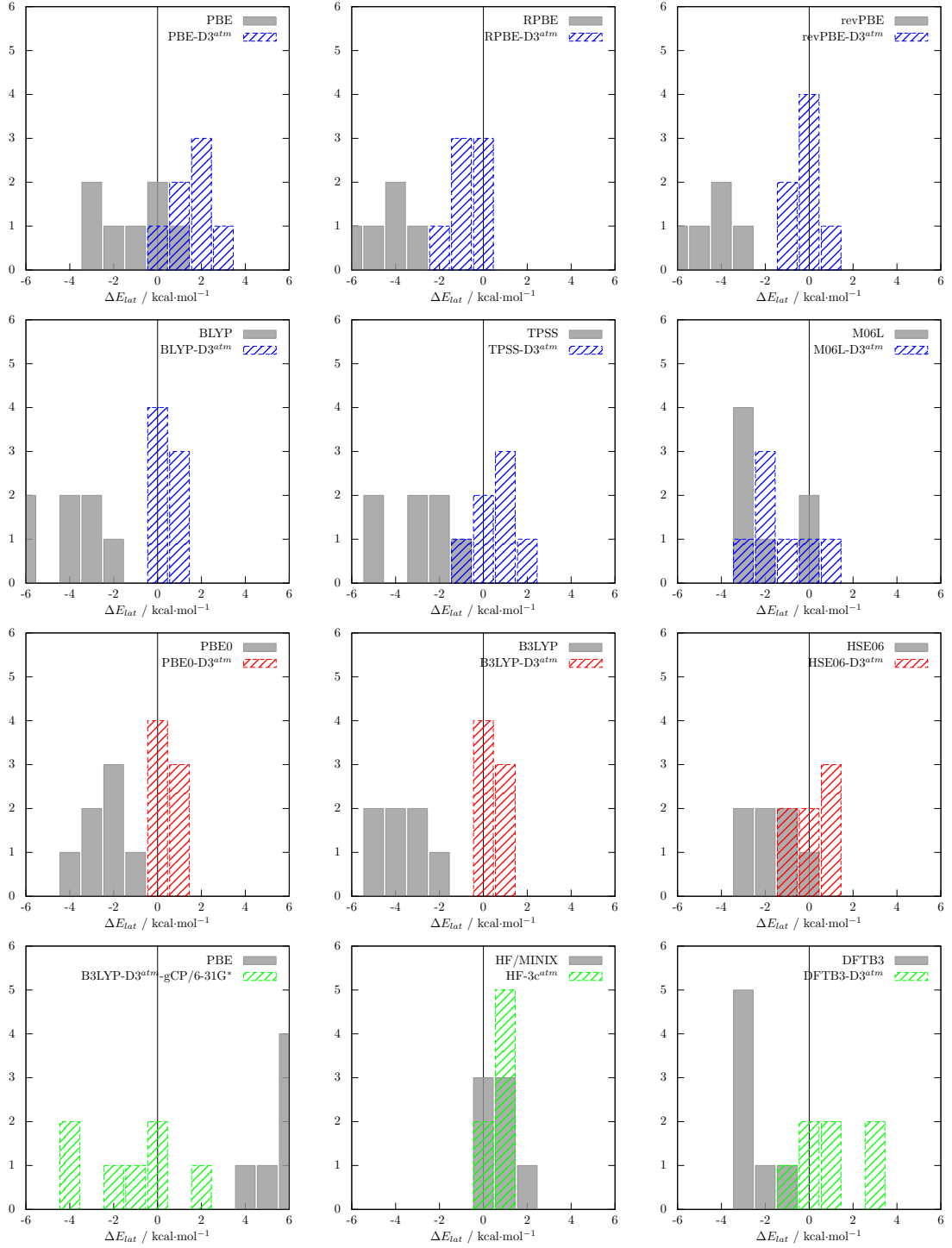


Figure A.1.: Deviation of selected methods from the reference lattice energy. Histograms are given with 1 kcal/mol bin width. The color code distinguishes GGA functionals (blue), hybrid functionals (red), and low-cost methods (green).

A. Supporting Information to Chapter 2

Table A.2.: Comparison of calculated lattice energies of the ICE10 benchmark set to the experimental reference. Energies, mean deviations (MD), and mean absolute deviations (MAD) are given in kcal/mol.

	1	2	3	4	5	6	7	8	9	10	MD	MAD
Ref.	14.07	14.05	13.85	13.68	13.07	13.31	13.97	–	–	–	–	–
DFT												
PBE	15.33	13.50	14.10	12.62	10.90	10.90	14.21	13.31	13.98	12.55	-0.6	1.1
RPBE	11.28	9.43	10.16	8.56	6.84	6.84	10.26	9.29	8.95	8.52	-4.7	4.7
revPBE	11.51	9.46	10.28	8.50	6.66	6.65	10.39	9.30	8.90	8.46	-4.7	4.7
BLYP	12.80	10.81	11.52	9.85	8.04	8.04	11.53	10.62	10.26	9.80	-3.3	3.3
TPSS	14.02	11.88	12.64	10.88	9.05	9.06	12.80	11.67	11.28	10.82	-2.2	2.2
M06L	11.23	11.96	11.56	12.35	13.45	13.36	11.64	12.04	12.18	12.32	-1.5	1.6
DFT-D												
PBE-D3	17.36	16.27	16.57	15.73	14.40	14.40	16.69	16.19	15.98	15.64	2.2	2.2
RPBE-D3	15.14	14.14	14.43	13.66	12.54	12.54	14.52	13.96	13.86	13.59	0.1	0.5
revPBE-D3	15.23	14.36	14.60	13.95	12.83	12.82	14.71	14.30	14.13	13.88	0.4	0.6
BLYP-D3	15.92	15.18	15.31	14.77	13.67	13.67	15.46	15.14	14.97	14.69	1.1	1.1
TPSS-D3	16.44	15.22	15.60	14.65	13.45	13.44	15.73	15.14	14.90	14.57	1.2	1.2
M06L-D3(0)	12.01	12.86	12.45	13.34	14.59	14.45	12.53	12.98	13.14	13.31	-0.5	1.3
DFT-D^{atm}												
PBE-D3 ^{atm}	17.24	16.10	16.41	15.51	14.15	14.15	16.53	16.00	15.77	15.43	2.0	2.0
RPBE-D3 ^{atm}	15.04	13.98	14.28	13.47	12.30	12.30	14.37	13.77	13.68	13.40	0.0	0.6
revPBE-D3 ^{atm}	15.13	14.20	14.45	13.75	12.60	12.60	14.56	14.13	13.94	13.68	0.2	0.5
BLYP-D3 ^{atm}	15.81	15.01	15.15	14.57	13.43	13.43	15.30	14.95	14.78	14.49	1.0	1.0
TPSS-D3 ^{atm}	16.33	15.05	15.44	14.44	13.17	13.16	15.57	14.95	14.69	14.36	1.0	1.1
M06L-D3(0) ^{atm}	11.90	12.69	12.29	13.13	14.32	14.20	12.37	12.79	12.94	13.10	-0.7	1.3
DFT-D^{atm} Hybrid												
PBE0-D3 ^{atm}	15.75	14.87	15.02	14.41	13.35	13.34	15.20	14.80	14.59	14.33	0.8	0.8
B3LYP-D3 ^{atm}	15.46	14.87	14.87	14.51	13.57	13.57	15.09	14.83	14.66	14.43	0.8	0.8
HSE06-D3 ^{atm}	15.70	14.93	15.19	14.40	13.26	13.25	15.36	14.89	14.64	14.35	0.9	0.9
Low-Cost												
B3LYP/6-31G*	21.14	20.20	20.49	19.71	18.13	18.13	20.50	18.28	19.90	19.56	6.0	6.0
HF-3c ^{atm}	16.04	15.26	15.32	14.61	14.16	14.16	15.56	15.15	14.77	14.63	1.3	1.3
DFTB3-D3 ^{atm}	13.41	14.35	14.94	14.72	16.67	16.67	14.41	14.73	14.53	14.91	1.3	1.5

Unit Cell Parameters

Table A.3.: Unit cell parameters of the ten ice polymorphs at various theoretical levels. The reference is estimated from the X-ray structures with isotropic zero-point and thermal corrections at the HF-3c level. Length are given in Å, angle in deg, and volumes in Å³ per molecule.

system		Ref.	PBE	RPBE	revPBE	BLYP	TPSS
Ih (hexagonal)	a	8.886	8.818	9.141	9.087	9.022	8.852
	c	7.233	7.182	7.461	7.410	7.335	7.206
	V	30.910	30.153	33.680	33.045	32.173	30.471

II (rhombohedral)	a	7.636	7.705	8.137	8.047	7.883	7.731
	α	113.1	112.346	111.990	112.051	112.264	112.367
	V	23.971	24.463	28.950	28.009	26.290	24.772
III (tetragonal)	a	6.631	6.630	6.919	6.874	6.774	6.657
	c	6.903	7.205	7.642	7.479	7.582	7.237
	V	25.291	26.402	30.487	29.455	28.997	26.733
VI (tetragonal)	a	6.198	6.226	6.638	6.548	6.380	6.244
	c	5.723	5.740	6.102	6.031	5.879	5.765
	V	21.983	22.252	26.883	25.856	23.931	22.477
VII (cubic)	a	6.757	6.980	7.646	7.471	7.176	6.925
	V	19.285	20.384	26.189	24.630	22.022	19.978
VIII (tetragonal)	a	4.715	4.833	5.243	5.123	4.950	4.799
	c	6.857	6.972	7.636	7.453	7.173	6.927
	V	19.059	20.360	26.237	24.453	21.946	19.940
IX (tetragonal)	a	6.662	6.572	6.868	6.815	6.745	6.600
	c	6.733	7.277	7.686	7.607	7.207	7.379
	V	24.899	26.184	30.215	29.434	27.320	26.785
XIII (monoclinic)	a	9.126	9.161	9.705	9.587	9.378	9.194
	b	7.379	7.467	7.891	7.809	7.644	7.508
	c	10.169	10.232	10.839	10.725	10.497	10.275
	β	109.687	109.468	109.354	109.472	109.497	109.200
	V	23.028	23.570	27.968	27.035	25.333	23.923
XIV (orthorhombic)	a	8.245	8.381	8.966	8.841	8.606	8.422
	b	8.037	8.172	8.668	8.581	8.364	8.222
	c	4.031	4.004	4.229	4.180	4.096	4.014
	V	22.263	22.851	27.392	26.425	24.571	23.163
XV (triclinic) (\approx tetragonal)	a	6.151	6.207	6.598	6.521	6.352	6.221
	b	6.162	6.212	6.611	6.532	6.363	6.227
	c	5.714	5.805	6.184	6.109	5.955	5.827
	α	90.06	90.110	89.763	89.880	90.037	90.091
	β	89.99	89.710	89.556	89.506	89.544	89.651
	γ	89.92	90.525	90.504	90.498	90.509	90.485
	V	21.657	22.383	26.970	26.022	24.069	22.573
system		Ref.	PBE-D3	RPBE-D3	revPBE-D3	BLYP-D3	TPSS-D3
Ih (hexagonal)	a	8.886	8.711	8.863	8.850	8.858	8.734
	c	7.233	7.099	7.216	7.198	7.193	7.113
	V	30.910	29.098	30.619	30.414	30.393	29.287
II (rhombohedral)	a	7.636	7.580	7.705	7.677	7.653	7.578
	α	113.1	112.456	112.539	112.544	112.486	112.479
	V	23.971	23.217	24.442	24.163	23.881	23.207
III (tetragonal)	a	6.631	6.550	6.665	6.640	6.644	6.563
	c	6.903	6.822	6.907	6.922	6.776	6.853
	V	25.291	24.380	25.551	25.434	24.926	24.590
VI (tetragonal)	a	6.198	6.105	6.206	6.186	6.158	6.094
	c	5.723	5.614	5.707	5.682	5.652	5.612
	V	21.983	20.926	21.981	21.743	21.430	20.840
VII (cubic)	a	6.757	6.797	6.899	6.903	6.858	6.711
	V	19.285	18.894	19.678	19.733	19.365	18.219
VIII (tetragonal)	a	4.715	4.729	4.775	4.784	4.753	4.658
	c	6.857	6.816	6.887	6.892	6.859	6.705
	V	19.059	19.051	19.631	19.713	19.375	18.180
IX (tetragonal)	a	6.662	6.499	6.615	6.584	6.598	6.509
	c	6.733	6.832	6.973	7.011	6.751	6.945
	V	24.899	24.045	25.425	25.323	24.493	24.517

A. Supporting Information to Chapter 2

XIII (monoclinic)	a	9.126	8.996	8.996	9.099	9.066	8.988
	b	7.379	7.318	7.318	7.414	7.378	7.321
	c	10.169	10.031	10.031	10.141	10.102	10.029
	β	109.687	109.464	109.464	109.237	109.323	109.339
	V	23.028	22.237	22.237	23.067	22.773	22.238
XIV (orthorhombic)	a	8.245	8.183	8.345	8.283	8.265	8.183
	b	8.037	7.973	8.117	8.079	8.040	7.982
	c	4.031	3.954	4.008	3.985	3.969	3.944
	V	22.263	21.499	22.624	22.223	21.977	21.466
XV (triclinic) (\approx tetragonal)	a	6.151	6.089	6.185	6.162	6.131	6.077
	b	6.162	6.091	6.188	6.163	6.137	6.080
	c	5.714	5.679	5.777	5.751	5.721	5.673
	α	90.06	90.168	90.110	90.161	90.144	90.175
	β	89.99	89.762	89.725	89.678	89.711	89.712
	γ	89.92	90.457	90.558	90.574	90.520	90.498
	V	21.657	21.059	22.107	21.840	21.523	20.960
system	Ref.	PBE-D3 ^{atm}	RPBE-D3 ^{atm}	revPBE-D3 ^{atm}	BLYP-D3 ^{atm}	TPSS-D3 ^{atm}	
Ih (hexagonal)	a	8.886	8.719	8.871	8.858	8.866	8.742
	c	7.233	7.105	7.223	7.205	7.200	7.120
	V	30.910	29.179	30.704	30.499	30.478	29.369
II (rhombohedral)	a	7.636	7.591	7.715	7.688	7.664	7.589
	α	113.1	112.456	112.539	112.544	112.486	112.479
	V	23.971	23.315	24.545	24.265	23.981	23.305
III (tetragonal)	a	6.631	6.569	6.685	6.660	6.664	6.582
	c	6.903	6.842	6.928	6.943	6.796	6.873
	V	25.291	24.596	25.778	25.659	25.147	24.808
VI (tetragonal)	a	6.198	6.116	6.217	6.197	6.168	6.105
	c	5.723	5.624	5.717	5.692	5.662	5.622
	V	21.983	21.037	22.098	21.859	21.544	20.951
VII (cubic)	a	6.757	6.813	6.915	6.918	6.874	6.726
	V	19.285	19.024	19.813	19.869	19.498	18.344
VIII (tetragonal)	a	4.715	4.739	4.786	4.795	4.763	4.668
	c	6.857	6.832	6.903	6.908	6.875	6.720
	V	19.059	19.182	19.766	19.849	19.508	18.305
IX (tetragonal)	a	6.662	6.517	6.634	6.603	6.617	6.528
	c	6.733	6.852	6.993	7.031	6.771	6.965
	V	24.899	24.253	25.645	25.542	24.704	24.729
XIII (monoclinic)	a	9.126	9.011	9.011	9.115	9.082	9.004
	b	7.379	7.331	7.331	7.427	7.391	7.334
	c	10.169	10.049	10.049	10.159	10.120	10.046
	β	109.687	109.464	109.464	109.237	109.323	109.339
	V	23.028	22.354	22.354	23.190	22.894	22.356
XIV (orthorhombic)	a	8.245	8.197	8.360	8.297	8.279	8.198
	b	8.037	7.987	8.131	8.093	8.054	7.996
	c	4.031	3.961	4.015	3.992	3.976	3.951
	V	22.263	21.612	22.743	22.340	22.091	21.579
XV (triclinic) (\approx tetragonal)	a	6.151	6.100	6.196	6.173	6.142	6.088
	b	6.162	6.102	6.199	6.174	6.148	6.091
	c	5.714	5.689	5.787	5.762	5.731	5.683
	α	90.06	90.168	90.110	90.161	90.144	90.175
	β	89.99	89.762	89.725	89.678	89.711	89.712
	γ	89.92	90.457	90.558	90.574	90.520	90.498
	V	21.657	21.173	22.226	21.957	21.639	21.073
system	Ref.	B3LYP/6-31G*	HF-3c ^{atm}	DFTB3-D3 ^{atm}			

Ih (hexagonal)	a	8.886	8.810	8.846	9.104
	c	7.233	7.174	7.272	7.158
	V	30.910	30.115	30.146	30.061
II (rhombohedral)	a	7.636	7.595	7.476	7.569
	α	113.100	113.051	113.178	113.852
	V	23.971	23.670	23.635	21.784
III (tetragonal)	a	6.631	6.572	6.622	6.752
	c	6.903	6.902	7.239	6.305
	V	25.291	24.849	26.453	18.697
VI (tetragonal)	a	6.198	6.148	6.143	5.887
	c	5.723	5.638	5.577	5.463
	V	21.983	21.312	21.021	18.691
VII (cubic)	a	6.757	6.891	6.969	6.746
	V	19.285	19.085	17.971	15.488
VIII (tetragonal)	a	4.715	4.708	4.542	4.286
	c	6.857	6.891	6.969	6.745
	V	19.059	19.096	17.975	15.488
IX (tetragonal)	a	6.662	6.526	6.686	6.381
	c	6.733	6.969	6.697	6.048
	V	24.899	24.730	24.947	20.595
XIII (monoclinic)	a	9.126	9.966	9.150	9.246
	b	7.379	7.385	7.209	7.004
	c	10.169	10.556	10.070	9.250
	β	109.687	115.782	109.345	114.765
	V	23.028	24.939	22.383	19.427
XIV (orthorhombic)	a	8.245	8.313	8.121	7.214
	b	8.037	8.017	8.063	8.131
	c	4.031	3.942	3.970	3.962
	V	22.263	21.889	21.662	19.355
XV (triclinic) (\approx tetragonal)	a	6.151	6.139	6.071	5.862
	b	6.162	6.124	6.069	5.874
	c	5.714	5.693	5.730	5.614
	α	90.060	89.747	89.418	83.907
	β	89.990	89.653	89.128	78.830
	γ	89.920	90.378	89.982	80.180
	V	21.657	21.402	21.110	18.633

Table A.4.: Statistical deviations of unit cell lengths calculated at various methods. MD, SD, and MAD are given in Å and MRD, SRD, and MARD are given in %.

method	MD	SD	MAD	MRD	SRD	MARD
PBE	0.07	0.13	0.09	1.02	1.87	1.41
RPBE	0.49	0.21	0.49	7.26	3.09	7.26
revPBE	0.40	0.18	0.40	5.87	2.59	5.87
BLYP	0.23	0.13	0.23	3.44	1.88	3.44
TPSS	0.08	0.13	0.10	1.23	1.97	1.48
PBE-D3	-0.08	0.06	0.09	-1.10	0.88	1.28
RPBE-D3	0.02	0.07	0.05	0.31	1.02	0.77
revPBE-D3	0.01	0.07	0.04	0.23	1.05	0.68
BLYP-D3	-0.02	0.04	0.04	-0.30	0.70	0.59
TPSS-D3	-0.09	0.07	0.10	-1.30	1.02	1.51

A. Supporting Information to Chapter 2

PBE-D3 ^{atm}	-0.06	0.06	0.08	-0.91	0.90	1.15
RPBE-D3 ^{atm}	0.03	0.07	0.06	0.51	1.04	0.85
revPBE-D3 ^{atm}	0.03	0.07	0.05	0.43	1.07	0.75
BLYP-D3 ^{atm}	-0.01	0.04	0.04	-0.10	0.71	0.57
TPSS-D3 ^{atm}	-0.08	0.07	0.09	-1.10	1.03	1.33
B3LYP/6-31G*	0.01	0.19	0.10	0.01	2.22	1.32
HF-3c ^{atm}	-0.05	0.14	0.11	-0.81	2.10	1.63
DFTB3-D3 ^{atm}	-0.33	0.36	0.36	-4.87	5.00	5.32

Unit Cell Volumes

Table A.5.: Comparison of calculated unit cell volumes of the ICE10 benchmark set with the experimental reference. Volumes are given in Å³. Mean relative deviations (MRD) and mean absolute relative deviations (MARD) are given in %.

	1	2	3	4	5	6	7	8	9	10	MRD	MARD
Ref.	30.91	23.97	25.29	21.98	19.28	19.06	24.97	23.03	22.26	21.58	–	–
DFT												
PBE	30.15	24.46	26.40	22.25	20.38	20.36	26.18	23.57	22.85	22.38	3.1	3.6
RPBE	33.68	28.95	30.49	26.88	26.19	26.24	30.22	27.97	27.39	26.97	23.7	23.7
revPBE	33.05	28.01	29.76	25.87	24.63	24.45	29.78	27.10	26.43	26.02	19.1	19.1
BLYP	32.17	26.29	29.00	23.93	22.02	21.95	27.48	25.33	24.45	24.07	10.9	10.9
TPSS	30.47	24.77	26.73	22.48	19.98	19.94	26.97	23.92	23.16	22.57	3.9	4.1
DFT-D												
PBE-D3	29.10	23.22	24.38	20.93	18.89	19.05	24.05	22.24	21.50	21.06	-3.2	3.2
RPBE-D3	30.61	24.44	25.59	21.98	19.68	19.63	25.43	22.24	22.62	22.11	1.0	1.8
revPBE-D3	30.43	24.16	25.43	21.74	19.73	19.71	25.32	23.07	22.28	21.84	0.7	1.3
BLYP-D3	30.39	23.88	24.93	21.43	19.37	19.38	24.49	22.77	21.98	21.52	-0.8	1.3
TPSS-D3	29.29	23.21	24.59	20.84	18.22	18.18	24.52	22.24	21.47	20.96	-3.8	3.8
DFT-D^{atm}												
PBE-D3 ^{atm}	29.18	23.31	24.60	21.03	19.02	19.18	24.25	22.35	21.61	21.17	-2.7	2.8
RPBE-D3 ^{atm}	30.70	24.55	25.81	22.10	19.81	19.77	25.64	22.35	22.74	22.23	1.6	2.3
revPBE-D3 ^{atm}	30.51	24.26	25.66	21.86	19.86	19.84	25.54	23.19	22.40	21.96	1.3	1.7
BLYP-D3 ^{atm}	30.48	23.98	25.15	21.54	19.50	19.51	24.71	22.89	22.09	21.64	-0.3	1.0
TPSS-D3 ^{atm}	29.37	23.31	24.81	20.95	18.34	18.30	24.73	22.36	21.58	21.07	-3.2	3.2
Low-Cost												
B3LYP/6-31G*	30.11	23.67	24.85	21.31	19.08	19.10	24.73	22.86	21.89	21.40	-0.5	2.1
HF-3c ^{atm}	30.31	23.68	26.55	21.03	17.96	17.96	25.14	22.43	21.72	21.15	-2.2	3.3
DFTB3-D3 ^{atm}	30.06	21.78	18.70	18.70	15.49	15.49	20.59	19.43	19.36	18.63	-15.1	15.1

B. Supporting Information to Chapter 3

Appendix B contains:

- Sublimation energy contributions for the X23 test set
- Unit cells of the optimized X23 geometries

Table B.1.: Electronic sublimation energy contribution for the X23 test set on the PBE/CBS, PBE-D3/CBS, PBE/SVP, PBE-D3/SVP, PBE-D3-gCP/SVP, PBE-D3-gCP+ $E^{(3)}$, B3LYP-D3/SVP, B3LYP-D3-gCP/SVP, and B3LYP-D3-gCP/SVP+ $E^{(3)}$ level. All values in kcal/mol.

Func/basis:	PBE/CBS			PBE/SVP			B3LYP/CBS			Exp
Disp-Corr:	-	-D3	-D3	-	-D3	-D3	-D3	-D3	-D3	
BSSE-Corr:	-	-	-	-	-	-gCP	-gCP	-	-gCP	
3-body:	-	-	+ $E^{(3)}$	-	-	-	+ $E^{(3)}$	-	-	+ $E^{(3)}$
C1	6.9	21.6	20.4	16.8	31.8	20.8	19.6	33.2	22.3	21.2
C2	9.7	17.2	16.7	18.9	26.7	16.6	16.1	26.6	16.7	17.4
C3	-1.5	17.1	15.4	2.0	20.4	14.4	12.7	20.5	14.6	13.6
C4	6.6	10.3	10.1	13.1	16.5	13.0	12.8	15.3	11.8	11.6
C5	-1.4	25.4	23.1	3.4	31.5	24.3	22.0	37.8	30.6	28.3
C6	0.3	13.1	12.2	4.1	16.6	12.9	12.1	17.4	13.7	12.8
C7	1.9	6.0	5.8	5.7	9.9	2.6	2.4	11.0	3.7	3.5
C8	15.2	22.4	22.0	21.3	28.4	22.8	22.4	27.9	22.4	21.8
C9	21.1	37.8	36.5	34.2	50.5	37.3	36.1	51.8	38.9	37.6
C10	10.9	21.4	20.7	20.0	30.4	20.1	19.4	31.1	21.0	20.3
C11	13.4	20.1	19.7	20.7	27.4	18.9	18.4	27.6	19.1	18.7
C12	11.8	22.0	21.3	18.2	28.1	22.6	21.9	28.4	22.9	22.2
C13	-1.2	19.1	17.5	2.4	21.9	16.2	14.6	26.1	20.5	18.9
C14	13.1	23.0	22.2	24.4	34.4	15.3	14.6	41.3	21.0	20.3
C15	13.9	23.6	23.0	24.9	35.3	15.9	15.2	41.4	20.8	20.1
C16	3.3	15.7	14.9	8.9	20.9	15.9	15.1	21.9	17.0	16.2
C17	9.2	19.4	18.7	15.3	25.6	19.7	19.0	25.7	19.9	19.1
C18	3.2	14.6	13.7	9.6	21.1	15.0	14.3	23.3	17.3	16.5
C19	4.1	14.1	13.3	14.4	25.8	12.8	12.0	27.8	15.1	14.3
C20	18.9	33.5	32.4	29.5	44.8	30.8	29.7	47.0	33.6	32.5
C21	19.0	27.3	26.8	28.5	36.8	25.5	25.0	37.0	25.9	25.4
R22	3.0	20.7	19.1	11.8	30.6	22.1	19.9	32.9	24.6	22.9
R23	17.7	32.0	30.8	33.3	47.6	26.9	26.1	47.6	27.2	26.0

B. Supporting Information to Chapter 3

Table B.2.: Crystal structures of a subgroup of the X23 benchmark set. We calculate the geometries on the PBE-D3/CBS, PBE-D3/SVP, PBE-D3- $\frac{1}{2}$ gCP/SVP, PBE-D3-gCP/SVP, B3LYP-D3/SVP, and B3LYP-D3- $\frac{1}{2}$ gCP/SVP levels. Cell length are in Å, volumes in Å³, and cell angles in degree.

Func/basis:		PBE/CBS	PBE/SVP		B3LYP/CBS		Exp
Disp-Corr:		-D3	-D3	-D3	-D3	-D3	
BSSE-Corr:		-	-	- $\frac{1}{2}$ gCP	-gCP	-	- $\frac{1}{2}$ gCP
C2	<i>a</i>	12.12	13.04	13.20	13.25	13.01	13.15
	<i>b</i>	4.07	3.83	3.99	4.13	3.75	3.92
	<i>c</i>	6.09	5.64	5.82	5.93	5.60	5.76
	Vol	300.2	281.5	306.5	324.8	272.9	294.0
C3	<i>a</i>	6.62	6.58	6.71	6.83	6.51	6.64
	<i>c</i>	8.88	8.86	8.98	9.04	8.69	8.85
	Vol	389.4	383.3	403.7	422.3	368.1	393.1
C4	<i>a</i>	5.01	4.91	5.04	5.13	4.91	5.12
	Vol	126.0	118.3	128.2	135.1	118.6	125.9
C6	<i>a</i>	7.35	7.13	7.29	7.49	—	7.39
	<i>b</i>	9.39	9.09	9.23	9.43	—	9.42
	<i>c</i>	6.77	6.96	7.01	7.02	—	6.81
	Vol	467.5	451.3	471.4	324.8	—	474.1
C7	<i>a</i>	5.68	5.52	5.74	5.94	5.40	5.62
	Vol	183.3	168.1	189.5	209.6	157.4	177.9
C8	<i>a</i>	6.86	6.75	6.91	7.07	6.66	6.86
	<i>b</i>	6.69	6.50	6.61	6.76	6.44	6.63
	<i>c</i>	8.94	8.90	8.94	9.00	8.99	9.15
	Vol	410.4	390.4	408.3	430.5	396.0	415.7
C14	<i>a</i>	6.69	6.59	6.97	7.38	6.41	6.55
	<i>b</i>	7.64	7.45	7.64	7.71	7.50	7.84
	<i>c</i>	6.15	5.96	6.25	6.35	5.94	6.09
	Vol	314.6	292.6	332.9	361.4	285.2	312.6
C16	<i>a</i>	9.23	9.17	9.21	9.31	9.14	9.33
	<i>b</i>	5.80	5.64	5.70	5.84	5.53	5.85
	<i>c</i>	3.71	3.65	3.80	3.86	3.57	3.73
	Vol	198.4	188.7	199.6	209.8	180.8	203.6
C17	<i>a</i>	8.20	8.05	8.04	8.25	—	8.19
	<i>b</i>	12.16	12.22	12.46	12.68	—	12.59
	<i>c</i>	6.80	6.74	6.99	7.10	—	6.77
	Vol	702.5	662.9	700.3	742.2	—	698.3
C21	<i>a</i>	5.56	5.41	5.56	5.67	5.39	5.57
	<i>c</i>	4.67	4.66	4.70	4.78	4.64	4.68
	Vol	144.4	136.5	145.0	153.6	134.9	145.1

C. Supporting Information to Chapter 4

Appendix C contains:

- General DFTB3-D3 methodology
- Explicit binding energies for the S66 and X23 test set
- Extended statistic on various databases including linear correlation plots

We utilize the DFTB Hamiltonian with full third-order correction and self consistent charges (SCC). The SCC tolerance is 10^{-7} au. The divide-and-conquer algorithm is used for fast diagonalization and the DIIS to enhance the SCC convergence. In the periodic calculations, the Brillouin zone is sampled with a Γ centered grid with at least 0.05 \AA^{-1} k -points, generated via the Monkhorst-Pack scheme.⁶⁸ We use the most recent Slater-Koster files provided by the group of M. Elstner. The hydrogen containing pair potentials are additionally damped with an exponent of 4.2, which is the recommended value for proton transfers.^{59,174,175,233} This SCC-DFTB3 method is abbreviated DFTB in the following.

The atom-pairwise D3 correction solely uses the geometry information to calculate the dispersion energy

$$E_{disp} = -\frac{1}{2} \sum_{n=6,8} \sum_{i,j} s_n \frac{C_n^{ij}}{\|\mathbf{r}_{ij}\|^n + f(R_0^{ij})^n}, \quad (\text{C.1})$$

where $C_{6/8}^{ij}$ are the leading order dipole-dipole and dipole-quadrupole dispersion coefficients and r_{ij} is the distance between the atom pairs i, j .²⁶ The s_6 scaling coefficient is set to unity to ensure the correct long-range behavior. The rational damping function $f(R_0^{ij})$ is used to match the long- and medium-range dispersion contribution from D3 with the semilocal correlation captured by DFTB.⁸⁷ The C_6 dispersion coefficients depend geometrically on the molecular environment and are pre-calculated by time-dependent DFT and utilizing the Casimir-Polder relation.^{97,235} The London dispersion correction D3 is used in the Becke-Johnson damping variant with parameters $s_8 = 0.5883$, $a_1 = 0.5719$, and $a_2 = 3.6017$. The parameters were fitted on the S66 reference energies similar to the procedure in the original publication.²⁶ A similar DFTB-D3 parametrization was already tested on the S12L set and for calculations of electron impact mass spectra.^{98,242}

For the molecular benchmark sets S22, S66, S66x8, X40, L7, and S12L, single point energy calculations on the reference geometry are performed. This is the standard procedure, but it cannot be applied to experimental (X-ray) reference structures. While the unit cell is typically very accurate, the molecular geometries (especially C-H bond length) typically deviate significantly. Therefore, we optimize the X23 geometries with fixed unit cell with the approximate normal coordinate rational function optimizer ANCOPT¹²⁴ until the atomic forces are below 10^{-4} au. Phonon frequencies are calculated at the Γ -point in a supercell approach. We repeat the primitive unit cell to create a supercell with a minimum cell length of 9 to 10 \AA (in all directions). The vibrational corrections to the lattice enthalpy are calculated in the harmonic approximation similar to the reference approach.³⁹

$$\Delta H_{sub}(T) = -E_{lat} + \Delta E_{vib}(T) + 4RT \quad (\text{C.2})$$

$$E_{vib}(T) = \sum_q \left(\frac{\hbar\omega_q}{2} + \frac{\hbar\omega_q}{\exp\left(\frac{\hbar\omega_q}{k_B T}\right) - 1} \right) \quad (\text{C.3})$$

In the X40 test set, systems including Br or I are excluded, and the Fe-containing complex in the S12L set is also disregarded due to missing Slater-Koster files.

S66 Energies

S66 dissociation energies are compared to basis set extrapolated CCSD(T) reference energies. The DFTB3 and DFTB3-D3 energies are calculated on fixed reference geometries.

C. Supporting Information to Chapter 4

Table C.1.: DFTB and DFTB-D3 dissociation energies on S66 test set. All values are in kcal/mol.

No.	DFTB3	DFTB3-D3	Ref.
1	4.23	4.68	4.89
2	4.17	4.92	5.57
3	3.42	4.2	6.88
4	7.02	8.11	8.08
5	3.86	4.81	5.75
6	3.19	4.53	7.54
7	6.24	7.72	8.22
8	3.83	4.4	5.00
9	1.37	2.44	3.04
10	0.91	2.35	4.15
11	2.24	4.27	5.41
12	3.48	4.43	7.25
13	3.52	5.08	6.18
14	2.82	4.74	7.45
15	5.54	7.88	8.62
16	3.66	4.46	5.12
17	12.86	15.13	17.18
18	3.1	4.02	6.83
19	2.92	4.27	7.40
20	17.3	19.06	19.09
21	13.62	15.46	16.26
22	16.41	18.41	19.49
23	15.39	17.49	19.19
24	-0.59	3.68	2.74
25	-0.21	4.22	3.83
26	3.18	9.16	9.82
27	-0.38	3.97	3.37
28	0.89	6.12	5.71
29	1.88	7.02	6.81
30	-0.37	2.03	1.41
31	0.84	3.56	3.38
32	1.45	3.8	3.74
33	-0.26	2.18	1.86
34	0.43	4.82	3.77
35	0.39	3.44	2.61
36	0.33	2.51	1.77
37	0.55	3.52	2.41
38	0.25	3.65	3.00
39	-0.11	3.8	3.57
40	0.22	3.32	2.89
41	0.68	5.5	4.84
42	0.41	4.62	4.13
43	0.16	3.61	3.70
44	0.21	2.36	1.99
45	0.03	1.91	1.75
46	0.62	4.62	4.24
47	0.05	2.74	2.87
48	0.37	3.03	3.53
49	0.18	2.89	3.32
50	0.84	2.55	2.86
51	0.63	1.26	1.52

52	2.28	4.55	4.70
53	1.7	3.74	4.36
54	1.62	3.09	3.27
55	1.36	3.78	4.19
56	0.47	2.97	3.23
57	1.26	4.84	5.28
58	0.53	2.06	4.15
59	2.3	2.7	2.85
60	2.93	4.16	4.86
61	0.51	3.32	2.88
62	0.8	3.93	3.51
63	0.78	3.77	3.80
64	0.77	2.78	2.99
65	1.92	2.81	3.99
66	0.59	2.86	3.97

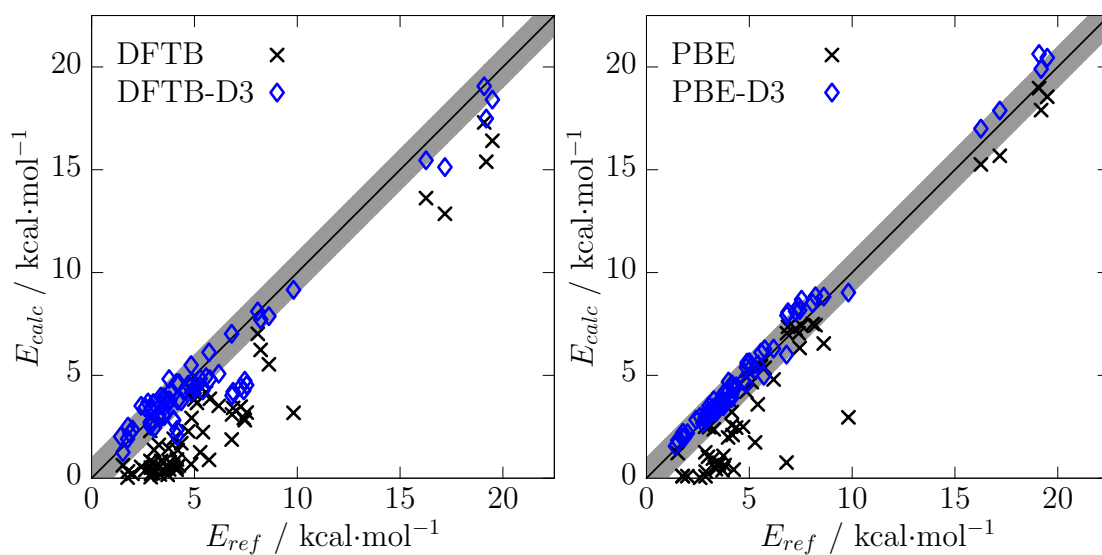


Figure C.1.: Correlation between the calculated DFTB, DFTB-D3, PBE, and PBE-D3 dissociation energies with the reference values. The gray shading highlights a ± 1 kcal/mol interval.

X23 Energies

X23 lattice energies are compared to thermodynamically back-corrected experimental sublimation energies. The DFTB and DFTB-D3 energies are calculated on optimized structures with fixed lattice constants. Additionally, we give thermodynamic contributions to sublimation energies compared to PBE-TS references.^{39,117}

Table C.2.: DFTB and DFTB-D3 lattice energies on X23 test set. Thermodynamic contributions are given in the harmonic approximation. All values are in kcal/mol.

No.	Name	Lattice Energies			Therm. Correction	
		DFTB	DFTB-D	Ref.	DFTB-D	Ref.
1	cyclohexanedione	5.71	24.07	21.18	-1.43	-1.79
2	acetic acid	10.73	18.59	17.40	-1.17	-1.48
3	adamantane	0.85	19.39	16.59	-1.62	-1.91
4	ammonia	4.09	7.70	8.89	-1.17	-1.77
5	anthracene	0.59	26.81	26.93	-1.32	-1.82
6	benzene	0.94	13.38	12.36	-1.28	-1.58
7	CO ₂	1.28	5.53	6.50	-1.52	-0.91
8	cyanamide	8.65	15.88	19.05	-1.53	-1.00
9	cytosine	18.69	35.16	40.58	-1.14	-1.53
10	ethylcarbanate	11.42	23.05	20.63	-1.51	-1.82
11	formamide	11.61	18.37	18.93	-1.77	-0.89
12	imidazole	0.05	25.02	20.60	-2.37	-1.73
13	hexamine	6.34	16.91	20.75	-1.42	-1.31
14	naphtalene	0.43	17.12	19.53	-1.85	-1.89
15	oxalic acid α	17.61	27.32	23.02	-1.58	-1.12
16	oxalic acid β	18.69	26.85	22.97	-1.87	-0.57
17	pyrazine	1.60	13.35	14.65	-2.09	-1.20
18	pyrazole	3.84	14.17	18.57	-1.31	-1.29
19	succinic acid	19.20	33.44	31.14	-1.96	-1.03
20	triazine	1.17	12.35	14.75	-1.64	-1.43
21	trioxane	3.76	15.59	15.87	-1.43	-1.98
22	uracil	19.21	35.53	32.43	-1.70	-1.55
23	urea	18.78	27.18	24.50	-1.87	-1.58

Extended Statistics

We give the mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) of the methods PBE-D3/def2-QZVP (1000 eV PAW for X23), DFTB, DFTB-D3, and PM7 on the benchmark sets S22, S66x8, X40, L7, S12L, and X23. The PBE-D3/1000 eV values for the X23 set are taken from reference⁴¹, PBE-D3/def2-QZVP for S12L from reference⁹⁸, PBE-D3/def2-QZVP for S22, S66x8 re-calculated but similar to reference¹⁷⁹. The density functional for the L7 set is TPSS-D3 with values from reference²³⁹. Due to inconsistencies in the original L7 references, we use new DLPNO-CCSD(T) energies which will be published elsewhere.³⁵⁹ PM7 is not conducted for the X23 set, because it is currently not directly applicable to molecular crystals. In total, more than 2000 data points are encoded.

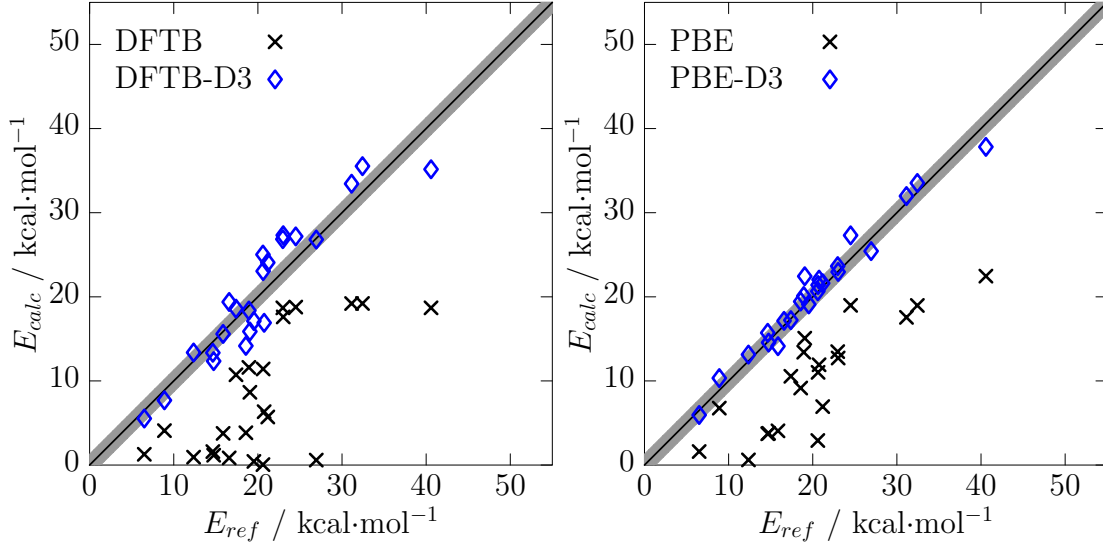


Figure C.2.: Correlation between the calculated DFTB, DFTB-D3, PBE, and PBE-D3 dissociation energies with the reference values. The gray shading highlights a ± 1 kcal/mol interval.

Table C.3.: Statistics for lattice and dissociation energies of several methods on the test sets S22, S66x8, X40, L7, S12L, and X23. We give mean absolute deviation (MAD), mean deviation (MD), and standard deviation (SD) in kcal/mol. Additionally the mean absolute relative deviations MARD in % are given.

Method	MAD	MD	SD	MARD
S22				
PBE-D3	0.58	0.11	0.79	10.1
DFTB	3.50	-3.50	4.23	65.3
DFTB-D3	0.95	-0.80	1.56	14.2
PM7	0.77	0.04	0.91	18.6
S66x8				
PBE-D3	0.35	0.24	0.48	14.4
DFTB	2.17	-2.17	2.54	74.5
DFTB-D3	0.79	-0.24	1.14	29.3
PM7	0.73	-0.13	0.96	35.6
X40				
PBE-D3	0.48	0.31	0.59	23.5
DFTB	2.12	-1.09	2.99	67.6
DFTB-D3	1.66	0.14	2.56	68.1
PM7	1.69	-1.01	3.22	52.6
L7				
PBE-D3	1.58	0.26	1.85	12.2
DFTB	14.15	-14.15	15.95	138.3
DFTB-D3	1.74	1.31	2.28	24.2
PM7	7.61	-7.61	8.33	92.7
S12L				
PBE-D3	2.01	1.21	2.50	6.2
DFTB	19.79	-19.79	22.11	75.9
DFTB-D3	5.90	4.60	7.99	15.4

C. Supporting Information to Chapter 4

PM7	17.51	-17.51	21.07	51.2
X23				
PBE-D3	1.07	0.43	1.34	5.7
DFTB	12.29	-12.29	13.58	64.0
DFTB-D3	2.48	-0.22	2.87	12.4
PM7	—	—	—	—

D. Supporting Information to Chapter 5

Appendix D contains:

- Parametrization of HF-3c(v) methodologies
- Individual energy contributions for all NCI test sets

Parametrization

The HF-3cv method is based on a Hartree-Fock with effective core potentials and minimal basis for the valence shells augmented with three correction schemes:

$$E_{\text{HF-3cv}} = E_{\text{HF/ECP-2G}} + E_{\text{D3}} + E_{\text{gCP}} + E_{\text{SRB}}. \quad (\text{D.1})$$

The standard Stuttgart-Dresden effective core potentials are used.^{256,257} The valence basis set consists of two Gaussian functions per shell for the elements H-Ar. For higher elements, the def2-SVP basis is used. The exponents of the ECP-2G basis are given in Table D.1.

Table D.1.: Exponents ζ and coefficients c of the ECP-2G basis set.

ζ	c	ζ	c	ζ	c	ζ	c
H (2s)/[1s]		He (2s)/[1s]		Li (2s1p)/[1s1p]		Be (2s1p)/[1s1p]	
2s							
1.44073	0.43013	4.09776	0.29152	0.05287	0.74709	0.11811	0.74709
0.25645	0.67891	0.53215	0.84993	0.02008	0.28560	0.04486	0.28560
1p							
				0.10000	1.00000	0.25000	1.00000
B (2s2p)/[1s1p]		C (2s2p)/[1s1p]		N (2s2p)/[1s1p]		O (2s2p)/[1s1p]	
2s							
0.21441	0.74709	0.33428	0.74709	0.35866	0.74709	0.48887	0.74709
0.08144	0.28560	0.12697	0.28560	0.13623	0.28560	0.18569	0.28560
2p							
0.63375	0.45226	1.06292	0.45226	1.58904	0.45226	2.14372	0.45226
0.15674	0.67131	0.26289	0.67131	0.39302	0.67131	0.53021	0.67131
F (2s2p)/[1s1p]		Ne (2s2p)/[1s1p]		Na (2s1p)/[1s1p]		Mg (2s1p)/[1s1p]	
2s							
0.63710	0.74709	0.80347	0.74709	0.46751	-0.15296	0.81367	-0.15296
0.24200	0.28560	0.30519	0.28560	0.04077	1.05137	0.07095	1.05137

D. Supporting Information to Chapter 5

2p				1p			
2.81156	0.45226	3.58451	0.45226	0.05000	1.00000	0.20000	1.00000
0.69539	0.67131	0.88656	0.67131				
Al (2s2p1d)/[1s1p1d]		Si (2s2p1d)/[1s1p1d]		P (2s2p1d)/[1s1p1d]		S (2s2p1d)/[1s1p1d]	
2s							
1.26075	-0.15296	1.78793	-0.15296	2.36742	-0.15296	3.01416	-0.15296
0.10993	1.05137	0.15590	1.05137	0.20643	1.05137	0.26283	1.05137
2s							
0.26798	0.53497	0.29761	0.53497	0.38696	0.53497	0.48705	0.53497
0.10407	0.52996	0.11557	0.52996	0.15027	0.52996	0.18914	0.52996
1d							
0.30000	1.00000	0.35000	1.00000	0.45000	1.00000	0.55000	1.00000
Cl (2s2p1d)/[1s1p1d]		Ar (2s2p1d)/[1s1p1d]					
2s							
3.71632	-0.15296	4.47522	-0.15296				
0.32406	1.05137	0.39023	1.05137				
2s							
0.60630	0.53497	0.74152	0.53497				
0.23544	0.52996	0.28795	0.52996				
1d							
0.65000	1.00000	0.70000	1.00000				

The short-range rational damping of the atom-pairwise D3 London dispersion correction

$$E_{disp} = -\frac{1}{2} \sum_{n=6,8} \sum_{i,j}^N s_n \frac{C_n^{ij}}{\|\mathbf{r}_{ij}\|^n + f(R_0^{ij})^n}, \quad (\text{D.2})$$

is fitted on the S66 reference energies similar to the original publication.^{26,87} R_0^{ij} are the pre-calculated cutoff radii, the dispersion coefficients C_n^{ij} are calculated for model systems utilizing a modified Casimir-Polder relation.⁹⁷ The individual parameters for the HF-3c, HF-3c, INDO-D3H+, NDDO-D3H+, DFTB3-D3, and OM2-D3 methods are summarized in Table D.2.

Table D.2.: D3 damping parameters for the utilized method combinations.

Method	s_6	s_8	a_1	a_2
HF-3cv	1.0000	0.5022	0.3063	3.9856
HF-3c	1.0000	0.8777	0.4171	2.9149
INDO-D3H+	1.0000	1.2377	0.3431	1.8463
NDDO-D3H+	1.0000	0.7505	0.3356	1.7474
DFTB3-D3	1.0000	0.5883	0.5719	3.6017
OM2-D3	1.0000	0.5310	0.6900	3.4460

The Hartree-Fock calculations in the (near) minimal basis sets ECP-2G and MINIX lead to basis set errors, which are corrected with semiempirical corrections. The geometrical counterpoise scheme gCP corrects the basis set superposition error (BSSE). Therein, we map the Boys-Bernardi counterpoise correction onto a semiempirical, repulsive pair

potential $V_{gCP}(r)$, which decays exponentially with the interatomic distance r

$$V_{gCP}(r) = e_A^{miss} \frac{\exp(-\alpha \cdot r^\beta)}{\sqrt{S \cdot N_B^{virt}}}. \quad (\text{D.3})$$

The pre-calculated parameters e_A^{miss} measure the basis incompleteness for each element and the four parameters α , β , η (averaging of s- and p-overlap), and σ (global scaling) are fitted onto the Boys-Bernardi CP energies of the S66x8 benchmark set.^{56,57} The individual parameters for the HF-3cv and HF-3c method are summarized in Table D.3

Table D.3.: gCP parameters for the utilized method combinations.

Method	σ	η	α	β
HF-3cv	0.1061	1.4239	0.8699	1.5000
HF-3c	0.1290	1.1526	1.1549	1.1763

The e^{miss} values and number of virtual basis functions N^{virt} for elements H-Kr are given in Table D.4. The parameters for the elements K-Kr are identical, because the same def2-SVP basis set is used.

Table D.4.: Pre-computed gCP values e^{miss} and N^{virt} (number of virtual basis functions) for the utilized method combinations.

Elem.	HF-3cv		HF-3c	
	e^{miss}	N^{virt}	e^{miss}	N^{virt}
H	0.06595	1	0.04240	1
He	0.11614	1	0.02832	1
Li	0.13832	5	0.17787	5
Be	0.15111	5	0.17160	5
B	0.21980	5	0.22424	5
C	0.39190	5	0.27995	5
N	0.59504	5	0.35791	5
O	0.99356	5	0.47901	5
F	1.52956	5	0.63852	5
Ne	2.21012	5	0.83235	5
Na	0.13227	9	1.11411	9
Mg	0.15747	9	1.27115	9
Al	0.15220	9	1.44695	14
Si	0.10318	9	1.61098	14
P	0.12986	9	1.76661	14
S	0.18404	9	1.98823	14
Cl	0.24730	9	2.22845	14

D. Supporting Information to Chapter 5

Ar	0.44966	9	2.48796	14
K	0.37425	11	0.37425	11
Ca	0.46097	11	0.46097	11
Sc	0.44489	21	0.44489	21
Ti	0.40499	21	0.40499	21
V	0.37841	21	0.37841	21
Cr	0.37344	21	0.37344	21
Mn	0.36125	21	0.36125	21
Fe	0.36001	21	0.36001	21
Co	0.36293	21	0.36293	21
Ni	0.24380	21	0.24380	21
Cu	0.40530	21	0.40530	21
Zn	0.39651	21	0.39651	21
Ga	0.35002	32	0.35002	32
Ge	0.34578	32	0.34578	32
As	0.34953	32	0.34953	32
Se	0.36731	32	0.36731	32
Br	0.38201	32	0.38201	32
Kr	0.39971	32	0.39971	32

Additional basis set errors, which mainly affects the covalent bond length are corrected via the short-ranged basis set correction SRB. The original SRB correction has the form

$$E_{\text{SRB1}} = -\frac{1}{2}s \sum_{i,j}^N (Z_i Z_j)^{3/2} \exp \left[-\gamma (R_0^{ij})^{3/4} r_{ij} \right], \quad (\text{D.4})$$

where Z_i are the nuclear charges, R_0 the D3 cutoff radii, and r_{ij} the atomic distances.⁶³ This correction is slightly modified for the HF-3cv method and given by

$$E_{\text{SRB2}} = -\frac{1}{4} \sum_{i,j}^N (s_i + s_j) \exp \left[-\gamma \left(\frac{r_{ij}}{R_0^{ij}} \right)^\kappa \right]. \quad (\text{D.5})$$

Here, element specific parameters are needed in addition to the two global parameters. All SRB parameters are given in Table D.5.

Table D.5.: SRB parameters for the HF-3c and HF-3cv methods (SRB1 and SRB2, respectively).

SRB1	
global	
$s = 0.03$	$\gamma = 0.70$
SRB2	
global	

$\kappa = 4$		$\gamma_{\text{H-Ne}} = 54.5$		$\gamma_{\text{Na-Ar}} = 42.5$		$\gamma_{\text{Li-Ar}} = 35.0$			
element specific s_i									
H	-0.0898	He	0.0400	Li	0.0400	Be	0.0400	B	0.1647
C	0.1568	N	0.2230	O	0.2273	F	0.1742	Ne	0.1981
Na	0.0109	Mg	0.0217	Al	0.0326	Si	0.0435	P	0.0544
S	0.0652	Cl	0.0761	Ar	0.0870	K	0.0109	Ca	0.0217
Sc	0.0326	Ti	0.0435	V	0.0544	Cr	0.0652	Mn	0.0761
Fe	0.0870	Co	0.0978	Ni	0.1087	Cu	0.1196	Zn	0.0217
Ga	0.0326	Ge	0.0435	As	0.0544	Se	0.0652	Br	0.0761
Kr	0.0870	Rb	0.0109	Sr	0.0217	Y	0.0326	Zr	0.0435
Nb	0.0544	Mo	0.0652	Tc	0.0761	Ru	0.0870	Rh	0.0978
Pd	0.1087	Ag	0.1196	Cd	0.0217	In	0.0326	Sn	0.0435
Sb	0.0544	Te	0.0652	I	0.0761	Xe	0.0870	Cs	0.0109
Ba	0.0217	La	0.0326	Ce	0.0435	Pr	0.0544	Nd	0.0652
Pm	0.0761	Sm	0.0870	Eu	0.0978	Gd	0.1087	Tb	0.1196
Dy	0.1304	Ho	0.1413	Er	0.1522	Tm	0.1631	Yb	0.1739
Lu	0.1848	Hf	0.0000	Ta	0.0544	W	0.0652	Re	0.0761
Os	0.0870	Ir	0.0978	Pt	0.1087	Au	0.1196	Hg	0.0217
Tl	0.0326	Pb	0.0435	Bi	0.0544	Po	0.0652	At	0.0761
Rn	0.0870								

S22 Energies

Table D.6.: HF/ECP-2G, HF-3cv, INDO, INDO-D3H+, NDDO und NDDO-D3H+ interaction energies on S22 test set. All values are in kcal/mol.

No.	Name	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	NDDO	NDDO-D3H+	Ref.
1	(NH_3) ₂ (C_{2h})	-2.63	-3.44	2.96	-0.07	2.25	-0.35	-3.17
2	(H_2O) ₂ (C_s)	-8.34	-7.96	-0.13	-5.28	-1.60	-6.50	-5.02
3	Formic acid dimer (C_{2h})	-24.42	-24.46	-1.31	-20.19	-1.29	-18.41	-18.61
4	Formamide dimer (C_{2h})	-17.99	-19.04	0.07	-14.49	-0.64	-13.77	-15.96
5	Uracil dimer (C_{2h})	-23.26	-24.79	-0.37	-17.74	-1.64	-17.41	-20.65
6	2-pyridoxine · 2-aminopyridine (C_1)	-19.72	-22.60	4.39	-13.67	2.95	-13.63	-16.71
7	Adenine · thymine WC (C_1)	-17.77	-20.75	3.16	-15.15	2.28	-14.54	-16.37
8	(CH_4) ₂ (D_{3d})	0.19	-0.60	0.77	-0.37	0.52	-0.47	-0.53
9	(C_2H_4) ₂ (D_{2d})	0.44	-1.74	3.49	-0.15	2.67	-0.43	-1.51
10	Benzene · CH_4 (C_3)	0.31	-1.81	1.33	-1.87	1.05	-1.69	-1.50
11	Benzene dimer (C_{2h})	2.2	-3.98	1.62	-6.41	1.01	-5.83	-2.73
12	Pyrazine dimer (C_s)	0.81	-5.58	1.86	-6.97	0.77	-6.70	-4.42
13	Uracil dimer (C_2)	-1.58	-9.19	1.58	-11.64	1.93	-9.33	-10.12
14	Indole · benzene (C_1)	2.29	-6.32	2.83	-8.64	1.58	-8.14	-5.22
15	Adenine · thymine stack (C_1)	-0.09	-11.62	3.43	-14.58	2.01	-13.21	-12.23
16	Ethene · ethine (C_{2v})	-0.34	-1.46	0.52	-1.17	0.45	-1.00	-1.69
17	Benzene · H_2O (C_s)	-1.44	-3.52	1.11	-2.90	0.98	-2.40	-3.28
18	Benzene · NH_3 (C_s)	-0.65	-2.82	1.39	-2.11	1.02	-1.96	-2.35
19	Benzene · HNC (C_s)	-0.62	-3.74	2.70	-2.83	2.04	-2.57	-4.46
20	Benzene dimer (C_{2v})	1.33	-2.49	2.86	-2.95	2.27	-2.68	-2.74
21	Indole · benzene T-shape (C_1)	-1.25	-6.24	5.08	-3.44	4.08	-3.11	-5.73
22	Phenol dimer (C_1)	-6.96	-10.56	2.45	-7.22	1.01	-7.65	-7.05

D. Supporting Information to Chapter 5

Table D.7.: HF/minix, HF-3c, DFTB3, DFTB3-D3, OM2, OM2-D3, PM6, PM6-D3H+ and PM7 interaction energies on S22 test set. All values are in kcal/mol.

No.	Name	HF/minix	HF-3c	DFTB3	DFTB3-D3	OM2	OM2-D3	PM6	PM6-D3H+	PM7
1	(NH ₃) ₂ (C _{2h})	-1,83	-2,64	-0,99	-1,64	-1,99	-2,51	-2,29	-3,00	-4,31
2	(H ₂ O) ₂ (C _s)	-6,30	-5,93	-4,24	-4,71	-4,16	-4,53	-3,87	-6,62	-4,88
3	Formic acid dimer (C _{2h})	-18,01	-18,05	-17,03	-18,68	-13,58	-14,84	-11,14	-16,88	-18,62
4	Formamide dimer (C _{2h})	-14,43	-15,47	-13,33	-15,07	-12,97	-14,35	-12,55	-17,32	-16,63
5	Uracil dimer (C _{2h})	-19,22	-20,75	-15,83	-18,15	-17,49	-19,39	-13,32	-19,19	-19,02
6	2-pyridoxine · 2-aminopyridine	-15,19	-18,07	-9,88	-12,65	-11,02	-13,31	-9,98	-16,66	-17,92
7	Adenine · thymine WC (C ₁)	-14,29	-17,27	-8,75	-11,71	-11,06	-13,51	-9,04	-15,99	-17,23
8	(CH ₄) ₂ (D _{3d})	0,24	-0,56	-0,06	-0,68	0,08	-0,45	-0,06	-0,44	-0,35
9	(C ₂ H ₄) ₂ (D _{2d})	0,45	-1,73	-0,19	-1,56	-0,37	-1,49	-0,35	-1,47	-1,01
10	Benzene · CH ₄ (C ₃)	0,55	-1,57	-0,16	-1,66	-0,16	-1,42	-0,47	-1,65	-1,79
11	Benzene dimer (C _{2h})	2,72	-3,47	0,57	-3,74	1,14	-2,49	0,11	-2,82	-4,22
12	Pyrazine dimer (C _s)	1,75	-4,64	-0,09	-4,43	-0,92	-4,50	-1,77	-4,81	-5,67
13	Uracil dimer (C ₂)	-0,79	-8,40	-3,18	-9,16	-4,34	-9,29	-4,46	-9,15	-8,58
14	Indole · benzene (C ₁)	3,78	-4,82	0,69	-5,37	1,64	-3,42	0,13	-4,06	-5,85
15	Adenine · thymine stack (C ₁)	1,11	-10,43	-3,44	-11,81	-4,00	-10,89	-4,94	-11,12	-11,37
16	Ethene · ethine (C _{2v})	-0,51	-1,64	-0,62	-1,33	-1,06	-1,64	-0,57	-1,09	-0,99
17	Benzene · H ₂ O (C _s)	-1,44	-3,52	-1,62	-3,09	-2,34	-3,55	-2,25	-3,37	-2,76
18	Benzene · NH ₃ (C _s)	-0,38	-2,55	-0,69	-2,21	-1,27	-2,54	-1,56	-2,72	-2,94
19	Benzene · HCN (C _s)	-0,80	-3,92	-1,40	-3,25	-3,06	-4,53	-1,99	-3,44	-3,05
20	Benzene dimer (C _{2v})	1,00	-2,82	-0,02	-2,55	-0,68	-2,83	-0,73	-2,67	-3,32
21	Indole · benzene T-shape (C ₁)	-0,58	-5,57	-1,25	-4,67	-2,37	-5,26	-2,43	-4,97	-5,97
22	Phenol dimer (C ₁)	-4,37	-7,98	-3,29	-6,09	-3,93	-6,28	-3,27	-7,40	-6,33

S66 Energies

Table D.8.: HF/ECP-2G, HF-3cv, INDO, INDO-D3H+, NDDO und NDDO-D3H+ interaction energies on S66 test set. All values are in kcal/mol.

No.	Name	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	NDDO	NDDO-D3H+	Ref.
1	H ₂ O · H ₂ O	-7.47	-6.54	-0.75	-5.22	-2.22	-6.53	-4.92
2	H ₂ O · MeOH	-7.23	-6.43	0.51	-5.12	-0.87	-6.16	-5.59
3	H ₂ O · MeNH ₂	-8.41	-8.29	1.24	-4.86	-0.83	-6.54	-6.91
4	H ₂ O · peptide	-8.40	-7.58	-1.07	-8.24	2.70	-9.03	-8.10
5	MeOH · MeOH	-7.56	-7.17	1.38	-4.79	-0.21	-6.00	-5.76
6	MeOH · MeNH ₂	-8.95	-9.87	2.80	-4.92	0.32	-6.82	-7.55
7	MeOH · peptide	-8.31	-8.21	1.01	-7.11	-0.66	-7.85	-8.23
8	MeOH · H ₂ O	-7.66	-7.05	0.16	-4.69	1.52	-6.17	-5.01
9	MeNH ₂ · MeOH	-3.26	-3.93	1.31	-2.89	0.58	-3.48	-3.06
10	MeNH ₂ · MeNH ₂	-3.16	-4.82	4.15	-1.60	3.24	-1.99	-4.16
11	MeNH ₂ · peptide	-3.05	-5.42	2.76	-4.06	1.52	-4.55	-5.42
12	MeNH ₂ · H ₂ O	-6.56	-7.50	2.04	-5.02	-0.47	-6.92	-7.27
13	peptide · MeOH	-6.28	-6.96	1.58	-4.29	0.65	-4.55	-6.19
14	peptide · MeNH ₂	-7.52	-9.44	4.09	-4.07	3.11	-4.44	-7.45
15	peptide · peptide	-6.91	-8.34	1.76	-6.89	1.04	-6.75	-8.63
16	peptide · H ₂ O	-6.37	-6.39	-0.14	-4.13	-0.90	-4.57	-5.12
17	uracil · uracil (BP)	-19.78	-17.91	0.46	-15.91	-0.41	-15.35	-17.18
18	H ₂ O · pyridine	-6.67	-6.73	1.15	-5.58	-1.25	-7.50	-6.86
19	MeOH · pyridine	-5.82	-8.06	2.25	-6.03	-0.11	-7.70	-7.41

20	<i>AcOH</i> · <i>AcOH</i>	-24.03	-18.91	-2.56	-20.76	-2.96	-19.55	-19.09
21	<i>AcNH₂</i> · <i>AcNH₂</i>	-17.72	-16.15	-0.48	-15.08	-1.00	-14.35	-16.26
22	<i>AcOH</i> · uracil	-23.20	-19.78	-1.73	-19.35	-2.91	-18.95	-19.49
23	<i>AcNH₂</i> · uracil	-21.52	-19.48	-0.97	-17.16	-1.78	-16.53	-19.19
24	benzene · benzene (π - π)	1.72	-3.25	1.53	-6.27	0.86	-5.80	-2.82
25	pyridine · pyridine (π - π)	1.18	-4.33	1.77	-6.80	0.80	-6.49	-3.90
26	uracil · uracil (π - π)	-1.58	-8.91	1.58	-11.64	1.93	-9.33	-9.83
27	benzene · pyridine (π - π)	1.19	-3.83	1.67	-6.50	0.77	-6.18	-3.44
28	benzene · uracil (π - π)	1.50	-5.58	1.70	-9.01	1.11	-7.97	-5.71
29	pyridine · uracil (π - π)	-0.10	-6.22	1.23	-9.27	0.17	-8.71	-6.82
30	benzene · ethene	1.11	-1.89	0.93	-3.55	0.47	-3.36	-1.43
31	uracil · ethene	0.34	-2.97	0.83	-4.82	0.42	-4.39	-3.38
32	uracil · ethyne	-0.40	-2.79	0.06	-4.99	-0.60	-4.86	-3.74
33	pyridine · ethene	0.58	-2.42	1.06	-3.62	0.41	-3.58	-1.87
34	pentane · pentane	1.00	-4.16	9.73	-0.09	7.61	-0.86	-3.78
35	neopentane · pentane	0.69	-2.99	6.67	0.02	5.38	-0.39	-2.61
36	neopentane · neopentane	0.60	-2.05	4.04	-0.35	3.21	-0.62	-1.78
37	cyclopentane · neopentane	-0.50	-3.03	6.49	0.60	5.23	0.13	-2.40
38	cyclopentane · cyclopentane	0.89	-3.06	8.61	0.83	6.83	0.12	-3.00
39	benzene · cyclopentane	0.91	-3.69	3.84	-4.16	2.78	-4.08	-3.58
40	benzene · neopentane	0.80	-3.46	2.28	-4.17	1.91	-3.69	-2.90
41	uracil · pentane	0.53	-5.13	4.08	-6.02	2.98	-5.68	-4.85
42	uracil · cyclopentane	0.74	-4.21	3.40	-5.35	2.48	-5.03	-4.14
43	uracil · neopentane	-0.07	-3.35	2.64	-3.77	1.66	-3.83	-3.71
44	ethene · pentane	0.61	-2.32	4.38	-0.60	3.55	-0.76	-2.01
45	ethyne · pentane	0.34	-1.77	1.55	-2.18	1.13	-2.07	-1.75
46	peptide · pentane	0.47	-4.01	5.61	-3.09	4.20	-3.26	-4.26
47	benzene · benzene (TS)	1.59	-2.82	2.81	-3.38	2.20	-3.07	-2.88
48	pyridine · pyridine (TS)	0.33	-3.17	2.68	-3.39	2.07	-3.09	-3.54
49	benzene · pyridine (TS)	0.41	-2.88	2.97	-3.04	2.14	-2.97	-3.33
50	benzene · ethyne (CH - π)	-0.87	-1.80	2.09	-2.07	1.63	-1.85	-2.87
51	ethyne · ethyne (TS)	-0.44	-1.14	0.40	-1.30	0.28	-1.15	-1.52
52	benzene · <i>AcOH</i> (OH - π)	-1.74	-4.31	2.20	-4.18	1.91	-3.44	-4.71
53	benzene · <i>AcNH₂</i> (NH - π)	-1.13	-3.27	1.99	-3.08	1.86	-2.42	-4.36
54	benzene · <i>H₂O</i> (OH - π)	-1.43	-3.24	1.12	-2.91	0.99	-2.40	-3.28
55	benzene · <i>MeOH</i> (OH - π)	-0.74	-3.97	2.57	-3.83	1.95	-3.45	-4.19
56	benzene · <i>MeNH₂</i> (NH - π)	-0.35	-3.26	2.63	-3.06	1.70	-3.13	-3.23
57	benzene · peptide (NH - π)	-0.41	-4.89	4.00	-4.31	3.10	-3.96	-5.28
58	pyridine · pyridine (NH - π)	-1.98	-3.76	3.34	-1.21	2.70	-1.13	-4.15
59	ethyne · <i>H₂O</i> (CH - O)	-1.85	-2.27	-0.42	-2.25	-0.96	-2.50	-2.85
60	ethyne · <i>AcOH</i> (OH - π)	-2.22	-3.51	0.80	-4.26	0.29	-3.88	-4.87
61	pentane · <i>AcOH</i>	0.15	-3.16	2.76	-3.20	2.10	-3.03	-2.91
62	pentane · <i>AcNH₂</i>	-0.01	-3.60	3.66	-3.06	2.84	-2.94	-3.53
63	benzene · <i>AcOH</i>	0.12	-3.26	1.65	-4.66	1.33	-4.03	-3.80
64	peptide · ethene	-0.31	-2.39	1.71	-2.88	1.25	-2.66	-3.00
65	pyridine · ethyne	-2.52	-4.11	0.95	-2.09	0.53	-2.01	-3.99
66	<i>MeNH₂</i> · pyridine	-1.59	-3.78	2.64	-3.15	1.82	-3.11	-3.97

Table D.9.: HF/minix, HF-3c, DFTB3, DFTB3-D3, OM2, OM2-D3, PM6, PM6-D3H+ and PM7 interaction energies on S66 test set. All values are in kcal/mol.

No.	Name	HF/minix	HF-3c	DFTB3	DFTB3-D3	OM2	OM2-D3	PM6	PM6-D3H+	PM7
-----	------	----------	-------	-------	----------	-----	--------	-----	----------	-----

D. Supporting Information to Chapter 5

1	$H_2O \cdot H_2O$	-6.05	-5.78	-4.23	-4.68	-3.93	-4.29	-3.81	-6.51	-4.84
2	$H_2O \cdot MeOH$	-5.77	-5.81	-4.17	-4.92	-3.91	-4.50	-4.15	-7.08	-4.95
3	$H_2O \cdot MeNH_2$	-7.18	-7.64	-3.42	-4.20	-4.45	-5.08	-4.07	-6.82	-6.84
4	$H_2O \cdot \text{peptide}$	-7.52	-7.93	-7.02	-8.11	-6.19	-7.05	-6.17	-8.61	-7.29
5	$MeOH \cdot MeOH$	-5.69	-6.13	-3.86	-4.81	-3.22	-4.01	-3.50	-6.66	-4.61
6	$MeOH \cdot MeNH_2$	-7.10	-8.58	-3.19	-4.53	-3.75	-4.86	-3.23	-6.49	-6.18
7	$MeOH \cdot \text{peptide}$	-6.93	-8.05	-6.24	-7.72	-4.83	-6.05	-4.93	-7.68	-6.61
8	$MeOH \cdot H_2O$	-5.90	-5.92	-3.83	-4.40	-3.26	-3.73	-3.21	-6.03	-4.32
9	$MeNH_2 \cdot MeOH$	-2.31	-3.26	-1.37	-2.44	-1.44	-2.35	-2.23	-4.72	-4.55
10	$MeNH_2 \cdot MeNH_2$	-2.12	-4.06	-0.91	-2.35	-1.51	-2.70	-1.91	-4.10	-5.38
11	$MeNH_2 \cdot \text{peptide}$	-2.66	-5.13	-2.24	-4.27	-2.76	-4.44	-3.86	-6.52	-6.25
12	$MeNH_2 \cdot H_2O$	-5.80	-7.12	-3.48	-4.43	-4.16	-4.91	-4.17	-6.86	-6.73
13	$\text{peptide} \cdot MeOH$	-4.99	-6.32	-3.52	-5.08	-3.85	-5.17	-4.24	-6.32	-6.51
14	$\text{peptide} \cdot MeNH_2$	-5.84	-8.23	-2.82	-4.74	-4.10	-5.72	-4.23	-7.55	-9.39
15	$\text{peptide} \cdot \text{peptide}$	-5.85	-8.26	-5.54	-7.88	-5.24	-7.22	-5.90	-9.34	-9.44
16	$\text{peptide} \cdot H_2O$	-5.22	-5.64	-3.66	-4.46	-3.70	-4.38	-3.83	-5.55	-6.11
17	uracil · uracil (BP)	-15.85	-17.28	-12.86	-15.13	-13.73	-15.60	-11.37	-17.30	-16.07
18	$H_2O \cdot \text{pyridine}$	-6.01	-6.75	-3.10	-4.02	-3.84	-4.60	-3.22	-6.16	-6.22
19	$MeOH \cdot \text{pyridine}$	-5.58	-7.33	-2.92	-4.27	-3.23	-4.35	-2.08	-5.41	-5.46
20	$AcOH \cdot AcOH$	-18.15	-18.29	-17.30	-19.06	-13.61	-14.99	-11.24	-17.44	-18.82
21	$AcNH_2 \cdot AcNH_2$	-14.38	-15.53	-13.62	-15.46	-12.80	-14.28	-12.37	-17.82	-16.67
22	$AcOH \cdot \text{uracil}$	-18.51	-19.33	-16.41	-18.41	-15.26	-16.86	-12.07	-18.08	-18.18
23	$AcNH_2 \cdot \text{uracil}$	-17.81	-19.17	-15.39	-17.49	-16.01	-17.71	-14.06	-19.69	-18.91
24	benzene · benzene (π - π)	2.51	-3.48	0.59	-3.68	1.36	-2.25	0.12	-2.77	-4.26
25	pyridine · pyridine (π - π)	1.93	-4.44	0.21	-4.22	-0.14	-3.84	-0.98	-4.00	-5.15
26	uracil · uracil (π - π)	-0.79	-8.40	-3.18	-9.16	-4.34	-9.29	-4.46	-9.15	-8.58
27	benzene · pyridine (π - π)	2.12	-4.00	0.38	-3.97	0.58	-3.07	-0.52	-3.49	-4.75
28	benzene · uracil (π - π)	1.90	-5.58	-0.89	-6.12	-1.15	-5.50	-1.51	-5.22	-5.39
29	pyridine · uracil (π - π)	0.73	-6.30	-1.88	-7.02	-2.35	-6.60	-3.13	-6.77	-6.68
30	benzene · ethene	1.48	-1.91	0.37	-2.03	1.07	-0.95	0.16	-1.46	-2.04
31	uracil · ethene	0.51	-3.08	-0.84	-3.56	-0.87	-3.14	-0.96	-2.90	-2.64
32	uracil · ethyne	-0.05	-3.15	-1.45	-3.80	-1.70	-3.63	-0.90	-2.57	-2.14
33	pyridine · ethene	1.11	-2.34	0.26	-2.18	0.50	-1.53	-0.26	-1.92	-2.27
34	pentane · pentane	1.40	-4.63	-0.43	-4.82	-0.37	-4.13	-0.78	-4.12	-4.19
35	neopentane · pentane	0.92	-3.19	-0.39	-3.44	-0.45	-3.12	-0.66	-3.04	-3.43
36	neopentane · neopentane	0.64	-2.20	-0.33	-2.51	-0.60	-2.55	-0.38	-2.07	-2.58
37	cyclopentane · neopentane	0.62	-3.19	-0.55	-3.52	-0.68	-3.28	-0.95	-3.21	-3.55
38	cyclopentane · cyclopentane	1.23	-3.43	-0.25	-3.65	-0.37	-3.30	-0.38	-2.94	-3.41
39	benzene · cyclopentane	1.38	-4.06	0.11	-3.80	0.33	-2.99	-0.55	-3.47	-4.23
40	benzene · neopentane	0.73	-3.55	-0.22	-3.32	0.10	-2.59	-0.64	-3.00	-3.84
41	uracil · pentane	1.21	-4.80	-0.68	-5.50	-0.59	-4.68	-1.76	-5.47	-5.06
42	uracil · cyclopentane	1.30	-3.96	-0.41	-4.62	-0.29	-3.87	-1.23	-4.45	-4.45
43	uracil · neopentane	0.77	-3.22	-0.16	-3.61	-0.77	-3.73	-0.98	-3.64	-3.51
44	ethene · pentane	0.53	-2.56	-0.21	-2.36	0.16	-1.67	-0.35	-1.98	-2.02
45	ethyne · pentane	0.68	-1.86	-0.03	-1.91	0.02	-1.57	-0.23	-1.56	-1.65
46	peptide · pentane	0.99	-4.14	-0.62	-4.62	-0.74	-4.13	-1.32	-4.46	-3.85
47	benzene · benzene (TS)	1.06	-3.08	-0.05	-2.74	-0.60	-2.88	-0.74	-2.81	-3.25
48	pyridine · pyridine (TS)	0.48	-3.42	-0.37	-3.03	-1.17	-3.39	-1.17	-3.25	-3.33
49	benzene · pyridine (TS)	0.62	-3.36	-0.18	-2.89	-1.03	-3.32	-1.15	-3.25	-3.53
50	benzene · ethyne (CH - π)	-0.08	-2.52	-0.84	-2.55	-1.84	-3.24	-1.08	-2.43	-2.12
51	ethyne · ethyne (TS)	-0.40	-1.36	-0.63	-1.26	-1.23	-1.73	-0.47	-0.95	-0.84
52	benzene · $AcOH$ (OH - π)	-1.43	-4.69	-2.28	-4.55	-2.59	-4.46	-2.58	-4.33	-3.93
53	benzene · $AcNH_2$ (NH - π)	-0.97	-3.68	-1.70	-3.74	-2.95	-4.65	-2.38	-3.97	-3.65
54	benzene · H_2O (OH - π)	-1.43	-3.52	-1.62	-3.09	-2.35	-3.55	-2.25	-3.37	-2.76
55	benzene · $MeOH$ (OH - π)	-0.66	-4.26	-1.36	-3.78	-1.77	-3.77	-1.90	-3.70	-3.33

56	benzene · <i>MeNH</i> ₂ (<i>NH</i> - π)	0.05	-3.50	-0.47	-2.97	-0.88	-2.98	-1,43	-3,28	-3.57
57	benzene · peptide (<i>NH</i> - π)	-0.24	-5.42	-1.26	-4.84	-1.86	-4.87	-2,27	-4,93	-5.37
58	pyridine · pyridine (<i>NH</i> - π)	-1.64	-3.85	-0.53	-2.06	-2.61	-3.89	-2,53	-3,77	-3.61
59	ethyne · <i>H</i> ₂ <i>O</i> (<i>CH</i> - <i>O</i>)	-1.58	-2.18	-2.30	-2.70	-2.87	-3.20	-1,84	-2,19	-1.57
60	ethyne · <i>AcOH</i> (<i>OH</i> - π)	-2.37	-4.26	-2.93	-4.16	-3.90	-4.87	-1,85	-2,84	-2.32
61	pentane · <i>AcOH</i>	0.46	-2.94	-0.51	-3.32	-0.68	-3.08	-1,21	-3,36	-2.70
62	pentane · <i>AcNH</i> ₂	0.37	-3.48	-0.80	-3.93	-0.76	-3.42	-1,43	-3,85	-3.40
63	benzene · <i>AcOH</i>	0.44	-3.68	-0.78	-3.77	-1.44	-3.94	-1,62	-3,88	-3.45
64	peptide · ethene	-0.02	-2.55	-0.77	-2.78	-1.06	-2.74	-1,07	-2,64	-2.20
65	pyridine · ethyne	-3.14	-4.66	-1.92	-2.81	-3.03	-3.77	-1,23	-1,85	-1.69
66	<i>MeNH</i> ₂ · pyridine	-0.72	-3.71	-0.59	-2.86	-1.16	-3.05	-1,42	-3,27	-4.38

X40 Energies

Table D.10.: HF/ECP-2G, HF-3cv, INDO, INDO-D3H+, NDDO und NDDO-D3H+ interaction energies on X40 test set. All values are in kcal/mol.

No.	Name	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	NDDO	NDDO-D3H+	Ref.
1	methane · <i>F</i> ₂	-0.19	-0.25	0.06	-0.73	0.04	-0.63	-0.49
2	methane · <i>Cl</i> ₂	0.09	-1.00	0.21	-2.05	0.59	-1.26	-1.08
3	methane · <i>Br</i> ₂	-0.50	-1.84	-1.38	-4.29	-0.69	-3.02	-1.30
4	methane · <i>I</i> ₂	-0.66	-2.15	—	—	—	—	-1.35
5	fluoromethane · methane	0.13	-0.81	1.26	-0.53	0.87	-0.66	-0.75
6	chloromethane · methane	0.42	-0.90	1.50	-0.83	1.08	-0.92	-0.98
7	trifluoromethane · methane	-0.65	-1.36	0.17	-1.22	0.05	-1.15	-0.69
8	trichloromethane · methane	0.26	-1.40	0.47	-2.16	0.35	-1.88	-1.15
9	fluoromethane dimer	-0.54	-0.74	-0.53	-1.78	-0.90	-1.96	-1.65
10	chloromethane dimer	0.00	-1.05	-0.35	-2.35	-0.51	-2.16	-1.34
11	trifluorobenzene · benzene	0.61	-5.00	1.20	-7.36	0.73	-6.56	-4.40
12	hexafluorobenzene · benzene	-0.55	-6.86	0.96	-8.90	0.64	-7.72	-6.12
13	chloromethane · formaldehyde	0.20	-0.86	0.92	-1.30	1.01	-0.81	-1.17
14	bromomethane · formaldehyde	-2.58	-3.91	-1.07	-4.10	-0.40	-2.81	-1.72
15	iodomethane · formaldehyde	-5.50	-7.10	—	—	—	—	-2.38
16	F3Cl-methane · formaldehyde	-0.14	-1.26	0.71	-1.92	0.89	-1.24	-2.25
17	F3Br-methane · formaldehyde	-5.87	-7.25	-2.91	-6.50	-2.01	-4.84	-3.10
18	F3I-methane · formaldehyde	-8.87	-10.57	—	—	—	—	-4.08
19	chlorobenzene · acetone	0.02	-1.41	1.50	-1.12	2.09	-0.09	-1.49
20	bromobenzene · acetone	-4.77	-6.57	-1.13	-4.85	0.74	-2.26	-2.43
21	iodobenzene · acetone	-8.31	-10.43	—	—	—	—	-3.46
22	chlorobenzene · trimethylamine	-0.20	-2.11	0.50	-2.92	1.89	-0.95	-2.11
23	bromobenzene · trimethylamine	-6.57	-9.32	-5.29	-11.06	-2.37	-7.01	-3.78
24	iodobenzene · trimethylamine	-10.96	-14.72	—	—	—	—	-5.81
25	bromobenzene · methanethiol	-1.16	-3.20	-1.13	-5.44	-0.75	-4.17	-2.32
26	iodobenzene · methanethiol	-2.98	-5.48	—	—	—	—	-3.08
27	bromomethane · benzene	-0.59	-3.04	-1.48	-5.70	-0.95	-1.81	-1.81
28	iodomethane · benzene	-1.10	-4.13	—	—	—	—	-2.48
29	trifluorobromomethane · benzene	-0.96	-3.85	-2.60	-7.83	-2.12	-6.31	-3.11
30	trifluoroiodomethane · benzene	-1.52	-5.09	—	—	—	—	-3.91
31	trifluoromethanol · water	-15.28	-13.24	-2.39	-10.31	-5.43	-12.79	-9.67
32	trichloromethanol · water	-16.55	-15.17	-1.71	-10.99	-3.28	-11.79	-10.41
33	<i>HF</i> · methanol	-9.51	-6.13	0.97	-5.55	-2.89	-8.41	-9.59
34	<i>HCl</i> · methanol	-9.50	-8.65	-0.27	-5.59	-2.51	-6.90	-6.30

D. Supporting Information to Chapter 5

35	<i>HBr</i> · methanol	-10.30	-9.86	-0.95	-5.91	-1.75	-5.83	-5.36
36	<i>HI</i> · methanol	-6.72	-6.90	—	—	—	—	-3.97
37	<i>HF</i> · methylamine	-9.29	-7.93	3.81	-5.52	-2.72	-10.62	-14.32
38	<i>HCl</i> · methylamine	-14.58	-14.54	1.70	-8.94	-0.61	-9.53	-11.42
39	methanol · fluoromethane	-4.35	-2.60	1.08	-2.62	0.20	-2.89	-3.89
40	methanol · chloromethane	-1.60	-3.03	1.89	-2.62	1.51	-2.18	-3.78

Table D.11.: HF/minix, HF-3c, DFTB3, DFTB3-D3, OM2, OM2-D3, PM6, PM6-D3H+ and PM7 interaction energies on X40 test set. All values are in kcal/mol.

No.	Name	HF/minix	HF-3c	DFTB3	DFTB3-D3	OM2	OM2-D3	PM6	PM6-D3H+	PM7
1	methane · <i>F</i> ₂	0,12	0,38	-0,03	-0,31	-0,01	-0,25	0,39	0,08	0,20
2	methane · <i>Cl</i> ₂	0,14	-0,49	-0,29	-1,00	—	—	-0,98	-1,74	-1,72
3	methane · <i>Br</i> ₂	0,05	-1,70	—	—	—	—	-1,80	-2,74	-0,77
4	methane · <i>I</i> ₂	-0,05	-2,02	—	—	—	—	-1,82	-2,82	-0,52
5	fluoromethane · methane	0,35	-0,47	0,02	-0,78	0,00	-0,66	-0,10	-0,68	-0,44
6	chloromethane · methane	0,52	-0,74	-0,35	-1,39	—	—	-0,35	-1,12	-1,25
7	trifluoromethane · methane	0,08	0,02	-0,08	-0,79	-0,02	-0,63	0,01	-0,52	-0,52
8	trichloromethane · methane	0,73	-0,46	-1,14	-2,45	—	—	-0,54	-1,52	-2,81
9	fluoromethane dimer	-0,83	-0,49	-1,61	-2,08	-0,68	-1,07	-0,08	-0,30	-0,54
10	chloromethane dimer	-0,09	-0,76	-0,68	-1,40	0,00	-0,56	-0,98	-1,29	-1,67
11	trifluorobenzene · benzene	1,56	-4,52	-0,52	-5,09	0,01	-3,83	-0,74	-3,92	-5,44
12	hexafluorobenzene · benzene	0,84	-5,64	-1,81	-6,91	-1,49	-5,75	-1,44	-5,07	-6,68
13	chloromethane · formaldehyde	0,54	-0,16	-1,85	-2,55	—	—	-0,50	-1,24	-1,36
14	bromomethane · formaldehyde	-1,37	-3,13	—	—	—	—	-1,92	-2,81	-0,67
15	iodomethane · formaldehyde	-2,90	-5,07	—	—	—	—	-2,64	-3,57	2,30
16	F3Cl-methane · formaldehyde	-0,93	-1,74	-5,23	-5,96	—	—	-1,87	-2,60	-2,33
17	F3Br-methane · formaldehyde	-4,49	-6,34	—	—	—	—	-3,71	-4,50	-2,01
18	F3I-ethane · formaldehyde	-6,95	-9,30	—	—	—	—	-5,71	-6,50	2,99
19	chlorobenzene · acetone	0,48	-0,64	-3,58	-4,64	—	—	-0,70	-1,69	-1,98
20	bromobenzene · acetone	-2,57	-4,85	—	—	—	—	-2,73	-3,93	-1,58
21	iodobenzene · acetone	-4,64	-7,43	—	—	—	—	-3,92	-5,17	2,75
22	chlorobenzene · trimethylamine	0,09	-1,63	-6,27	-7,70	—	—	-0,66	-1,87	-3,25
23	bromobenzene · trimethylamine	-3,46	-7,00	—	—	—	—	-9,57	-11,23	-4,24
24	iodobenzene · trimethylamine	-6,60	-11,64	—	—	—	—	-5,66	-7,74	-0,46
25	bromobenzene · methanethiol	0,27	-2,40	—	—	0,00	-0,87	-0,18	-1,42	-1,84
26	iodobenzene · methanethiol	-0,50	-3,85	—	—	0,00	-0,89	-3,92	-5,16	-2,15
27	bromomethane · benzene	-0,07	-3,18	—	—	—	—	-1,61	-3,29	-2,43
28	iodomethane · benzene	-0,56	-4,45	—	—	—	—	-2,52	-4,56	-2,26
29	F3Br- benzene	-1,46	-5,19	—	—	—	—	-2,68	-4,66	-3,20
30	F3I- benzene	-2,43	-7,09	—	—	—	—	-4,12	-6,43	-2,84
31	trifluoromethanol · water	-12,77	-12,04	-8,16	-8,97	—	—	-6,19	-9,51	-9,34
32	trichloromethanol · water	-12,20	-12,39	-7,85	-9,11	—	—	-7,08	-10,80	-11,40
33	<i>HF</i> · methanol	-10,60	-9,72	-7,98	-8,58	-4,69	-5,16	-3,58	-4,05	-2,71
34	<i>HCl</i> · methanol	-6,96	-7,25	-6,98	-7,81	—	—	-5,72	-6,38	-5,65
35	<i>HBr</i> · methanol	-6,64	-7,52	—	—	—	—	-5,47	-6,21	-4,42
36	<i>HI</i> · methanol	-4,24	-5,28	—	—	—	—	-4,70	-5,48	-4,76
37	<i>HF</i> · methylamine	-12,93	-13,04	-6,55	-7,38	-6,76	-7,40	-3,08	-3,67	-3,42
38	<i>HCl</i> · methylamine	-11,73	-13,10	-7,38	-8,66	—	—	-8,28	-9,21	-10,84
39	methanol · fluoromethane	-3,33	-2,66	-2,69	-3,46	-1,39	-2,02	1,21	0,51	2,01
40	methanol · chloromethane	-0,85	-1,37	-1,20	-2,26	—	—	-1,96	-2,92	-2,45

L7 Energies

Table D.12.: HF/ECP-2G, HF-3cv, INDO, INDO-D3H+, NDDO und NDDO-D3H+ interaction energies on L7 test set. All values are in kcal/mol.

No.	Name	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	NDDO	NDDO-D3H+	E_{disp}^{abc}	Ref.
1	CBH	2.61	-13.56	20.49	-4.02	16.55	-4.93	0.73	
2	C2C2PD	6.19	-24.55	5.00	-32.89	2.99	-29.45	1.23	
3	C3A	2.96	-18.56	4.75	-22.81	2.46	-21.16	2.31	
4	C3GC	6.62	-31.32	7.87	-41.45	4.40	-37.82	1.72	
5	GCGC	1.64	-18.08	5.06	-25.98	3.48	-24.00	1.02	
6	GGG	4.35	-2.50	3.94	-6.54	5.11	-4.22	0.26	
7	PHE	-16.00	-23.79	-0.80	-24.83	-3.11	-24.61	0.39	

Table D.13.: HF/minix, HF-3c, DFTB3, DFTB3-D3, OM2, OM2-D3, PM6, PM6-D3H+ and PM7 interaction energies on L7 test set. All values are in kcal/mol.

No.	Name	HF/minix	HF-3c	DFTB3	DFTB3-D3	OM2	OM2-D3	PM6	PM6-D3H+	PM7
1	CBH	3.29	-12.88	-1.23	-14.10	-1.56	-13.10	-1.18	-10.54	-14.95
2	C2C2PD	8.61	-22.12	1.49	-20.92	3.60	-15.65	-0.65	-16.06	-28.23
3	C3A	4.78	-16.74	0.41	-15.54	-0.49	-14.22	-3.37	-14.91	-20.74
4	C3GC	10.03	-27.91	0.93	-27.07	1.42	-22.60	-4.48	-24.78	-35.87
5	GCGC	5.25	-14.47	0.79	-14.84	0.26	-13.09	-6.54	-20.70	-26.03
6	GGG	5.83	-1.03	5.31	-0.34	4.46	-0.43	1.69	-3.20	-6.91
7	PHE	-15.47	-23.26	-16.97	-24.63	-17.42	-24.14	-17.28	-27.09	-28.86

S30L Energies

The S30L binding energies on the different levels are given in Table D.14

Table D.14.: HF/ECP-2G, HF-3cv, INDO, INDO-D3H+, NDDO und NDDO-D3H+ interaction energies on S30l test set. All values are in kcal/mol.

No.	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	NDDO	NDDO-D3H+	E_{disp}^{abc}	Ref.
1	6.21	-26.08	22.25	-34.46	17.26	-30.87	1.83	
2	6.20	-17.98	17.34	-24.19	12.73	-22.54	1.27	
3	5.86	-18.81	14.81	-36.01	12.04	-31.37	1.95	
4	-8.07	-19.34	-2.60	-26.80	-3.98	-24.80	0.74	
5	7.53	-32.70	5.04	-51.28	3.39	-44.73	2.28	
6	16.39	-22.69	33.94	-29.78	26.99	-26.98	1.99	
7	23.18	-36.82	15.48	-66.56	10.77	-59.23	1.99	
8	25.58	-42.36	15.94	-74.91	11.27	-66.37	2.22	
9	16.31	-31.18	17.20	-51.33	11.63	-46.60	3.30	
10	16.93	-33.09	20.55	-52.45	14.35	-47.76	3.61	
11	26.40	-35.48	20.15	-68.84	8.74	-67.38	5.12	
12	26.47	-36.04	19.91	-68.13	8.35	-67.02	5.07	
13	2.77	-28.89	11.85	-33.58	7.37	-32.03	2.93	
14	5.92	-30.24	11.89	-41.00	7.55	-37.75	3.04	

D. Supporting Information to Chapter 5

15	-5.14	-20.98	19.29	-34.22	17.20	-29.45	1.05
16	-0.06	-14.51	16.67	-28.38	14.13	-25.24	1.03
17	-3.56	-17.92	19.13	-12.52	15.33	-13.32	1.63
18	-0.46	-21.17	36.01	-9.40	29.88	-10.61	2.33
19	5.30	-29.78	12.09	-39.37	10.60	-34.94	3.42
20	10.25	-26.12	14.88	-43.37	13.83	-37.11	3.57
21	-41.76	-41.31	19.25	-30.32	14.63	-30.13	0.11
22	-70.40	-73.51	3.79	-38.42	-3.89	-42.59	-0.24
23	-41.34	-73.86	22.62	-59.53	11.24	-58.99	2.22
24	-45.22	-78.55	-2.74	-64.81	-9.71	-63.49	2.21
25	13.01	-31.14	17.34	-39.72	15.14	-34.08	2.70
26	12.93	-31.40	17.55	-39.57	15.37	-33.92	2.74
27	-65.79	-87.10	-46.98	-92.23	-50.49	-91.02	2.33
28	-65.68	-83.09	-46.88	-85.72	-49.40	-84.26	1.92
29	-78.81	-122.52	-16.06	-132.95	-26.39	-128.05	4.82
30	-87.37	-134.88	-54.77	-150.09	-64.77	-147.16	5.88

Table D.15.: HF/minix, HF-3c, DFTB3, DFTB3-D3, OM2, OM2-D3, PM6, PM6-D3H+ and PM7 interaction energies on S30l test set. All values are in kcal/mol.

No.	HF/minix	HF-3c	DFTB3	DFTB3-D3	OM2	OM2-D3	PM6	PM6-D3H+	PM7
1	9,21	-30,89	-0,14	-28,62	-8,09	-32,28	-9,1	-29.42	-38,49
2	8,14	-21,26	0,75	-19,88	-2,93	-20,47	-4,79	-19.64	-26,70
3	9,68	-20,78	0,70	-23,94	-4,00	-24,94	-6,37	-24.23	-32,04
4	-6,19	-19,22	-6,48	-18,79	—	—	-10,1	-18.93	-24,23
5	5,86	-34,38	-2,85	-34,07	-8,09	-34,64	-12,18	-34.74	-45,56
6	17,09	-25,71	4,74	-24,61	1,16	-23,59	-5,55	-26.01	-35,51
7	24,24	-39,88	5,28	-38,78	7,38	-30,04	-0,37	-31.20	-49,45
8	26,29	-45,32	5,55	-44,03	8,36	-33,89	-0,67	-35.17	-57,31
9	19,42	-37,37	3,89	-36,37	1,86	-32,71	-1,07	-30.39	-57,19
10	20,83	-39,26	4,88	-38,10	3,32	-33,68	-0,41	-31.82	-60,79
11	34,60	-39,28	8,29	-44,49	—	—	-2,42	-41.48	-75,20
12	34,13	-39,69	8,53	-44,33	—	—	-2,27	-41.69	-75,97
13	3,02	-28,06	-1,88	-26,75	-3,65	-25,45	-8,91	-27.77	-35,81
14	6,70	-29,68	-2,78	-30,64	—	—	-8,14	-29.68	-36,04
15	-8,28	-28,56	-12,46	-31,15	-20,67	-36,47	-22,81	-42.53	-45,09
16	-1,81	-20,47	-5,97	-23,09	-12,11	-26,67	-14,75	-32.42	-34,48
17	-1,96	-16,41	-2,91	-17,15	-2,63	-15,43	-5,42	-21.89	-19,26
18	1,52	-19,79	-2,14	-22,62	-2,41	-20,75	-6,44	-27.38	-25,23
19	4,89	-27,62	-0,31	-28,22	1,03	-23,82	-4,53	-28.33	-33,16
20	12,10	-23,49	2,58	-27,79	2,24	-24,53	-2,14	-27.74	-32,81
21	-31,70	-39,52	-22,48	-33,76	-25,38	-34,99	-24,05	-42.22	-49,26
22	-57,78	-66,17	-33,96	-41,69	-43,83	-50,40	-43,36	-58.79	-72,11
23	-29,35	-68,07	-24,20	-55,71	-31,80	-58,68	-19,87	-45.67	-60,40
24	-41,69	-74,13	-35,62	-64,14	-36,68	-61,43	-28,76	-53.28	-70,83
25	10,94	-34,08	3,70	-29,38	—	—	-0,95	-24.33	-50,48
26	10,77	-34,28	3,68	-29,49	2,12	-26,97	-0,74	-24.11	-50,85
27	-74,51	-92,41	-75,34	-95,17	-77,48	-94,98	-75,15	-94.92	-112,65
28	-74,30	-88,76	-73,26	-89,88	-76,53	-91,20	-74,78	-91.95	-110,49
29	-80,47	-129,21	-96,57	-138,17	-110,77	-147,04	-105,85	-144.62	-180,77
30	-100,35	-144,1	-115,69	-162,05	-121,24	-162,21	-126,22	-164.68	-183,77

X23 Energies

The lattice energies for the X23 organic crystals (on the different levels) are given in Table D.16. X23 lattice energies are compared to thermodynamically back-corrected experimental sublimation energies. The lattice energies are calculated on optimized structures with fixed (experimental) unit cell.³⁹

Table D.16.: HF/ECP-2G, HF-3cv, INDO and INDO-D3H+ lattice energies on X23 test set. All values are in kcal/mol.

No.	Name	HF/ECP-2G	HF-3cv	INDO	INDO-D3H+	E_{disp}^{abc}	Ref.
1	cyclohexanedione	—	21.19	9.49	-24.50	1.44	-21.18
2	acetic acid	-12.46	-17.00	-0.28	-23.54	0.73	-17.40
3	adamantane	5.25	-17.66	30.85	-7.50	2.29	-16.59
4	ammonia	-8.74	-11.54	6.12	-4.84	0.23	-8.89
5	anthracene	—	—	—	—	2.32	-26.93
6	benzene	2.11	-12.94	11.09	-12.18	1.12	-12.36
7	CO ₂	-5.82	-10.45	-2.73	-19.63	0.27	-6.50
8	cyanamide	-10.05	-19.73	—	—	0.53	-19.05
9	cytosine	-24.87	-38.41	8.24	-34.68	1.64	-40.58
10	ethylcarbanate	—	-11.61	3.58	-21.36	0.74	-20.63
11	formamide	-15.09	-18.62	1.30	-21.06	0.56	-18.93
12	imidazole	-2.50	-14.64	21.55	-13.64	1.86	-20.60
13	hexamine	—	—	7.81	-14.52	0.81	-20.75
14	naphtalene	—	—	16.52	-20.13	1.67	-19.53
15	oxalic acid α	-23.20	-29.26	-16.29	-50.18	0.94	-23.02
16	oxalic acid β	—	—	-14.29	-51.28	0.71	-22.97
17	pyrazine	—	-16.46	8.41	-15.94	1.13	-14.65
18	pyrazole	-9.83	-20.31	—	—	0.86	-18.57
19	succinic acid	-22.37	-26.25	-0.79	-37.89	1.49	-33.44
20	triazine	—	—	5.59	-15.33	0.96	-14.75
21	trioxane	—	—	9.62	-11.77	0.94	-15.87
22	uracil	—	—	1.01	-35.40	1.20	-32.43
23	urea	-20.10	-22.73	2.60	-28.15	0.71	-24.50

Table D.17.: HF/minix, HF-3c, DFTB3 and DFTB3-D3 lattice energies on X23 test set.

All values are in kcal/mol.

No.	Name	HF/minix	HF-3c	DFTB3	DFTB3-D3
1	cyclohexanedione	—	—	-5.71	-24.07
2	acetic acid	-10.08	-16.88	-10.73	-18.59
3	adamantane	4.91	-18.53	-0.85	-19.39
4	ammonia	-5.44	-9.55	-4.09	-7.7
5	anthracene	—	—	-0.59	-26.81
6	benzene	2.13	-14.31	-0.94	-13.38
7	CO ₂	-3.88	-7.7	-1.28	-5.53
8	cyanamide	-8.08	-17.74	-8.65	-15.88
9	cytosine	-20.96	-38.05	-18.69	-35.16
10	ethylcarbanate	—	—	-11.42	-23.05
11	formamide	-12.28	-17.34	-11.61	-18.37
12	imidazole	-1.59	-17.16	-0.05	-25.02
13	hexamine	—	-2564	-6.34	-16.91

14	naphtalene	—	—	-0.43	-17.12
15	oxalic acid α	-20.55	-27.6	-17.61	-27.32
16	oxalic acid β	—	—	-18.69	-26.85
17	pyrazine	—	—	-1.6	-13.35
18	pyrazole	-6.09	-19.3	-3.84	-14.17
19	succinic acid	-17.15	-26.27	-19.2	-33.44
20	triazine	—	—	-1.17	-12.35
21	trioxane	—	—	-3.76	-15.59
22	uracil	—	—	-19.21	-35.53
23	urea	-17.88	-23.93	-18.78	-27.18

Acknowledgments

First, I am grateful to my supervisor Prof. Dr. Stefan Grimme. I could work on highly interesting research topics and his support was always encouraging. His (occasionally) impatience and his productiveness significantly increased my own efficiency.

I thank Prof. Dr. Thomas Bredow for our joint research projects and for being my second referee.

Thanks go also to all colleagues in the Grimme group with great working (and non-working) atmosphere. I thank all my collaborators in Bonn, Torino, Braunschweig, Warsaw, Bielefeld, Münster, and Berlin.

Finally, my dearest thank go to my family. Without their support, my studies would not have been possible.

Curriculum Vitae

Jan Gerit Brandenburg

Personal Data

Address

Mobile

E-mail

Languages

German (native)

English (fluent, C1 classified by ZSL Heidelberg)

Education

03/2012 – 07/2015

Mulliken Center for Theoretical Chemistry
University of Bonn

Postgraduate studies

PhD thesis on

“Development and Application of Electronic Structure Methods for Noncovalent Interactions in Organic Solids” (grade A*, *summa cum laude*)

10/2006 – 02/2012

Heidelberg University

Diploma in physics (grade A)

Thesis at the Institute for Theoretical Physics “Dimensionality of Spinmodulations in Superconducting Cuprates” (grade A*)

08/1997 – 07/2006

High School “Gymnasium Herzogenrath”

Abitur, certificate of general university maturity (grade A*, grant of German Physical Society DPG)

Work Experience

Since 03/2012

Mulliken Center for Theoretical Chemistry
University of Bonn

Researcher

Working in the group of **Prof. Dr. Stefan Grimme** on development and evaluation of density functional methods for non-covalent interactions in supramolecular and periodic systems

Teaching assistant for the lectures “Quantum Chemistry I/II” and “Theoretical Methods for the Condensed Phase” (organisation of lecture and tutorials, devising exercises and exams, and teaching a tutorial)

10/2011 – 02/2012

Institute for Theoretical Physics
Heidelberg University

Research assistant

Teaching tutorials for the lecture “Theoretical Mechanics”

10/2008 – 02/2011

Interdisciplinary Center for Scientific Computing
Heidelberg University

Research assistant

Teaching assistant for the lectures “Mathematics for Scientists I/II” (organisation of lecture and tutorials, devising exercises and exams, and teaching a tutorial)

03/2005 – 03/2005

National Aeronautics and Space Research Center

Trainee at the Institute for Propulsion Technology

11/2002 – 11/2002

RWTH Aachen University

Trainee at the Leibniz Institute for Interactive Materials

Scientific Conferences (Selection)

- 09/2014 *Symposium on Theoretical Chemistry* in Vienna, Austria
Poster: “Semiempirical Methods for Organic Crystal Modeling”
- 07/2014 *International Conference on Computational and Mathematical Methods in Science and Engineering* in Cádiz, Spain
Invited talk: “Electronic Structure Modeling of Organic Molecular Crystals”
- 03/2014 *IFF Spring School: Computational Materials Physics* in Jülich, Germany
- 09/2013 *European Summerschool in Quantum Chemistry* in Palermo, Italy
Poster: “Dispersion Corrected DFT and HF for Organic Crystal Structure Prediction”
- 03/2013 *Symposium on Weak Molecular Interactions* in Pécs, Hungary
Poster: “A Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic DFT Calculations”
- 09/2012 *Symposium on Theoretical Chemistry* in Karlsruhe, Germany

Scientific Publications (Selection)

- J. G. Brandenburg, G. Bender, J. Ren, A. Hansen, S. Grimme, H. Eckert, C. G. Daniliuc, G. Kehr, and G. Erker, *Crystal Packing Induced Carbon-Carbon Double-Triple Bond Isomerization in a Zirconocene Complex*, *Organometallics* **2014**, *33*, 5358–5364.
- J. G. Brandenburg and S. Grimme, *Accurate Modeling of Organic Molecular Crystals by Dispersion-corrected Density Functional Tight-Binding (DFTB)*, *J. Phys. Chem. Lett.* **2014**, *5*, 1785–1789.
- J. G. Brandenburg and S. Grimme, *Dispersion Corrected Hartree-Fock and Density Functional Theory for Organic Crystal Structure Prediction*, *Top Curr Chem* **2014**, *345*, 1–23.
- J. G. Brandenburg and S. Grimme, *A Dispersion-Corrected Density Functional Theory Case Study on Ethyl Acetate Conformers, Dimer, and Molecular Crystal*, *Theor. Chem. Acc.* **2013**, *132*, 1399.
- J. G. Brandenburg, M. Alessio, B. Civalleri, M. F. Peintinger, T. Bredow, and S. Grimme, *Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic Density Functional Theory Calculations*, *J. Phys. Chem. A* **2013**, *117*, 9282–9292.
- J. G. Brandenburg, S. Grimme, P. G. Jones, G. Markopoulos, H. Hopf, M. K. Cyranski, and D. Kuck, *Unidirectional Molecular Stacking of Tribenzotriquinacenes in the Solid State: A Combined X-Ray and Theoretical Study*, *Chem. Eur. J.* **2013**, *19*, 9930–9938.

IT Skills

- OS: Windows, Linux (daily use) and Mac (tutorials at the IWR)
- CAS: Mathematica, Maple (teaching Maple tutorials 09/2009 – 02/2011)
- Fortran (published code and contributions to CRYSTAL14.1)
Python and bash scripting, L^AT_EX, Gnuplot, MS/Open Office

Volunteer Work and Interests

- Volleyball: trained referee and former youth coach (DJK Hockenheim)
- Ballroom dancing
- Fantasy and historical novels

Bonn, July 6, 2015