



PhD-FSTC-2018-73
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defense held on 16/11/2018 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN PHYSIQUE

by

Johannes HOJA

FIRST-PRINCIPLES MODELING OF MOLECULAR
CRYSTALS: CRYSTAL STRUCTURE PREDICTION
AND VIBRATIONAL PROPERTIES

Dissertation defense committee

Dr. Alexandre Tkatchenko, Dissertation Supervisor
Professor, Université du Luxembourg

Dr. Ludger Wirtz, Chairman
Professor, Université du Luxembourg

Dr. Josh Berryman, Vice Chairman
Research Scientist, Université du Luxembourg

Dr. Sarah L. Price
Professor, University College London

Dr. Graeme M. Day
Professor, University of Southampton

This PhD thesis is based on (and parts of it have been published in) the following articles:

- J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio Jr., A. Tkatchenko, Reliable and Practical Computational Description of Molecular Crystal Polymorphs, *Sci. Adv.* **2019**, *5*, eaau3338.
- J. Hoja, A. Tkatchenko, First-Principles Stability Ranking of Molecular Crystal Polymorphs with the DFT+MBD Approach, *Faraday Discuss.* **2018**, *211*, 253–274.
- G. Folpini, K. Reimann, M. Woerner, T. Elsaesser, J. Hoja, A. Tkatchenko, Strong Local-Field Enhancement of the Nonlinear Soft-Mode Response in a Molecular Crystal, *Phys. Rev. Lett.* **2017**, *119*, 097404.
- J. Hoja, A. M. Reilly, A. Tkatchenko, First-Principles Modeling of Molecular Crystals: Structures and Stabilities, Temperature and Pressure, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, 1294.
- A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, J. Hoja, E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman, B. Kahr, Powder Diffraction and Crystal Structure Prediction Identify Four New Coumarin Polymorphs, *Chem. Sci.* **2017**, *8*, 4926–4940.
- A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr., A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu, C. R. Groom, Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods, *Acta Crystallogr. Sect. B: Struct. Sci.* **2016**, *72*, 439–459.

For the three publications, in which the author of this thesis is listed as first author, he performed and interpreted all or the majority of calculations and wrote the initial manuscript. In the case of the other publications, only calculations which were completely or predominantly performed by the author are presented in this thesis.

Abstract

Understanding the structure and stability, as well as response properties of molecular crystals at certain thermodynamic conditions is crucial for the engineering of new molecular materials and the design of pharmaceuticals. A reliable description of the polymorphic energy landscape of a molecular crystal would provide an extensive insight into the development of drugs in terms of the existence and the likelihood of late-appearing polymorphs. Furthermore, accurate modeling of low-frequency vibrational spectra would be important for the characterization of molecular crystal polymorphs. However, an accurate description of molecular crystals is very challenging since many properties highly depend on the crystal-packing arrangement of the involved molecules and the temperature. The difficulties for computational predictions of molecular crystal polymorphs lie in the high dimensionality of crystallographic and conformational space, and the need for very accurate relative free energies. It was shown that accurate lattice energies can be obtained by using density-functional theory (DFT) calculations supplemented by a high-level model for long-range van der Waals (vdW) dispersion interactions, such as the many-body dispersion (MBD) model. Therefore, this thesis utilizes throughout vdW-inclusive DFT using the MBD and the related pairwise Tkatchenko-Scheffler (TS) dispersion model and the importance of dispersion interactions is highlighted for several properties. A hierarchical stability-ranking approach based on the DFT+MBD framework for the final stage of a molecular crystal structure prediction procedure is presented and analyzed. This approach provides excellent stability rankings over the diverse set of molecular crystals studied in the latest blind test of the Cambridge Crystallographic Data Centre. The results suggest that accounting for many-body dispersion effects and vibrational free energies can be crucial for the description of relative stabilities, especially for highly polymorphic systems. The presented approach enables the calculation of reliable structures and thermodynamic stabilities for pharmaceutically relevant systems, contributing to a better understanding of complex polymorphic energy landscapes. Furthermore, many first-principles calculations are performed by using fully optimized structures and free energies obtained within the harmonic approximation, neglecting the thermal expansion of the studied molecular crystal and further anharmonic effects. Therefore, this thesis illustrates that the majority of the thermal expansion of molecular crystals can be captured with the used methods by applying the quasi-harmonic approximation. In addition, we estimate further anharmonic effects on the vibrational frequencies by utilizing Morse oscillators.

Acknowledgements

First and foremost, I would like to thank my supervisor Alexandre Tkatchenko for the opportunity to work on this very interesting topic as well as for his guidance and support during the last four years. I would like to express my gratitude to all my collaborators but especially to Hsin-Yu Ko, Robert A. DiStasio Jr., Anthony Reilly, and Marcus Neumann for their valuable contributions and helpful discussions. Financial support from the Deutsche Forschungsgemeinschaft (SFB-951/A10 and SPP 1807), the European Research Council Consolidator Grant BeStMo, and the University of Luxembourg is gratefully acknowledged. Furthermore, I would like to thank the Fritz Haber Institute of the Max Planck Society, the Argonne National Laboratory, and the High-Performance Computing facilities of the University of Luxembourg for providing the computational resources for this thesis. Moreover, I would like to extend special thanks to all my co-workers at the Theory Department of the Fritz Haber Institute in Berlin and to all members of the Theoretical Chemical Physics group at the University of Luxembourg for many fruitful discussions and their support in difficult times. Finally, I would particularly like to thank my parents for their encouragement and continuous support.

Contents

Abstract	v
Acknowledgements	vii
Contents	ix
1 Introduction	1
1.1 Intermolecular Interactions	1
1.2 Molecular Crystals	4
1.3 Prediction of Crystal Structures and their Properties	4
1.4 Outline	6
I Theoretical Background	9
2 Electronic Structure Theory	11
2.1 The Schrödinger Equation	11
2.2 The Born-Oppenheimer Approximation	13
2.3 The Pauli Exclusion Principle	13
2.4 The Wave Function of an Electron	14
2.5 Many-Electron Wave Functions	15
2.6 The Hartree-Fock Method	16
3 Density Functional Theory	19
3.1 The Hohenberg-Kohn Theorems	19
3.2 Kohn-Sham DFT	20
3.3 Density-Functional Approximations	21
3.4 Van der Waals Dispersion Models	22
3.4.1 Tkatchenko-Scheffler Dispersion Model	23
3.4.2 Many-Body Dispersion Model	24
4 Lattice Vibrations	29
4.1 The Harmonic Approximation	29

CONTENTS

4.2	The Classical Harmonic Oscillator	31
4.3	The Quantum Harmonic Oscillator	34
4.4	The Morse Oscillator	35
4.5	Phonon Modes and Dispersion in Simple Model Systems	36
4.5.1	Infinite Chain of Atoms with One Atom in the Unit Cell	36
4.5.2	Infinite Chain of Atoms with Two Atoms in the Unit Cell	38
4.6	Vibrations in Molecules	40
II Calculation of Vibrational, Thermal, and Mechanical Properties		43
5	Illustration of Modeling Approaches Using a Deuteroammonia Crystal	45
5.1	Static Lattice Energies	46
5.2	Geometries	47
5.3	Temperature and Pressure	48
5.4	The Harmonic Approximation	49
5.4.1	Harmonic Helmholtz free energies	50
5.4.2	Harmonic vibrational spectra	51
5.4.3	What is missing in the harmonic approximation?	54
5.5	The Quasi-Harmonic Approximation	55
5.5.1	Thermal expansion	57
5.5.2	Elastic properties	59
5.6	Summary	63
6	Low-Frequency Spectra of a Purine Crystal	65
6.1	Computational Methods	66
6.2	Calculated Low-Frequency Spectra	67
6.3	Summary	71
7	Time-Resolved THz Spectroscopy of an Aspirin Crystal	73
7.1	Background	73
7.2	Computational Methods	75
7.3	THz Response of Aspirin Form I	75
7.4	Summary	79
III Crystal Structure Prediction		81
8	Sixth Blind Test of Organic Crystal Structure Prediction Methods	83
8.1	Background	83
8.2	Computational Methods	84
8.3	Initial Structure Screening	86
8.4	Overall Results of the Blind Test	87
8.5	Discussion	89
8.6	Summary	90

CONTENTS

9	Recommended Stability-Ranking Procedure Based on DFT+MBD	91
9.1	Computational Methods	91
9.2	Energy Ranking Approach	96
9.3	Polymorphic Energy Landscapes	96
9.4	Stability Ranking Results	99
9.5	Discussion of Most-Stable Unobserved Structures	100
9.6	Beyond the Harmonic Approximation	102
9.7	Summary	104
10	Benchmarking the DFT+MBD CSP Procedure	107
10.1	Computational Methods	107
10.2	Static Lattice Energies	109
10.3	Geometries	112
10.4	Vibrational Free Energies	115
10.5	Computational Cost	120
10.6	Beyond the Harmonic Approximation	121
10.7	Summary	125
11	Relative Stabilities of the Five Coumarin Polymorphs	127
11.1	Computational Methods	128
11.2	Polymorphic Energy Landscapes	129
11.3	Relative Stabilities of Forms I–V	130
11.4	Summary	132
12	Conclusion	133
IV	Appendices	137
A	Benchmark Calculations for Small Molecular Crystals	139
B	Accuracy of Optimized Structures for the Blind-Test Systems	141
C	Stability Rankings for the Blind-Test Systems	145
	Abbreviations	161
	Publication List	163
	List of Figures	165
	List of Tables	173
	Bibliography	179

Chapter 1

Introduction

Molecular crystals are very versatile materials which are used as pharmaceuticals, organic semiconductors, explosives, and in solid-state reactions [1–8]. These solids are composed of molecular moieties, which are held together by intermolecular interactions (non-covalent forces). While these interactions between molecules are generally weaker than covalent bonds, they govern many properties of molecular crystals. Therefore, a proper understanding of intermolecular interactions would be crucial for the field of crystal engineering, in which the primary goal is the design and ultimately the synthesis of new crystalline materials having a predefined arrangement of molecules, enabling certain properties [9, 10]. This would include an understanding of self-recognition and self-organization of molecules. Recognition events are often discussed in this context in terms of specific non-covalent interactions such as hydrogen bonds, halogen bonds, π - π stacking, dipole-dipole interactions, C–H $\cdots\pi$ interactions, van der Waals interactions, etc. [10]. Hence, it is vital to understand and adequately model these different types of interactions in a balanced and accurate manner.

1.1 Intermolecular Interactions

Between closed-shell molecules we generally distinguish four different intermolecular interactions: electrostatics, induction, dispersion, and exchange repulsion [11]. These interactions can be understood based on perturbation theory [12]. The static Coulomb interaction between fixed charge distributions constitutes the electrostatic interaction. In the picture of a multipole expansion, electrostatics is the interaction between permanent multipoles. Therein, the first term describes dipole-dipole interactions, followed by dipole-quadrupole, quadrupole-quadrupole, etc. The interaction between permanent multipoles and induced multipoles is called induction, which is always attractive. Furthermore, we can have an interaction between instantaneously created multipoles and hence induced multipoles, which is called dispersion. Such an instantaneous multipole moment can be created by electronic zero-point fluctuations in one molecule, which induce then a multipole moment in another molecule. Dispersion interactions do not

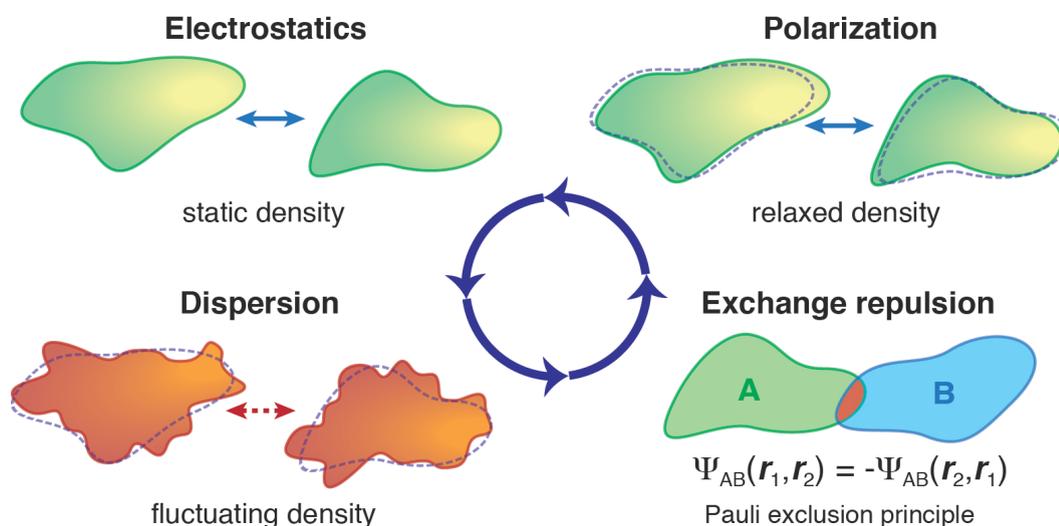


Figure 1.1: Schematic representation of intermolecular interactions. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

require the presence of permanent multipole moments and are therefore ubiquitous. Finally, we discuss the exchange repulsion, which is comprised of two different effects. The first one results from the Pauli exclusion principle, leading to a repulsive force between electrons having the same spin. The second effect originates due to the fact that electrons are at short intermolecular distances able to move over both molecules (in the case of a dimer), which results in an attractive interaction. The overall exchange repulsion is dominated by the first effect and is hence repulsive.

Furthermore, we can also discuss intermolecular interactions in terms of electron densities (see Fig. 1.1) [13]. Electronic structure calculations determine the electron density of a given system via an iterative approach called *self-consistent field* (SCF) [14] (see Chapter 2). Therein, a static density is initialized at the beginning, typically based on atomic-charge information. This static density can be interpreted as electrostatic interaction. With every step in the SCF cycle, the electron densities of the involved molecules relax due to the surrounding charge distribution. This effect leading to relaxed densities can be viewed as induction or polarization. When the intermolecular distance is small and the involved electron clouds overlap, an electron exchange is possible. Due to the Pauli exclusion principle, our total wave function has to be antisymmetric in terms of exchanging two electrons having the same spin. Such an adaptation of the wave function leads to an energy loss and constitutes therefore a repulsive interaction (exchange repulsion). These three effects constitute the description of intermolecular interactions in the Hartree-Fock method (see Chapter 2) or (semi)-local density functional approximations (see Chapter 3). In all methods accounting for long-range electron correlation, instantaneous electronic fluctuations are responsible for a further modification of the electron densities and we can loosely interpret the resulting intermolecular interaction as dispersion. Ferri *et al.* [15] have shown that dispersion interactions can lead

to a significantly modification of the electron density and related properties. Note that generally all discussed intermolecular interactions are closely linked together and they cannot be unambiguously separated in practice, especially when dealing with extended systems like molecular crystals. While first-principles methods enable in principle the accurate description of all discussed intermolecular interactions, many calculations are still performed by using empirical force fields due to their computational efficiency. However, empirical force fields are sometimes limited to electrostatic interactions or utilize simplistic models for describing the dispersion and exchange repulsion. Often, induction effects are completely neglected, notwithstanding the development of polarizable force fields [16–18].

The term *van der Waals (vdW) interactions* can have varying meaning depending on the community, as discussed by Dobson and Gould [19]. In solid-state physics this term typically refers to dispersion interactions, while in chemistry it is often also used for the sum of all intermolecular interactions. In this thesis, we will use the former definition throughout. Due to their ubiquity and relatively long range, vdW interactions often dominate the cohesive energy of molecular crystals. However, in the context of crystal engineering, hydrogen bonds [20] and halogen bonds can also be crucial due to their directionality at short distances, which could be utilized to control certain structural motifs during the formation of a crystal. The intermolecular interactions between hydrogen/halogen-bonded molecules are typically a combination of all discussed types of interactions, which was analyzed in Refs. 21–23 by utilizing symmetry-adapted perturbation theory [24, 25].

In case of small molecular dimers, interaction energies can be obtained extremely accurately on a first-principles level by using coupled-cluster approaches [26] or quantum Monte Carlo [27], which lead to several benchmark databases for non-covalent interactions [28–30]. Such methods can unfortunately only be applied to relatively small systems due to their tremendous computational cost. For larger systems, density functional theory (DFT) has emerged as the method of choice due to the continuous development of better density-functional approximations (DFA) and the inclusion of long-range vdW interactions via a variety of different dispersion models [31–37]. However, it should be noted that the description of intermolecular interactions on a vdW-inclusive DFA level can sometimes be inadequate, for instance in the case of large charge transfer effects or strong hydrogen bonds [38]. For periodic systems like molecular crystals we need a DFA which provides a sufficient description of covalent and intermolecular interactions while still being computationally efficient. The performance of DFA+vdW approaches is often evaluated by comparing the results with experimental structures and stabilities [39, 40] and we will discuss the accuracy of several vdW-inclusive DFA approaches in Chapter 4.

1.2 Molecular Crystals

For a given molecular crystal formed by the same molecular moieties, different crystal-packing arrangements (polymorphs) can be possible. It was recently shown by Cruz-Cabeza *et al.* that more than 50 % of compounds within a large dataset of solid form screenings exhibit polymorphism [41]. Typically, the energy difference between these solid forms amounts to less than 1 kcal/mol (4.2 kJ/mol) and sometimes it can even be smaller than 1 kJ/mol [41]. Therefore, methods with an accuracy of at least 1 kJ/mol would in principle be needed in order to ensure a correct description of relative stabilities of molecular crystal polymorphs. The calculation of accurate stabilities is further complicated by the fact that molecular crystals are experimentally always grown and studied at finite temperatures. So far, thermal effects have often been neglected in first-principles calculations, although they can be responsible for a significant re-ordering of low-energy polymorphs in terms of their relative stability [42, 43].

Moreover, the crystallization of molecular crystals is not always governed by thermodynamics. Therefore, it is sometimes not possible to crystallize the thermodynamically most stable form because of kinetic effects [44]. One example would be a thermodynamically stable crystal structure, in which the molecular conformation is very different from the one in the gas phase or solution. In such a case, the crystal structure could simply be inaccessible under the available crystallization conditions. However, the reason for the experimental unavailability of a computationally predicted solid form is often the fact that the appropriate crystallization experiment has simply not been performed yet [45]. Furthermore, it is often necessary to look beyond single-component molecular crystals and consider complex multi-component crystals like salts, co-crystals, hydrates, and solvates.

In addition, also the modeling of response properties of molecular crystals is quite challenging. In general, polymorphs can for example have completely different melting points, solubilities, vibrational spectra, heat capacities, elastic constants, refractive indices, densities, conductivities, vapor pressures and nuclear magnetic resonance (NMR) chemical shifts [9]. Furthermore, these properties can also highly depend on the temperature, mainly because of the thermal expansion of the crystal. Therefore, knowledge of the unit cell corresponding to the proper temperature is required for accurate calculations of vibrational spectra or elastic properties. Such properties are typically calculated within the harmonic approximation when first-principles methods are used. Hence, anharmonic effects are commonly neglected, although they can be significant at room temperature. [13]

1.3 Prediction of Crystal Structures and their Properties

When properties of already known molecular crystals are studied, the experimentally determined crystal structure typically serves as starting point for all calculations. However, the goal would be to predict properties prior to experimental observations. In such cases, the most stable crystal-packing arrangement of the involved molecules needs

to be determined from scratch in a molecular crystal structure prediction (CSP) procedure before any properties can be calculated. The Cambridge Crystallographic Data Centre (CCDC) organizes regular blind test for organic CSP methods [46–51]. Therein, participants try to predict the structure of a molecular crystal based exclusively on the two-dimensional structural formula of the involved molecule(s). Most approaches utilize force fields but the ratio of first-principles contributions is growing. One CSP method was able to correctly predict most target structures during the latest three blind tests [49–51].

The ability to correctly predict the structures and corresponding stabilities of a molecular crystal including its often numerous polymorphs without prior experimental information would also be crucial for the formulation and design of pharmaceuticals [41]. Accurate CSP methods would be able to provide detailed insight into the thermodynamic polymorphic energy landscape of a given molecular crystal. This would allow an assessment if the existing structure of a drug candidate is in fact the thermodynamically most stable form at room temperature, or if there is a potential risk of a so-called late-appearing polymorph, leading to the eventual disappearance of the current form in the manufacturing process [52]. In such a case, the emerging polymorph might have a similar stability but potentially different and unwanted properties, such as a lower solubility. Such an event could trigger a cascade of health-related and financial repercussions. Therefore, the utilization of computational CSP methods in combination with experimental polymorph screening would offer a solution to this challenge [53]. However, the computational CSP of molecular crystal polymorphs is extremely challenging due to the high dimensionality of the conformational and the crystallographic space, in combination with the need for very accurate relative free energies.

The success and reliability of a CSP method depends on two equally important aspects: a sufficiently complete sampling of the crystallographic and conformational space spanned by a certain molecular crystal, and a sufficiently accurate energetic ranking of the various low-energy polymorphs according to their relative stabilities at the desired thermodynamic conditions [51, 53]. In terms of the latter aspect, major advances have been made during the past several years, resulting in a significant progress in the field of molecular CSP [43, 54–60]. A typical molecular CSP procedure consists of three essential steps, which are illustrated in Fig. 1.2 and discussed below. For a more detailed discussion, the reader is referred to Ref. 53.

First, a three-dimensional molecular structure is obtained based on the available 2D structural formula via geometry optimization, leading to the most stable conformer in the gas phase. For rigid molecules this molecular conformation is typically already close to the one adopted in the crystal. For flexible molecules, energetically relevant conformational isomers must also be taken into account. Special care is needed when intermolecular hydrogen bonds can be formed since this is often accompanied by a significant change in the molecular conformation. Furthermore, also the possibility of tautomers should be considered. For instance, it was believed that barbituric acid occurs only in the keto form. However, a recently discovered polymorph contains the enol form [61].

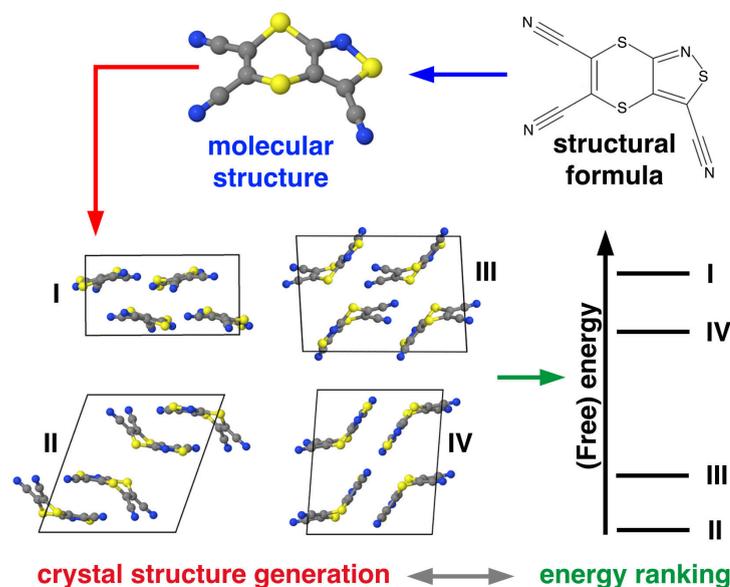


Figure 1.2: Illustration of a typical molecular CSP procedure.

Second, a vast number of different crystal-packing arrangements is generated. Even for rigid molecules, this is a non-trivial task since one has to sample all important space groups, unit-cell sizes, and intermolecular orientations. An important parameter for the complexity of such a crystallographic space sampling is the number of molecules, which constitute the asymmetric unit of a unit cell (Z'). Most searches limit themselves to $Z' = 1$ since any increase leads to a tremendous increase in the number of possible structures. In case of flexible molecules, it is also necessary to consider different torsion angles for rotatable bonds, leading to a further increase in the computational complexity [62, 63].

Finally, after a sufficient sampling of crystallographic space, the created structures have to be ranked according to their stability. This is typically achieved by calculating the lattice energy, which essentially describes the molecular energy gain by forming a crystal compared to infinitely separated molecules. Due to the tremendous number of sampled structures, the whole ranking cannot routinely be performed on a first-principles level. Instead, it is way more efficient to apply a hierarchical procedure utilizing increasingly accurate methods. After an initial stability ranking, a limited amount of structures is retained, which are subsequently re-optimized and re-ranked with a more accurate method. The number of structures depends on the obtained energy intervals and the expected accuracy of the initially applied method.

1.4 Outline

This thesis aims to accurately calculate the relative stabilities of molecular crystal polymorphs and related properties on a first-principles level. It was shown that accurate lattice energies can be obtained for a set of small molecular crystals by utilizing DFT

calculations supplemented by a high-level model for vdW interactions, such as the many-body dispersion [35, 64] (MBD) model [40]. Therefore, vdW-inclusive DFT is applied throughout this thesis utilizing the MBD model as well as the related pairwise Tkatchenko-Scheffler model [34]. This thesis is structured as follows: First, the basics of electronic structure theory are discussed in Chapter 2, followed by a short introduction into DFT in Chapter 3, which also includes a discussion about the used vdW models. Chapter 4 concludes the theoretical background by first discussing the concept of lattice vibrations within the harmonic approximation. Furthermore, the anharmonic Morse oscillator model is described and normal modes are discussed for selected model systems. The next part of the thesis focuses on the modeling of molecular crystal properties using vdW-inclusive DFT. Therefore, Chapter 5 gives an overview of state-of-the-art first-principles modeling approaches and necessary approximations for molecular crystals illustrated by using a cubic ammonia crystal. Next, Chapter 6 discusses the calculation of low-frequency vibrational spectra for a purine crystal, highlighting the importance of thermal expansion and further anharmonic effects captured by using Morse oscillators, resulting in an improved description of vibrational frequencies in the THz range. In addition, Chapter 7 illustrates that the used approaches can even provide a qualitative understanding for time resolved THz-spectroscopy experiments. The third part of this thesis focuses on organic CSP. We have developed a hierarchical computational procedure based on the DFT+MBD framework intended for the final stage of a molecular CSP, for which it is crucial to accurately rank an existing set of possible structures according to their stability. Therein, we also consider vibrational free energies. Chapter 8 shows the results of our contribution to the latest blind test for molecular CSP methods organized by the CCDC [51]. Furthermore, our procedure is extended by hybrid DFT calculations and systematically applied to a different set of initial structures of the blind-test systems in Chapter 9. Therein, we obtain excellent stability rankings for the quite diverse set of molecular crystals and predict a possible new form of one system. Moreover, we highlight the importance of many-body dispersion interactions and vibrational free energies for obtaining accurate stability rankings. In addition, we provide in Chapter 10 a further benchmark and discussion of the approximations and the computational setup used in our presented stability-ranking approach using a subset of structures from the previous chapter. Finally, we apply this approach to coumarin crystals in Chapter 11. By now, coumarin has 5 confirmed polymorphs and their relative stabilities were experimentally obtained. Therefore, this allows us to test the accuracy of our approach on a quantitative level. We show that our relative stabilities agree within 1 kJ/mol with the experimental observations. This thesis illustrates that accurate thermodynamic stabilities can now in principle be obtained for pharmaceutically relevant systems, which is a prerequisite for a better understanding of complex polymorphic energy landscapes.

Part I

Theoretical Background

Chapter 2

Electronic Structure Theory

This chapter provides a description of the basics of electronic structure theory. The discussion builds upon a basic understanding of quantum mechanics, as described for instance by McQuarrie [65]. For a more detailed description the reader is referred to the books by Szabo & Ostlund [14] and Piela [66]. The notation used in this chapter follows mostly the book by Szabo & Ostlund [14]. All equations in this chapter are expressed in atomic units (see for example page 41 of Ref. 14).

2.1 The Schrödinger Equation

The fundamental equation of electronic structure theory is the Schrödinger equation [67–70], which enables a complete quantum-mechanical account of systems containing atomic nuclei (or *ions*) and electrons. The most general version of the Schrödinger equation is given by

$$i \frac{\partial}{\partial t} |\Phi(\mathbf{R}, \mathbf{r}, s, t)\rangle = \hat{H} |\Phi(\mathbf{R}, \mathbf{r}, s, t)\rangle, \quad (2.1)$$

where \hat{H} describes the Hamiltonian and Φ is the wave function of the system. In general, \hat{H} is the sum of the kinetic energy operator \hat{T} and the potential energy operator \hat{V} . The wave function Φ depends on all nuclear coordinates \mathbf{R} , all spatial coordinates of the electrons \mathbf{r} , the electron spin s , and in this case also on the time t . Further on, we combine \mathbf{r} and s in the electronic coordinates \mathbf{x} . Eq. 2.1 is called the *time-dependent* Schrödinger equation and therefore describes the time evolution of a given system. Note, that the Schrödinger equation is only valid in the non-relativistic limit. In order to describe relativistic effects, the Dirac equation [71] is needed. However, relativistic effects play only a negligible role for the systems discussed within this thesis, and therefore we will limit ourselves to the Schrödinger equation.

In most problems of interest to the fields of chemical physics and chemistry, the studied system is in a stationary state, i.e., all observables are independent of time. In this case, Eq. 2.1 can be simplified to the *time independent* Schrödinger equation, given by

$$\hat{H}\Phi(\mathbf{R}, \mathbf{x}) = E\Phi(\mathbf{R}, \mathbf{x}). \quad (2.2)$$

The now time-independent Hamiltonian operator \hat{H} yields the total energy of the system E , when it acts on the wave function. A state is *stationary* if the corresponding wave function

$$\Phi(\mathbf{R}, \mathbf{x}, t) = \Phi(\mathbf{R}, \mathbf{x}) \exp\left(\frac{-iEt}{\hbar}\right). \quad (2.3)$$

In this case, the probability density $|\Phi|^2$ becomes independent of time

$$|\Phi(\mathbf{R}, \mathbf{x}, t)|^2 = \Phi^*(\mathbf{R}, \mathbf{x})\Phi(\mathbf{R}, \mathbf{x}) \underbrace{\exp\left(\frac{iEt}{\hbar}\right) \exp\left(\frac{-iEt}{\hbar}\right)}_1 = |\Phi(\mathbf{R}, \mathbf{x})|^2. \quad (2.4)$$

However, note that the wave function itself is not stationary but continually changes its phase factor, forming a standing wave. Eq. 2.2 is an eigenvalue equation with the wave function being an eigenfunction of \hat{H} . This eigenvalue equation has an infinite number of solutions and the lowest obtainable energy describes the ground state of the considered system.

The Hamiltonian operator \hat{H} for a molecular or periodic system containing nuclei and electrons can be written as

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{nn} + \hat{V}_{en}, \quad (2.5)$$

with \hat{T}_e and \hat{T}_n describing the kinetic energy operator of the electrons and nuclei, respectively; \hat{V}_{ee} represents the repulsive interaction between electrons; \hat{V}_{nn} is the repulsive interaction between nuclei; \hat{V}_{en} describes the attractive Coulomb interaction between electrons and nuclei. In more detail, \hat{H} for a system containing N electrons and M nuclei is given by

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{a=1}^M \frac{1}{2m_a} \nabla_a^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{a=1}^M \sum_{b>a}^M \frac{Z_a Z_b}{R_{ab}} - \sum_{i=1}^N \sum_{a=1}^M \frac{Z_a}{r_{ia}}, \quad (2.6)$$

where the ordering of the terms is the same as in Eq. 2.5. Indices i and j run over electrons, while a and b refer to nuclei. The distance between two electrons is described by r_{ij} and the distance between two nuclei by R_{ab} ; r_{ia} is the electron-nucleus distance. The respective atomic number enters as Z and m_a refers to the ratio between the nucleus mass of nucleus a and the electron mass. The Coulomb interaction between electrons and nuclei can also be written

$$-\sum_{i=1}^N \sum_{a=1}^M \frac{Z_a}{r_{ia}} = \sum_{i=1}^N \nu(\mathbf{r}_i), \quad (2.7)$$

where $\nu(\mathbf{r}_i)$ describes an external potential acting on electron i due to all nuclei.

2.2 The Born-Oppenheimer Approximation

The wave function $\Phi(\mathbf{R}, \mathbf{x})$ is, even for small molecules, a high-dimensional object and an approximation is needed to reduce the computational complexity. We can utilize the fact that atomic nuclei are much heavier than electrons. The smallest and lightest nucleus consisting of a single proton is already 1836 times heavier than an electron. Therefore, nuclei move much slower than electrons. Born and Oppenheimer [72] proposed in 1927 that one should consider the dynamics of electrons within the field of fixed atomic nuclei. This approach is often an excellent approximation and is referred to as the *Born-Oppenheimer approximation*. Therein, we neglect the kinetic energy of the nuclei and the repulsion between nuclei is described as constant. Since the addition of a constant to an operator does only add to the eigenvalues and does not modify the eigenfunctions, we do not consider the nucleus-nucleus repulsion in the Hamiltonian. Therefore, the remaining Hamiltonian describes electrons moving in the field of point charges. We call this Hamiltonian the *electronic Hamiltonian* \hat{H}_{elec} , which is given by

$$\hat{H}_{\text{elec}} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} - \sum_{i=1}^N \sum_{a=1}^M \frac{Z_a}{r_{ia}}. \quad (2.8)$$

The corresponding Schrödinger equation,

$$\hat{H}_{\text{elec}} \Phi_{\text{elec}} = E_{\text{elec}} \Phi_{\text{elec}}, \quad (2.9)$$

yields the electronic wave function Φ_{elec} and the corresponding electronic energy E_{elec} . Φ_{elec} describes the motion of the electrons and therefore explicitly depends on electronic spatial coordinates \mathbf{r} . However, Φ_{elec} also parametrically depends on nuclear coordinates \mathbf{R} , which is also true for E_{elec} . The parametric dependence means that Φ_{elec} will always be a different function of electronic coordinates for different nuclear coordinates. The total energy of a system within the Born-Oppenheimer approximation E_{tot} is given by the sum of the electronic energy and the nuclear repulsion

$$E_{\text{tot}} = E_{\text{elec}} + \sum_{a=1}^M \sum_{b>a}^M \frac{Z_a Z_b}{R_{ab}}. \quad (2.10)$$

2.3 The Pauli Exclusion Principle

As already mentioned in section 2.1, the wave function depends on electronic coordinates \mathbf{x} , which consist of the spatial coordinates \mathbf{r} and also the spin s . However, we have not considered the electron spin yet, since the non-relativistic \hat{H}_{elec} depends only on \mathbf{r} . The Pauli exclusion principle [73] states that two electrons cannot simultaneously occupy the same quantum state. This is a result of the so-called *antisymmetry principle*, which states that the total electronic wave function must be antisymmetric w.r.t. the exchange

of any two electrons

$$\Phi_{\text{elec}}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Phi_{\text{elec}}(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N). \quad (2.11)$$

Therefore, our exact electronic wave function must not only satisfy the Schrödinger equation but also the antisymmetry principle.

We specify the electron spin s by introducing the spin functions $\alpha(\omega)$ and $\beta(\omega)$ for spin up and down, respectively. ω refers in this context to an unspecified spin variable. We only require from the spin functions that they are complete and orthonormal with

$$\int d\omega \alpha^*(\omega)\alpha(\omega) = \int d\omega \beta^*(\omega)\beta(\omega) = 1 \quad (2.12)$$

and

$$\int d\omega \alpha^*(\omega)\beta(\omega) = \int d\omega \beta^*(\omega)\alpha(\omega) = 0. \quad (2.13)$$

2.4 The Wave Function of an Electron

The spatial distribution of an electron is described by a spatial orbital $\psi(\mathbf{r})$ and the probability P of finding the electron within a volume element $d\mathbf{r}$ around the coordinates \mathbf{r} is given by

$$P = |\psi(\mathbf{r})|^2 d\mathbf{r}. \quad (2.14)$$

If we add now the information about the spin to the spatial orbital, the resulting wave function is called a *spin orbital* $\chi(\mathbf{x})$, given by

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \psi(\mathbf{r})\beta(\omega) \end{cases} \quad (2.15)$$

Therefore, for every spatial orbital one can build two different spin orbitals.

Spatial orbitals centered on an atomic nucleus are called atomic orbitals (AOs) and can be represented by certain *basis functions*. In general, the basis expansion has the form

$$\psi(\mathbf{r}) = \sum_{i=1}^K c_i \phi_i(\mathbf{r}), \quad (2.16)$$

with $\phi_i(\mathbf{r})$ being the individual basis functions within the applied basis set and c_i are constant coefficients. The number K determines when the basis expansion is truncated, i.e. how many basis functions are used. An exact representation of $\psi(\mathbf{r})$ is only possible if the basis set is complete, which would require an infinite number of basis functions. Therefore, we always have to work with approximate representations and the size of the used basis set determines among other things the numerical accuracy of the calculations. In quantum chemistry mostly atom-centered Gaussian-type basis functions are

used, while most solid-state physics codes rely on so-called plane waves. In this work, we will use throughout numerical atom-centered basis functions, which are described in detail in Ref. 74. For the description of electrons within a molecular systems, so-called *molecular orbitals* (MOs) are used. They are typically obtained by a linear combination of atomic orbitals (LCAO).

2.5 Many-Electron Wave Functions

Since we have established the wave function of a single electron, we will now move on to the wave function of an N -electron system. First, we assume that we are dealing with a system containing N electrons, which do not interact with each other. The corresponding Hamiltonian of a non-interacting system is then just the sum of independent single-electron hamiltonians $\hat{h}(i)$ and can be expressed as

$$\hat{H}_{\text{non-int}} = \sum_{i=1}^N \hat{h}(i) = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 + \sum_{i=1}^N \nu(\mathbf{r}_i). \quad (2.17)$$

These single-electron operators $\hat{h}(i)$ describe the kinetic and potential energy of the corresponding electron i , and we completely neglect any electron-electron repulsion effects. Hence, the wave function of such a non-interacting system (Ψ^{HP}) can simply be a product of spin orbitals.

$$\Psi^{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \cdots \chi_N(\mathbf{x}_N) \quad (2.18)$$

This many-electron wave function is called a *Hartree product*. The corresponding eigenvalue amount to the sum of the individual spin orbital energies. However, this approach has one major drawback, even within the independent-particle approximation. In the Hartree product, electrons are not indistinguishable, i.e., electron number n will always occupy spin orbital χ_n . Hence, the Hartree product violates the antisymmetry principle described above.

However, it is possible to create a linear combination of Hartree products, which fulfills the antisymmetry principle. This antisymmetrized product is called a *Slater determinant* $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ and can be written as

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}, \quad (2.19)$$

where the factor $1/\sqrt{N!}$ provides for the proper normalization of the wave function. In this approach we have now N electrons, which occupy N spin orbitals but it is not specified which electron occupies which orbital. If we now interchange the coordinates of two electrons, the sign of the determinant changes. In case two electrons would occupy

the same spin orbital, the determinant would become zero. Therefore, this approach now completely satisfies the Pauli exclusion principle.

While the Hartree product is a completely independent-electron wave function, the antisymmetrization leading to the Slater determinant introduces so-called *exchange-correlation* effects. This means that now two electrons with the same spin are correlated. However, the motion of electrons having opposite spin, remains uncorrelated.

2.6 The Hartree-Fock Method

Now we discuss the Hartree-Fock (HF) method, which aims to approximately solve the electronic Schrödinger equation, and is the general starting point for more advanced wave-function based approaches. In the HF approximation we utilize the above discussed Slater determinant. The goal is to find the best possible approximation for the true electronic wave function, which can be described by using a single Slater determinant. According to the variational principle, the best wave function within the used functional form yields the lowest possible energy E_0 , determined by

$$E_{\text{trial}} = \langle \Psi_{\text{trial}} | \hat{H}_{\text{elec}} | \Psi_{\text{trial}} \rangle \geq \langle \Psi_0 | \hat{H}_{\text{elec}} | \Psi_0 \rangle = E_0. \quad (2.20)$$

Therefore, we start with a trial wave function Ψ_{trial} and minimize E_{trial} until E_0 is reached by varying the spin orbitals. The optimal spin orbitals can be obtained via the so-called *Hartree-Fock equations*, which have the general form

$$\left[\hat{h}(\mathbf{x}_i) + \sum_b \hat{J}_b(\mathbf{x}_i) - \sum_b \hat{K}_b(\mathbf{x}_i) \right] \chi_a(\mathbf{x}_i) = \epsilon_a \chi_a(\mathbf{x}_i), \quad (2.21)$$

where ϵ_a is the energy corresponding to spin orbital χ_a . The expression within the square bracket represents the effective one-electron *Fock operator* \hat{F} , which itself consists of the one-electron operator \hat{h} , the *Coulomb operator* \hat{J}_b , and the *Exchange operator* \hat{K}_b . The latter two operators are defined below by their effect on orbital χ_a :

$$\hat{J}_b(\mathbf{x}_i) \chi_a(\mathbf{x}_i) = \left[\int d\mathbf{x}_j \chi_b^*(\mathbf{x}_j) \frac{1}{r_{ij}} \chi_b(\mathbf{x}_j) \right] \chi_a(\mathbf{x}_i), \quad (2.22)$$

$$\hat{K}_b(\mathbf{x}_i) \chi_a(\mathbf{x}_i) = \left[\int d\mathbf{x}_j \chi_b^*(\mathbf{x}_j) \frac{1}{r_{ij}} \chi_a(\mathbf{x}_j) \right] \chi_b(\mathbf{x}_i). \quad (2.23)$$

In the HF approximation we replace the many-electron problem essentially by a one-electron problem, where we treat the electron-electron interaction in an average way. The HF potential $v^{\text{HF}}(\mathbf{x}_i)$, given by

$$v^{\text{HF}}(\mathbf{x}_i) = \sum_b \left(\hat{J}_b(\mathbf{x}_i) - \hat{K}_b(\mathbf{x}_i) \right), \quad (2.24)$$

describes the average potential, which is experienced by electron i due to the other electrons in the system. Thus, our electrons move essentially independently within the mean field of all other electrons. As can be seen in the equations above, $v^{\text{HF}}(\mathbf{x}_i)$ depends on all other spin orbitals. Therefore, it is only possible to solve the HF equations iteratively, which is called the *self-consistent field* (SCF) approach. Therein, we start with an initial guess for the spin orbitals in order to calculate the HF potential. Solving the HF equations yields then a new set of spin orbitals, which we utilize as input for the next iteration. This iterative procedure is continued until so-called *self-consistency* is reached. This is the case when the field is no longer changing, i.e. when the spin orbitals used for the construction of \hat{F} are essentially identical with the eigenfunctions of \hat{F} .

Since we treat the electron-electron interaction here only in an average way, the HF method is for example not able to describe dispersion interactions. Typically, the so-called *correlation energy* is defined as the difference between the true ground-state energy of a system and the HF solution. The HF method serves as starting point for a variety of more accurate *post-Hartree-Fock* methods. One way of improving upon HF is the *configuration interaction* (CI) method, which utilizes a linear combination of the HF Slater determinant with excited Slater determinants. The wave function can be written in a CI expansion according to

$$|\Phi\rangle = c_0|\Psi_0\rangle + \sum_{a,r} c_a^r |\Psi_a^r\rangle + \sum_{a<b} \sum_{r<s} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \dots, \quad (2.25)$$

where a and b denote indices referring to occupied orbitals while the indices r and s refer to virtual orbitals according to the HF description. We successively replace occupied orbitals by virtual orbitals and the corresponding contribution to the total wave function is described by the mixing coefficients c . When all possible excited Slater determinants are considered, this approach is called *full-CI* [14, 75]. At the complete basis set limit, full-CI provides the exact energy of a system (in the non-relativistic limit). Unfortunately, due to the computational complexity, this approach can only be used for small molecules. More practically and commonly used wave-function-based methods for the calculation of the correlation energy are for example Møller-Plesset perturbation theory [76] and the Coupled Cluster approach [26, 77].

Chapter 3

Density Functional Theory

In the previous chapter we have discussed the basics of wave-function based electronic structure theory. We have seen that the wave function is a very high-dimensional object with $3N$ spatial coordinates plus N spin functions. However, it is possible to reformulate the Schrödinger equation based on a much simpler, only three-dimensional quantity — the electron density. The one-particle density $\rho(\mathbf{r})$ (electron density) describes the number of electrons within a specific unit volume and can be evaluated at the position \mathbf{r}_1 as

$$\rho(\mathbf{r}_1) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)|^2 ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N. \quad (3.1)$$

This chapter provides an overview of the basics of density functional theory (DFT) and is based on the book by Parr & Yang [78] and Ref. 79. All equations within this chapter are expressed in atomic units and the discussion is limited to the spin-unpolarized case.

One practical drawback of DFT is the fact that common density functional approximations (DFAs) are unable to properly describe long-range correlation effects. This issue has long prevented DFT to be used for weakly-bound systems. However, these effects can be included by combining DFAs with appropriate models for van-der-Waals (vdW) dispersion interactions. This development lead to the broad applicability and the major success of DFT in recent years. Therefore, the second part of this chapter gives a short overview of dispersion models including a detailed description of the two methods used throughout this thesis.

3.1 The Hohenberg-Kohn Theorems

DFT is based on the Hohenberg-Kohn theorems [80]. It is claimed that the external potential $\nu(\mathbf{r})$ of a given system can — within an additive constant — be determined by the ground-state electron density $\rho(\mathbf{r})$. The knowledge of $\rho(\mathbf{r})$ enables the determination of the ground-state wave function and the total energy of a given system can be expressed as

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{ne}[\rho] = \int \rho(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + F[\rho], \quad (3.2)$$

where $F[\rho]$ describes the universal Hohenberg-Kohn functional, which is given by

$$F[\rho] = T[\rho] + V_{ee}[\rho]. \quad (3.3)$$

The second Hohenberg-Kohn theorem [80] established the density variation principle, which can be written as

$$E_0 = F[\rho] + \int \rho(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} \leq F[\rho'] + \int \rho'(\mathbf{r})\nu(\mathbf{r})d\mathbf{r}. \quad (3.4)$$

The exact ground-state energy E_0 , which corresponds to the ground-state density $\rho(\mathbf{r})$, is the lowest obtainable energy. This means that the energy corresponding to any trial density $\rho'(\mathbf{r})$ is guaranteed to be larger than or equal to E_0 . The trial density $\rho'(\mathbf{r})$ has to satisfy

$$\int \rho(\mathbf{r})d\mathbf{r} = N, \quad (3.5)$$

with N being the number of electrons. These two theorems show that all ground-state properties of a certain system could theoretically be determined by density functional theory. However, the exact Hohenberg-Kohn density functional $F[\rho]$ remains unfortunately unknown.

3.2 Kohn-Sham DFT

In 1965, Kohn and Sham addressed this issue by proposing the usage of a hypothetical non-interacting reference system [81]. They demanded that this non-interacting system has exactly the same ground-state electron density as the corresponding interacting system. The Hamiltonian of this non-interacting system \hat{H}_s is given by

$$\hat{H}_s = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_i^N \nu(\mathbf{r}_i). \quad (3.6)$$

It can be seen that we do not account here for the electron-electron repulsion. This approach has the advantage that the corresponding wave function Ψ_s is — just like in Hartree-Fock theory — a simple Slater determinant (see Eq. 2.19). Therefore, we can calculate the kinetic energy $T_s[\rho]$ according to

$$T_s[\rho] = \langle \Psi_s | -\frac{1}{2} \sum_i^N \nabla_i^2 | \Psi_s \rangle = -\frac{1}{2} \sum_i^N \langle \psi_i | \nabla^2 | \psi_i \rangle. \quad (3.7)$$

Next, we can redefine $F[\rho]$ as

$$F[\rho] = T[\rho] + V_{ee}[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho], \quad (3.8)$$

with $J[\rho]$ expressing the *self-interaction* energy of the electron cloud, which can be calculated according to

$$J[\rho] = \frac{1}{2} \int \int \frac{1}{|\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}) \rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \quad (3.9)$$

Then, we define a new energy term called *exchange-correlation* energy $E_{xc}[\rho]$, which gathers all still missing energy contributions as

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]. \quad (3.10)$$

However, no exact form for $E_{xc}[\rho]$ is known. Therefore, finding the exact exchange-correlation functional has become a central task within DFT. Many approximations of $E_{xc}[\rho]$ have been developed and several examples will be discussed below. Assuming the knowledge of $E_{xc}[\rho]$, the total energy $E[\rho]$ can now be written as

$$E[\rho] = \int \nu(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} + T_s[\rho] + J[\rho] + E_{xc}[\rho]. \quad (3.11)$$

This leads then to the so-called Kohn-Sham (KS) equations

$$\left(\nu(\mathbf{r}) - \frac{1}{2} \nabla^2 + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \nu_{xc}(\mathbf{r}) \right) \psi_i = \epsilon_i \psi_i \quad (3.12)$$

where

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2, \quad (3.13)$$

and

$$\nu_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \quad (3.14)$$

The KS equations can be solved iteratively by using the SCF approach, as described in the previous chapter.

3.3 Density-Functional Approximations

Now, we discuss several developed density-functional approximations (DFAs). The first and simplest approximation for the exchange-correlation functional ($E_{xc}[\rho]$) is the so-called *local density approximation* (LDA) [81], given by

$$E_{xc}^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r})) d\mathbf{r}, \quad (3.15)$$

with $\epsilon_{xc}(\rho(\mathbf{r}))$ describing the exact exchange-correlation energy within a homogeneous electron gas. While this approximation yields reasonable results for systems with slowly-varying electron density — like metallic systems — the accuracy is most often insufficient for molecules and vdW-bound system.

The next approximation is the so-called *generalized gradient approximation* (GGA), given by

$$E_{xc}^{\text{GGA}}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho) d\mathbf{r} = \int \rho(\mathbf{r})\epsilon_{xc}(\rho(\mathbf{r}))F_{xc}(\rho(\mathbf{r}), \nabla\rho) d\mathbf{r}. \quad (3.16)$$

Therein, (some) non-local effects are introduced by making $E_{xc}[\rho]$ also dependent on the first derivative of the density $\nabla\rho$. E_{xc}^{GGA} can be formulated in terms of E_{xc}^{LDA} and an additional dimensionless *enhancement factor* F_{xc} . This approximation lead to a variety of different DFAs, called semi-local functionals. One very commonly used GGA functional — and the one considered within this thesis — was developed by Perdew, Burke, and Ernzerhof [82], and is therefore termed the *PBE functional*. This functional is considered to be a non-empirical functional and satisfies several exact conditions. GGA functionals yield significantly better results than the LDA but still underestimate binding energies in weakly-bound systems.

Subsequently, so-called *meta-GGA* functionals were developed, which take also the second derivative of the density into account. While GGA and meta-GGA functionals provide already quite good results for a variety of systems [83], the description of the exchange still has one fundamental problem; every electron experiences a spurious interaction with itself, which is called *self-interaction error*. However, the exchange is described correctly in the previously discussed Hartree-Fock method. This realization lead to the development of so-called *hybrid functionals*, which incorporate a certain fraction of exact exchange (E_x) analogous to Hartree-Fock theory. The exchange-correlation energy can then be written as

$$E_{xc}^{\text{Hybrid}}[\rho] = \alpha E_x[\rho] + (1 - \alpha)E_x^{\text{GGA}}[\rho] + E_c^{\text{GGA}}[\rho], \quad (3.17)$$

where the amount of included exact exchange is controlled with the factor α . Within this thesis the PBE0 hybrid functional [84] is used. This method relies on PBE for the GGA part and uses 25% of exact exchange, i.e. $\alpha = 0.25$. It will be shown later on that PBE0 yields almost consistently better results than PBE but the calculation of the exact exchange also significantly increases the computational cost.

3.4 Van der Waals Dispersion Models

Traditional DFAs — like the ones discussed above — are not able to correctly capture long-range electron-correlation effects in molecular systems. Therefore, the missing dispersion energy E_{disp} is often determined *a posteriori* and subsequently added to the total energy of the used DFA. This approach is termed vdW-inclusive DFT. Most atom-pairwise dispersion models have the general form

$$E_{\text{disp}} = - \sum_{A>B} f_{\text{dmp}} \frac{C_{AB}^6}{R_{AB}^6}, \quad (3.18)$$

with R_{AB} being the interatomic distance, C_6^{AB} being the dipole-dipole dispersion coefficient between the atoms A and B , and f_{dmp} refers to a damping function. A detailed discussion about such damping functions can be found in Ref. 85. Such an approach is used for instance in DFT-D2 [31] and the Tkatchenko-Scheffler (TS) [34] vdW model. D2 utilizes fixed empirical dispersion coefficients while the TS model obtains them *on the fly* from the electron density.

Dispersion interaction can in general be expressed in a multipole expansion [11]. However, the so far mentioned D2 and TS models contain only dipole-dipole interactions. Hence, the qualitative description could be improved by including dipole-quadrupole interactions, which lead to an additional C_8^{AB}/R_{AB}^8 term. Such an approach is used within the DFT-D3 method [32], in which the dispersion coefficients depend now on atomic coordination numbers. Another approach for the calculation of dispersion interactions is the so-called exchange-dipole moment (XDM) model [33, 86, 87], which also depends on the electron density. Therein, pairwise dispersion interactions are considered up to quadrupole-quadrupole interactions.

However, dispersion interactions are in general not pairwise additive and consideration of many-body interactions can be important [88]. The DFT-D3 model approximates such many-body interactions by using a three-body Axilrod-Teller-Muto term to describe the interaction between three instantaneous dipoles. Furthermore, Tkatchenko *et al.* developed the many-body dispersion (MBD) model [35, 64]. Therein, dipolar many-body (many-atom) interactions are considered up to infinite order. In addition, MBD includes also electrodynamic response effects via a short-range self-consistent screening of atomic polarizabilities. Such effects can be crucial for molecular crystals since the presence of the crystal field leads to significant changes in the polarizabilities compared to isolated molecules. [13]

Within this thesis, the TS and MBD dispersion models will be used throughout. Therefore, these two models are described in detail below. It is noted in passing that the TS and MBD models are explicit density functionals, which enables the investigation of vdW effects for a variety of properties beyond structures and stabilities [15, 89].

3.4.1 Tkatchenko-Scheffler Dispersion Model

The dispersion energy within the TS model [34] is calculated in an effective-pairwise fashion using the general expression shown in Eq. 3.18. However, the dispersion parameters for each atom A are obtained from the electron density, as determined during a DFT calculation. Therefore, an effective atom-in-a-molecule volume V_A^{eff} is determined via a Hirshfeld [90] partitioning of the electron density, which is further on compared to a free-atom volume V_A^{free} . The resulting effective atomic dipole polarizability α_A^{eff} , dispersion coefficient $C_{6,A}^{\text{eff}}$, and vdW radius $R_A^{0,\text{eff}}$ can be calculated according to

$$\alpha_A^{\text{eff}} = \left(\frac{V_A^{\text{eff}}}{V_A^{\text{free}}} \right) \alpha_A^{\text{free}}, \quad (3.19)$$

$$C_{6,A}^{\text{eff}} = \left(\frac{V_A^{\text{eff}}}{V_A^{\text{free}}} \right)^2 C_{6,A}^{\text{free}}, \quad (3.20)$$

$$R_A^{0,\text{eff}} = \left(\frac{V_A^{\text{eff}}}{V_A^{\text{free}}} \right)^{1/3} R_A^{0,\text{free}}. \quad (3.21)$$

The corresponding free-atom values of these three quantities are tabulated for (almost) every element. The interatomic dispersion coefficient C_6^{AB} between atoms A and B is obtained via the following relation (the *eff* labels are subsequently dropped for clarity):

$$C_6^{AB} = \frac{2C_6^{AA}C_6^{BB}}{\frac{\alpha_B^0}{\alpha_A^0}C_6^{AA} + \frac{\alpha_A^0}{\alpha_B^0}C_6^{BB}}. \quad (3.22)$$

Finally, the damping function f_{dmp} is calculated in a Fermi-type way [91] according to

$$f_{\text{dmp}}(R_{AB}, R_{AB}^0) = \frac{1}{1 + \exp \left[-20 \left(\frac{R_{AB}}{s_R R_{AB}^0} - 1 \right) \right]}, \quad (3.23)$$

where

$$R_{AB}^0 = R_A^0 + R_B^0, \quad (3.24)$$

and s_R is a single empirical parameter determining the onset of the dispersion correction in terms of the distance, which is necessary to couple the dispersion model to any density functional [92]. This parameter was determined using the S22 benchmark set [28] and amounts for PBE and PBE0 to 0.94 and 0.96, respectively [34].

3.4.2 Many-Body Dispersion Model

The adiabatic connection fluctuation-dissipation theorem [93, 94] provides a framework, in which the exact electron correlation energy can be expressed in terms of density-density response functions. The MBD model [35, 64] uses this framework to describe the dipole response of atoms within the random-phase approximation. The resulting MBD dispersion energy (long-range correlation energy) amounts to

$$E_{\text{MBD}} = \frac{1}{2\pi} \int_0^\infty d\omega \text{Tr} [\ln(\mathbf{1} - \mathbf{A}\mathbf{T})], \quad (3.25)$$

with \mathbf{A} being a diagonal $3n \times 3n$ matrix containing isotropic atomic polarizabilities, where n refers to the number of atoms; \mathbf{T} describes the dipole-dipole coupling tensor. It was shown that this expression is equivalent to the diagonalization of a model Hamiltonian for coupled isotropic quantum harmonic oscillators [64].

The MBD method accounts for many-body dispersion effects but also for self-consistent dielectric screening effects, i.e. a modification of the interaction between atoms due to the presence of additional atoms or polarizable centers. MBD calculations involve three

separate steps and are done in a range-separated way. First, atomic polarizabilities are obtained via the TS approach described above. In a second step, screened atomic polarizabilities are calculated for the short range via a range-separated self-consistent screening (rsSCS) with

$$\boldsymbol{\alpha}^{\text{rsSCS}}(i\omega) = \boldsymbol{\alpha}(i\omega) - \boldsymbol{\alpha}(i\omega)\mathbf{T}^{\text{SR}}\boldsymbol{\alpha}^{\text{rsSCS}}(i\omega), \quad (3.26)$$

where $\boldsymbol{\alpha}(i\omega)$ describes the diagonal polarizability matrix derived from the TS approach and $\boldsymbol{\alpha}^{\text{rsSCS}}(i\omega)$ is the resulting screened polarizability matrix. The frequency dependence for an atom A within $\boldsymbol{\alpha}(i\omega)$ is given by

$$\alpha_A(i\omega) = \frac{\alpha_A^0}{1 + \left(\frac{\omega}{\omega_A}\right)^2}, \quad (3.27)$$

with α_A^0 and ω_A being the static polarizability and characteristic frequency, respectively, as determined within the TS approach [34]. The dipole-dipole coupling tensor $\mathbf{T}_{\mathbf{AB}}$ between atoms A and B within the MBD model is given by

$$\begin{aligned} \mathbf{T}_{\mathbf{AB}} &= \nabla_{\mathbf{R}_A} \otimes \nabla_{\mathbf{R}_B} \frac{\text{erf}(R_{AB}/\sigma_{AB})}{R_{AB}} = -\frac{3R_i R_j - R_{AB}^2 \delta_{ij}}{R_{AB}^5} \\ &\times \left[\text{erf}\left(\frac{R_{AB}}{\sigma_{AB}}\right) - \frac{2R_{AB}}{\sqrt{\pi}\sigma_{AB}} e^{-(R_{AB}/\sigma_{AB})^2} \right] + \frac{4R_i R_j}{\sqrt{\pi}\sigma_{AB}^3 R_{AB}^2} e^{-(R_{AB}/\sigma_{AB})^2}, \end{aligned} \quad (3.28)$$

with

$$\sigma_{AB} = \sqrt{\sigma_A^2 + \sigma_B^2}, \quad (3.29)$$

where

$$\sigma_A = \left(\frac{\sqrt{2}\alpha_A}{3\sqrt{\pi}} \right)^{1/3}. \quad (3.30)$$

The indices i and j indicate the Cartesian components of $\mathbf{R}_{\mathbf{AB}}$. Now, the short-range dipole-dipole tensor is obtained via

$$\mathbf{T}_{\mathbf{AB}}^{\text{SR}} = (1 - f_{\text{dmp}}(R_{AB}))\mathbf{T}_{\mathbf{AB}} \quad (3.31)$$

with the damping function

$$f_{\text{dmp}} = \frac{1}{1 + \exp[-6(R_{AB}/S - 1)]}, \quad (3.32)$$

where

$$S = s_R(R_A^0 + R_B^0). \quad (3.33)$$

In this case, the values for the scaling parameter s_R amount for PBE and PBE0 to 0.83 and 0.85, respectively [64].

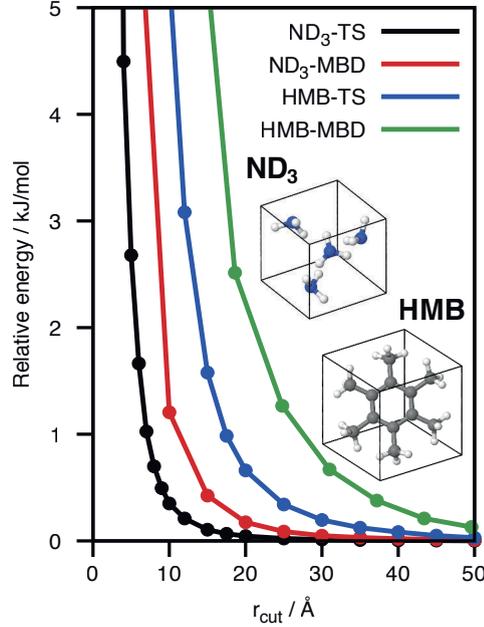


Figure 3.1: Convergence of TS and MBD vdW lattice energies w.r.t. a dipole-dipole cut-off radius and a MBD supercell cut-off radius for a ND_3 and HMB crystal. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

The so obtained screened polarizabilities are now utilized as input for step 3, which calculates the corresponding long-range correlation energy according to

$$E_{\text{MBD@rsSCS}} = \frac{1}{2\pi} \int_0^\infty d\omega \text{Tr} [\ln(\mathbf{1} - \mathbf{A}\mathbf{T}^{\text{LR}})]. \quad (3.34)$$

In this case only the long-range dipole-dipole tensor \mathbf{T}^{LR} is used, in order to avoid double-counting effects [64]. \mathbf{T}^{LR} is given by

$$\mathbf{T}^{\text{LR}} = \mathbf{T} - \mathbf{T}^{\text{SR}}, \quad (3.35)$$

but is in the MBD model approximated by its long-range limit, amounting between atoms A and B to

$$\mathbf{T}_{\text{AB}}^{\text{LR}} = -\frac{3R_i R_j - R_{AB}^2 \delta_{ij}}{R_{AB}^5}, \quad (3.36)$$

with i and j referring to the Cartesian components of \mathbf{R}_{AB} . Furthermore, it is also possible to decompose the MBD energy into orders n , i.e. n -atom contributions [95, 96], according to

$$E_n = -\frac{1}{2\pi} \int_0^\infty d\omega \sum_{n=2}^\infty \frac{1}{n} \text{Tr} [(\mathbf{A}\mathbf{T})^n]. \quad (3.37)$$

For a more in-depth discussion of the MBD method the reader is referred to Refs. 64, 88, 96, 97.

In molecular crystals, dispersion interactions can extend over a large distance, especially when described with MBD. We have illustrated this in Ref. 13 using a cubic deuterio-ammonia (ND_3) and a hexamethylbenzene (HMB) crystal (see Fig. 3.1). The convergence of the vdW energy normalized per molecule is plotted w.r.t. a cut-off radius r_{cut} . The completely converged energy defines the zero of the energy. In the case of TS, interactions are only considered for interatomic distances less than r_{cut} . In the case of MBD, interactions between two atoms are only included in the dipole-dipole tensor if the interatomic distance is less than r_{cut} . Moreover, we consider the coupling between the harmonic oscillators within a cubic supercell with lattice constant $a = r_{\text{cut}}$. It can be clearly seen in Fig. 3.1 that the MBD energies converge significantly slower than the TS energies. For instance, converging the lattice energy of HMB to 0.5 kJ/mol requires a distance of about 20 Å for TS and already a distance of around 33 Å for MBD. Both considered molecular crystals are quite symmetric. For less symmetric and more complex molecular crystals, we expect an even longer range for MBD interactions, as for instance shown for nanostructured materials in Refs. 98, 99.

Chapter 4

Lattice Vibrations

Up to this point we have only discussed the total energy of a system in terms of the electronic energy and the nuclear repulsion. In these cases we assumed that all the atomic positions are fixed. However, thermodynamic fluctuations always create displacements from the equilibrium positions of the atoms, i.e. the atoms vibrate around their equilibrium positions. As we will see later on, these displacements are even present at a temperature of 0 K due to quantum-mechanical effects. We will see that the energy of lattice vibrations is quantized and the corresponding quantum is called *phonon*. Accounting for this dynamics due to lattice vibrations is for instance essential for the description of free energies and vibrational spectra. In this chapter, we will first discuss the concept of the so-called *harmonic approximation* and the related harmonic oscillator model system. Then, we address also anharmonic effects by introducing the Morse oscillator. Furthermore, we focus on the understanding of normal/phonon modes, which will be illustrated for a few simple systems. This discussion and subsequent chapters require a basic understanding of crystal lattices and the concept of the reciprocal lattice, as for instance described in the book by Ashcroft & Mermin [100]. The content of this chapter is based on the textbooks by Ashcroft & Mermin [100], Hofmann [101], Kittel [102], McQuarrie [65], and Piela [66].

4.1 The Harmonic Approximation

In principle, a complete description of a given system would require the knowledge of the total potential-energy surface (PES), which is a $3n$ -dimensional object, with n being the number of atoms in the considered molecular system or in the unit cell. One very common approximation to this high-dimensional problem is the *harmonic approximation* (HA). Therein, we express the PES around the equilibrium geometry (with coordinates \mathbf{R}_{eq}) using a second-order Taylor expansion, given by

$$E(\mathbf{R}_{\text{eq}} + \Delta\mathbf{R}) = E(\mathbf{R}_{\text{eq}}) + \sum_i \frac{\partial E}{\partial \mathbf{R}_i} \Delta\mathbf{R}_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E}{\partial \mathbf{R}_i \partial \mathbf{R}_j} \Delta\mathbf{R}_i \Delta\mathbf{R}_j + \mathcal{O}(\Delta\mathbf{R}^3), \quad (4.1)$$

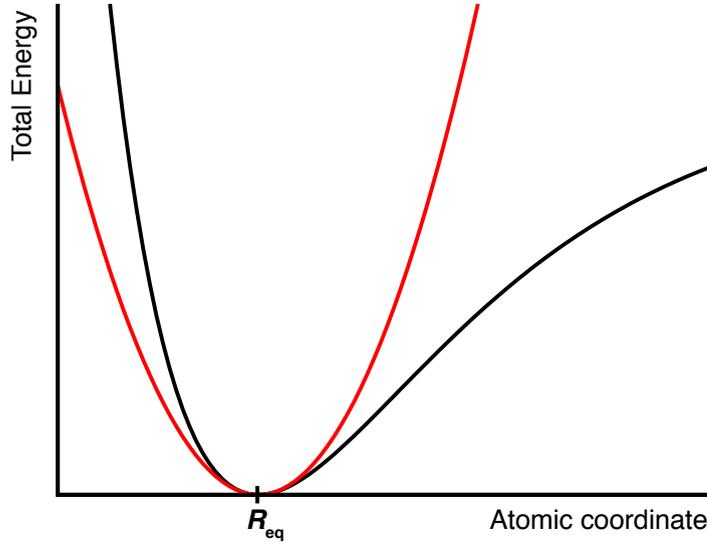


Figure 4.1: Illustration of the harmonic approximation. The actual potential curve of the total energy is illustrated in black, while the harmonic approximation around the equilibrium position (\mathbf{R}_{eq}) is shown in red.

where $E(\mathbf{R}_{\text{eq}})$ denotes the total (static) energy of the equilibrium geometry. This is illustrated in Fig. 4.1. The first-order term vanishes because the forces at a local minimum of the PES are zero. The second-order term leads to the harmonic force constants Φ_{ij}

$$\frac{\partial^2 E}{\partial \mathbf{R}_i \partial \mathbf{R}_j} = -\frac{\partial \mathbf{F}_j}{\partial \mathbf{R}_i} = \Phi_{ij}, \quad (4.2)$$

which form the so-called Hessian matrix, or in the case of periodic systems the *dynamical matrix* [100]

$$D_{ij}(\mathbf{q}) = \sum_{j'} \frac{e^{i(\mathbf{q} \cdot \mathbf{R}_{jj'})}}{\sqrt{m_i m_j}} \Phi_{ij}, \quad (4.3)$$

where m_i stands for the mass of atom i . All effects arising from the higher-order terms in Eq. 4.1 are called *anharmonic effects* and are neglected in the harmonic approximation.

The harmonic force constant Φ_{ij} describes how the force on atom j changes when atom i is displaced. Typically, the dynamical matrix is obtained numerically by finite differences, but it can also be calculated using density-functional perturbation theory [103]. The eigenvalues of $D_{ij}(\mathbf{q})$ are the phonon frequencies ω and the corresponding eigenvectors describe the phonon modes. In the case of non-periodic systems we speak of vibrational frequencies/modes. The complete dynamics of such a harmonic system is determined by a system of $3n$ independent quantum harmonic oscillators, one oscillator for each phonon mode. The harmonic Helmholtz free energy F^{HA} is given by

$$F^{\text{HA}}(T) = E_{\text{tot}} + F_{\text{vib}}^{\text{HA}}(T), \quad (4.4)$$

with

$$F_{\text{vib}}^{\text{HA}}(T) = \int d\omega g(\omega) \frac{\hbar\omega}{2} + \int d\omega g(\omega) k_{\text{B}}T \ln \left[1 - \exp\left(-\frac{\hbar\omega}{k_{\text{B}}T}\right) \right], \quad (4.5)$$

with $g(\omega)$ being the phonon density of states (pDOS), which describes the number of available vibrational states at a certain frequency. The pDOS can be calculated according to

$$g(\omega) = \sum_s \int \frac{d\mathbf{q}}{(2\pi)^3} \delta(\omega - \omega_s(\mathbf{q})). \quad (4.6)$$

The first integral in equation 4.5 describes so-called zero-point vibrations, which are always present in every quantum system — even at a temperature of 0K — while the second integral accounts for thermally induced vibrations including vibrational entropy. In the following, we discuss in detail the underlying harmonic oscillator model system.

4.2 The Classical Harmonic Oscillator

Before we move on to a quantum harmonic oscillator, we shall discuss the basics of a classical one-dimensional harmonic oscillator. Suppose we have a mass m , which is connected to a solid wall via a spring. We consider here only the force due to the spring, i.e. we neglect the gravitational force. The coordinate x describes the displacement of m w.r.t. its equilibrium position. Furthermore, we assume that the force (f) acting on m is directly proportional to the displacement x . This results in the so-called *Hooke's law*, given by

$$f = -kx, \quad (4.7)$$

where k is the positive proportionality constant, usually called *force constant*. Therefore, the corresponding Newton's equation of motion is

$$m \frac{d^2x}{dt^2} + kx = 0. \quad (4.8)$$

The general solution for this equation is given by

$$x(t) = c_1 \sin(\omega t) + c_2 \cos(\omega t), \quad (4.9)$$

with the angular frequency ω being

$$\omega = \sqrt{\frac{k}{m}}. \quad (4.10)$$

Assume the spring is initially stretched so that the displacement amounts to A and subsequently released. In this case, the initial velocity amounts to zero. This leads to

$$x(0) = A = c_1 \sin(0) + c_2 \cos(0) = c_2 \quad (4.11)$$

and

$$\left(\frac{dx}{dt}\right)_{t=0} = 0 = c_1\omega \cos(0) - c_2\omega \sin(0) = c_1\omega. \quad (4.12)$$

Therefore, $c_1 = 0$ and $c_2 = A$ in this case, and

$$x(t) = A \cos(\omega t). \quad (4.13)$$

This result illustrates that our mass m continuously oscillates between A and $-A$ with an angular frequency ω . The corresponding linear frequency $\nu = \omega/2\pi$.

The potential energy of the classical harmonic oscillator V can be obtained from the fact that the force is a derivative of the potential energy, which can be expressed as

$$f(x) = -\frac{dV}{dx}. \quad (4.14)$$

Therefore, $V(x)$ amounts to

$$V(x) = -\int f(x)dx + c = \frac{kx^2}{2} + c, \quad (4.15)$$

where c is an arbitrary constant. Typically, the equilibrium position is chosen to be the zero of the energy, which leads to $c = 0$. The kinetic energy T is

$$T = \frac{m}{2} \left(\frac{dx}{dt}\right)^2. \quad (4.16)$$

Hence, the total energy E of our classical harmonic oscillator amounts to

$$E = T + V = \frac{m\omega^2 A^2}{2} \sin^2(\omega t) + \frac{kA^2}{2} \cos^2(\omega t) = \frac{kA^2}{2} [\sin^2(\omega t) + \cos^2(\omega t)] = \frac{kA^2}{2}. \quad (4.17)$$

The total energy consists of potential and kinetic energy. As can be seen from the equation above, the total energy is constant, proving that the harmonic oscillator is a conservative system.

After discussing the basics of a general one-dimensional classical harmonic oscillator, we return now to our atomistic systems. But before we address polyatomic systems, we focus on the vibrations within a diatomic molecule. In this case we have two atoms with masses m_1 and m_2 , which are connected via a spring. The position of the atoms shall be x_1 and x_2 , respectively. The forces acting on the two individual atoms are equal but with opposite sign. The vibrational motion of our diatomic molecule depends only on the interatomic distance $x = x_2 - x_1$. In this case we have now two equations of motion

$$m_1 \frac{d^2 x_1}{dt^2} = kx \quad \text{and} \quad m_2 \frac{d^2 x_2}{dt^2} = -kx. \quad (4.18)$$

Adding both equations of motions leads to

$$\frac{d^2}{dt^2} (m_1 x_1 + m_2 x_2) = 0. \quad (4.19)$$

Introducing now a center-of-mass coordinate

$$X = \frac{m_1 x_1 + m_2 x_2}{M}, \quad \text{with } M = m_1 + m_2 \quad (4.20)$$

yields

$$M \frac{d^2 X}{dt^2} = 0. \quad (4.21)$$

This result illustrates that the center of mass moves with a constant momentum. Furthermore, subtracting both equations of motions and dividing by the respective masses leads to

$$\frac{d^2 x}{dt^2} = -kx \left(\frac{1}{m_1} + \frac{1}{m_2} \right). \quad (4.22)$$

Rewriting

$$\frac{1}{m_1} + \frac{1}{m_2} = \frac{m_1 + m_2}{m_1 m_2} = \frac{1}{\mu} \quad (4.23)$$

leads to

$$\mu \frac{d^2 x}{dt^2} + kx = 0, \quad (4.24)$$

where μ is the so-called *reduced mass*. This result is essentially equivalent to Eq. 4.10 except that the mass m is replaced by the reduced mass μ , leading to

$$\omega = \sqrt{\frac{k}{\mu}}. \quad (4.25)$$

The substitution with the reduced mass allows us to reduce the two-body problem to a simple one-body problem.

This concept can also be used for polyatomic and periodic systems by the introduction of so-called *normal-mode coordinates* [65]. For a system with N atoms this leads to $3N$ normal modes, which describe the motion of the system. For every mode we obtain an atomic displacement vector, which describes the corresponding atomic motion. The corresponding reduced mass can be calculated according to

$$\mu = \frac{1}{\sum_{i=1}^N \frac{1}{m_i}}. \quad (4.26)$$

4.3 The Quantum Harmonic Oscillator

Now, we discuss the description of vibrations using a quantum-harmonic oscillator. For simplicity, we return to our diatomic molecule described above. The Schrödinger equation is given by

$$-\frac{\hbar^2}{2\mu} \frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x), \quad (4.27)$$

where $V(x)$ is our harmonic potential

$$V(x) = \frac{1}{2}kx^2. \quad (4.28)$$

This leads to the following differential equation

$$\frac{d^2\psi}{dx^2} + \frac{2\mu}{\hbar^2} \left(E - \frac{1}{2}kx^2 \right) \psi(x) = 0. \quad (4.29)$$

Solving this equation leads to the energy of the quantum-harmonic oscillator, given by

$$E_v = \hbar\omega \left(v + \frac{1}{2} \right), \quad \text{with} \quad v = 0, 1, 2, \dots \quad (4.30)$$

Therefore, the energy of a quantum-harmonic oscillator is quantized, i.e. only discrete energy values are possible. An illustration of these energy levels is shown in Fig. 4.2. All energy levels are equally spaced and the energy difference between two levels amounts to $\hbar\omega$. The harmonic oscillator has an infinite number of vibrational levels and does not allow for dissociation. Note, that at the lowest level $v = 0$, the corresponding energy is not zero — as in the classical case — but $1/2\hbar\omega$. This is a result of the uncertainty principle and we call this energy the *zero-point energy*. Since $v = 0$ is the vibrational ground state, this vibrational energy is also present at zero temperature (at which an excitation to a higher level is not possible).

The corresponding wave functions of the harmonic oscillator have the following form:

$$\psi_v(x) = N_v H_v(\sqrt{\alpha}x) e^{-\alpha x^2/2} \quad (4.31)$$

with

$$\alpha = \sqrt{\frac{k\mu}{\hbar^2}} \quad (4.32)$$

and a normalization constant

$$N_v = \frac{1}{\sqrt{2^v v!}} \left(\frac{\alpha}{\pi} \right)^{1/4}. \quad (4.33)$$

The expression $H_v(\sqrt{\alpha}x)$ stands for the so-called *Hermite polynomials*. The first three Hermite polynomials are given by

$$H_0(\xi) = 1, \quad H_1(\xi) = 2\xi, \quad H_2(\xi) = 4\xi^2 - 2, \quad (4.34)$$

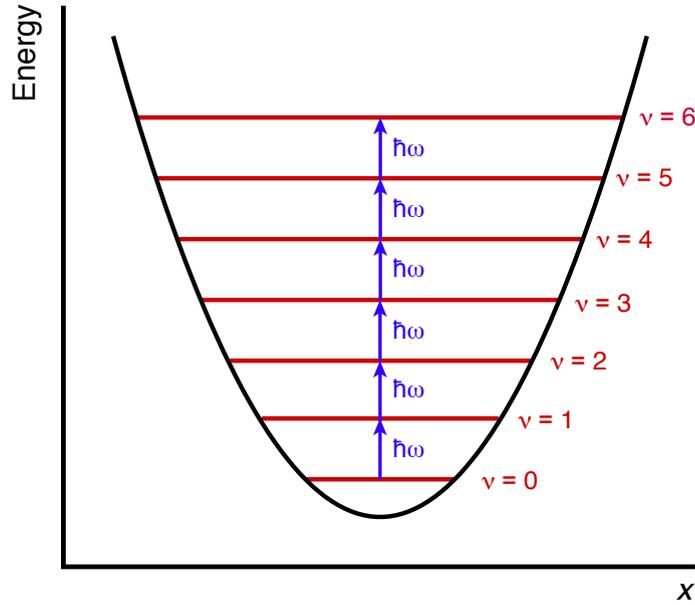


Figure 4.2: Illustration of a one-dimensional quantum harmonic oscillator.

where $\sqrt{\alpha}x$ has been replaced with ξ for simplicity. Therefore, the wave functions corresponding to the lowest three vibrational levels are

$$\psi_0 = \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2/2}, \quad \psi_1 = \left(\frac{4\alpha^3}{\pi}\right)^{1/4} x e^{-\alpha x^2/2}, \quad \psi_2 = \left(\frac{\alpha}{4\pi}\right)^{1/4} (2\alpha x^2 - 1) e^{-\alpha x^2/2}. \quad (4.35)$$

4.4 The Morse Oscillator

As mentioned at the beginning of this chapter, the harmonic approximation does not include any anharmonic effects, i.e. any higher-order terms in the discussed Taylor expansion. One quantum-mechanical model system, which is able to describe some anharmonic effects, is the *Morse oscillator* [104, 105]. The potential of a Morse oscillator is given by

$$V(x) = D_e \left(1 - e^{-a(x-x_0)}\right)^2, \quad (4.36)$$

with a , D_e , and x_0 being the width of the potential, the well depth, and the minimum of the potential, respectively. An illustration of a Morse oscillator is shown in Fig. 4.3. The Morse oscillator is a particularly useful model system since it is one of very few systems for which we know an analytical solution of the Schrödinger equation. For a

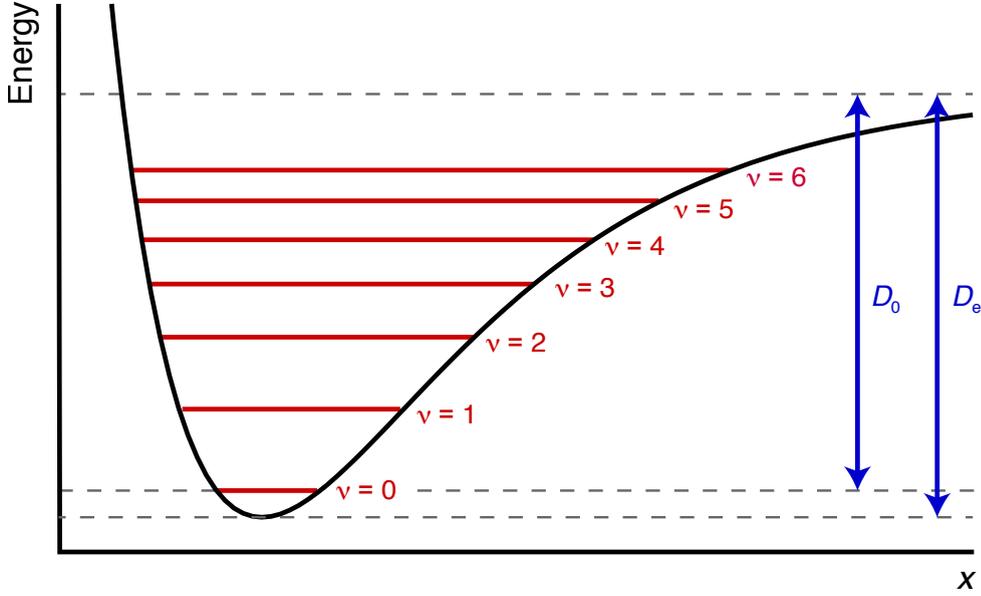


Figure 4.3: Illustration of a one-dimensional Morse oscillator.

vibrational mode in state v the energy can be calculated according to

$$E(v) = \hbar\omega_0 \left(v + \frac{1}{2} \right) - \frac{\hbar^2 \omega_0^2}{4D_e} \left(v + \frac{1}{2} \right)^2, \quad (4.37)$$

with

$$\omega_0 = \sqrt{\frac{2a^2 D_e}{\mu}}, \quad (4.38)$$

where μ describes the reduced mass. As can be seen in Fig. 4.3, the vibrational levels are now — in contrast to the harmonic oscillator — not evenly-spaced. The energy differences between them decrease with increasing v and the energy between two subsequent levels amounts to

$$E(v+1) - E(v) = \hbar\omega_0 - (v+1) \frac{\hbar^2 \omega_0^2}{2D_e}. \quad (4.39)$$

In contrast to the harmonic oscillator, the Morse oscillator allows for dissociation, i.e. there exists only a finite number of bound vibrational states. Note that the actual dissociation energy is not D_e but amounts to D_0 (see Fig. 4.3) due to the zero-point energy.

4.5 Phonon Modes and Dispersion in Simple Model Systems

4.5.1 Infinite Chain of Atoms with One Atom in the Unit Cell

Now we discuss the nature of phonon/normal modes and how their vibrational frequency varies in reciprocal space, which is called *phonon dispersion*. First, we consider a very

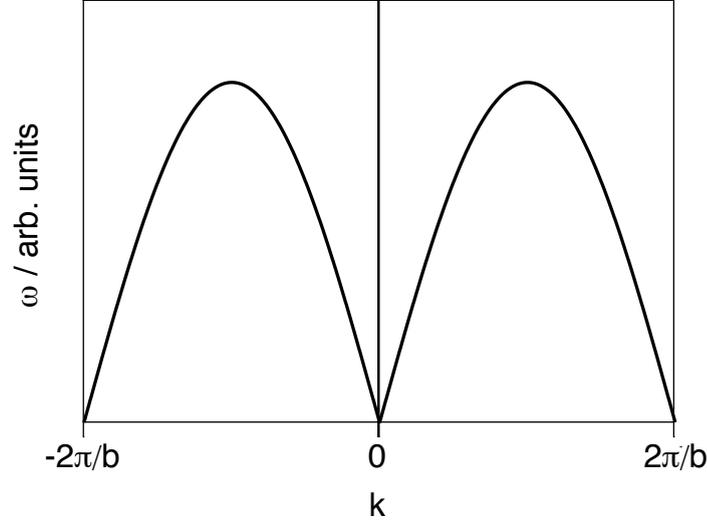


Figure 4.4: Phonon dispersion of a one-dimensional infinite chain of atoms with a single atom in the unit cell.

simple model system, which is a homogeneous infinite chain of atoms. This system is periodic in one dimension and has only one atom per unit cell. The atoms are separated by the cell length b and hence the length of the reciprocal space vector amounts to $2\pi/b$. All atoms are connected via springs with force constant γ and have a mass M . The location of atom n within the chain amounts to nb ($n = 1, 2, 3, \dots$). In this example we only take interactions between the nearest neighbors into account and express everything in terms of classical harmonic oscillators. Therefore, the equation of motion for an atom n is given by

$$M \frac{d^2 \xi_n}{dt^2} = -\gamma(\xi_n - \xi_{n-1}) + \gamma(\xi_{n+1} - \xi_n) = -\gamma(2\xi_n - \xi_{n-1} - \xi_{n+1}), \quad (4.40)$$

where ξ describes the displacement of the respective atom. The corresponding solution has the following form

$$\xi_n(t) = A e^{i(kbn - \omega t)}, \quad (4.41)$$

with A being the displacement amplitude and k being the wave vector. Inserting Eq. 4.41 into Eq. 4.40 yields

$$-M\omega^2 e^{i(kbn - \omega t)} = -\gamma(2 - e^{-ikb} - e^{ikb}) e^{i(kbn - \omega t)} = -2\gamma [1 - \cos(kb)] e^{i(kbn - \omega t)}. \quad (4.42)$$

Therefore, the frequency ω depends for every atom n on k according to

$$\omega(k) = \sqrt{\frac{2\gamma [1 - \cos(kb)]}{M}}. \quad (4.43)$$

This phonon dispersion is illustrated in Fig. 4.4. It can be seen that at a small k the

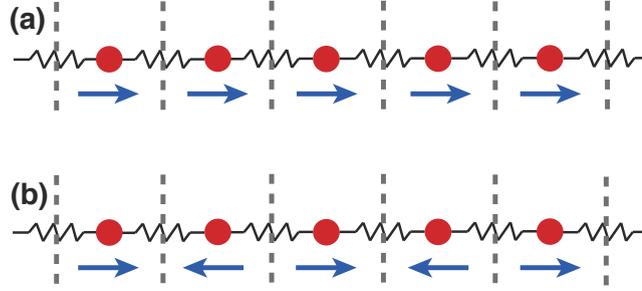


Figure 4.5: Illustration of the phonon mode of a homogeneous infinite atomic chain at the $k = 0$ (a) and at $k = \pm\pi/b$ (b). The dashed lines indicate the length of the unit cell.

frequency ω shows a linear behavior, which can be described by

$$\omega = \sqrt{\frac{\gamma b^2}{M}} |k|. \quad (4.44)$$

When k increases, the frequency reaches its maximum at $\pm\pi/b$. Fig. 4.5 illustrates the atomic motions at $k = 0$ (a) and $k = \pm\pi/b$ (b). In the first case, all atoms move in phase in the same direction. In the second case, we observe that now neighboring atoms move out of phase either towards or away from each other.

4.5.2 Infinite Chain of Atoms with Two Atoms in the Unit Cell

Now we move on to a slightly more complex one-dimensional atomic chain, which has two identical atoms per unit cell [100]. The length of the unit cell is again b and both atoms have a mass M . The positions of the two unit-cell atoms along the chain now amount to nb and $nb + d$ (see Fig. 4.7) and the corresponding displacements will be labeled with ξ_n and ζ_n , respectively. We assume that $d \leq b/2$, leading to interatomic distances between neighboring atoms of d and $b - d$, respectively. Hence, the springs connecting the atoms have now alternating force constants γ and γ' . Again, we only consider interactions between neighboring atoms, leading to the following equations of motion:

$$M \frac{d^2 \xi_n}{dt^2} = -\gamma(\xi_n - \zeta_{n-1}) - \gamma'(\xi_n - \zeta_n), \quad (4.45)$$

$$M \frac{d^2 \zeta_n}{dt^2} = -\gamma(\zeta_n - \xi_{n+1}) - \gamma'(\zeta_n - \xi_n). \quad (4.46)$$

Analogous to Section 4.5.1, we seek a solution of the form

$$\xi_n(t) = A_1 e^{i(kbn - \omega t)}, \quad (4.47)$$

$$\zeta_n(t) = A_2 e^{i(kbn - \omega t)}, \quad (4.48)$$

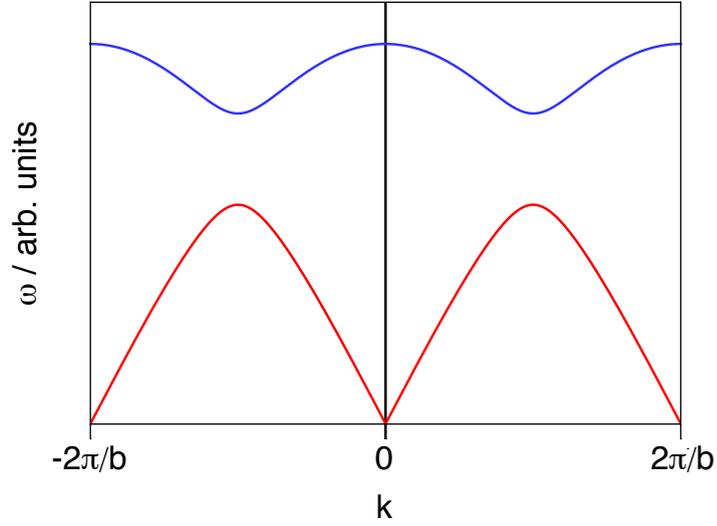


Figure 4.6: Phonon dispersion of a one-dimensional infinite chain of atoms with two atoms in the unit cell. The acoustic and optical mode is shown in red and blue, respectively.

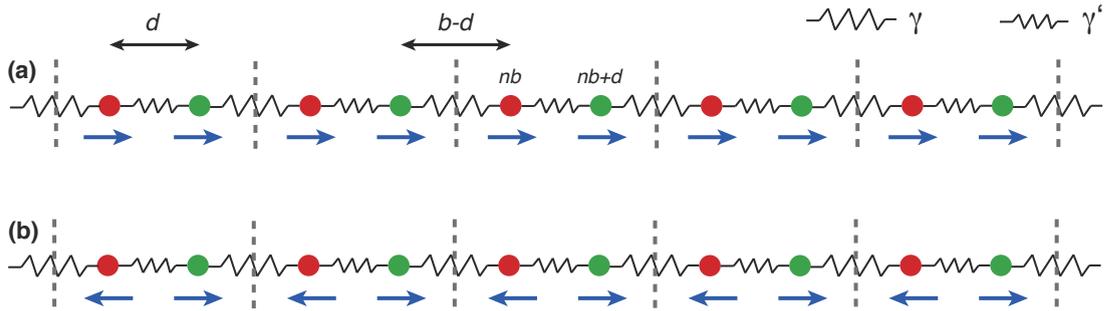


Figure 4.7: Illustration of the acoustic (a) and optical (b) modes in a diatomic infinite chain at $k = 0$.

with A_1 and A_2 being the respective displacement amplitudes. Inserting Eqs. 4.47 and 4.48 into Eqs. 4.45 and 4.46 yields the following coupled equations:

$$[M\omega^2 - (\gamma + \gamma')] A_1 + (\gamma e^{-ikb} + \gamma') A_2 = 0, \quad (4.49)$$

$$(\gamma e^{-ikb} + \gamma') A_1 + [M\omega^2 - (\gamma + \gamma')] A_2 = 0, \quad (4.50)$$

which leads finally to the following expression for the frequency ω :

$$\omega^2 = \frac{\gamma + \gamma'}{M} \pm \frac{1}{M} \sqrt{\gamma^2 + \gamma'^2 + 2\gamma\gamma' \cos(kb)}. \quad (4.51)$$

This leads to two different solutions for every n , one for each atom in our unit cell. The resulting phonon dispersion is plotted in Fig. 4.6, which consists now of two different

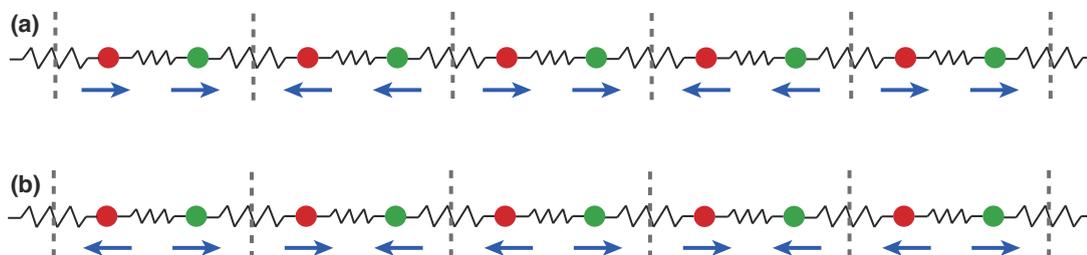


Figure 4.8: Illustration of the acoustic (a) and optical (b) modes in a diatomic infinite chain at $k = \pm\pi/b$.

branches. The lower branch (shown in red) resembles the result obtained in Section 4.5.1; the frequency ω vanishes at $k = 0$ and behaves linear at small values of k . This branch is called the *acoustic branch* since its behavior is similar to that of sound waves. The second branch (shown in blue) is located at a higher frequency and is always non-zero. This branch is called the *optical branch* since such modes are mainly responsible for the optical behavior of ionic crystals. The atomic displacements of both branches are visualized at $k = 0$ and $k = \pm\pi/b$ in Figs. 4.7 and 4.8, respectively. It can be seen that all motions at $k = 0$ are in phase, while all motions at $k = \pm\pi/b$ are out of phase. The acoustic mode behaves like the result of Section 4.5.1. In the optical mode at $k = 0$ the two atoms in the unit cell move against each other, leading to a non-zero frequency. The number of phonon modes per unit cell depends on the dimensionality and the number of atoms. In a three-dimensional molecular crystal with N atom per unit cell we observe $3N$ phonon modes per unit cell, from which always 3 are acoustic modes and the remaining ones are optical modes.

4.6 Vibrations in Molecules

After the discussion of simple periodic systems, we finally consider briefly the vibrations in the case of isolated molecules. As mentioned in Section 4.2, a molecule with N atoms has always $3N$ degrees of freedom. Three of them are needed to describe the center of mass and the corresponding motion consists of simple translations of the whole molecule. Therefore, these three normal modes constitute the *translational degrees of freedom*. For non-linear molecules, another three degrees of freedom are needed for the specification of the orientation about the center of mass, which are called the *rotational degrees of freedom*. This leaves for non-linear molecules $3N - 6$ vibrational degrees of freedom. In case of a linear molecule, we have two rotational degrees of freedom and $3N - 5$ vibrational degrees of freedom. We illustrate these vibrational motions for an ammonia molecule, which has 4 atoms (see Fig. 4.9). In this case we have 6 vibrational degrees of freedom describing one so-called *wagging* motion, one *scissoring* motion, and 4 N-H bond-stretching vibrations.

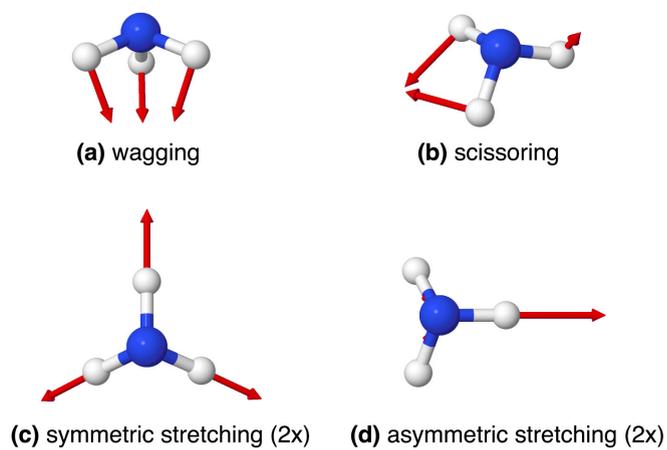


Figure 4.9: Vibrational modes of an ammonia molecule.

In the case of molecular crystals, our resulting normal/phonon modes are a combination of these intramolecular vibrations and intermolecular translations and rotations. Such modes will be visualized in Section 7.3.

Part II

Calculation of Vibrational, Thermal, and Mechanical Properties

Chapter 5

Illustration of Modeling Approaches Using a Deuteroammonia Crystal

For a long time, modeling of molecular crystals was primarily performed by using classical force-field approaches or semi-empirical calculations. Due to the development of sophisticated and efficient quantum-mechanical approximations together with the increase of available computational power, first-principles electronic structure calculations are nowadays possible for practically relevant molecular crystals. DFT has emerged in recent years as the method of choice (especially for larger crystals) due to the development of more accurate DFAs and the inclusion of long-range correlation energy via several models for vdW dispersion interactions. So far, DFT calculations of molecular crystals mainly focused on stabilities and lattice/geometry optimizations without considering any thermal effects.

This chapter provides an overview of current state-of-the-art approaches used for first-principles modeling of molecular crystals and discusses necessary approximations and challenges. We will mainly focus on how temperature effects can be included in the description of structure and stability and discuss the calculation of vibrational, thermal, and elastic properties. Furthermore, we also highlight the influence of dispersion

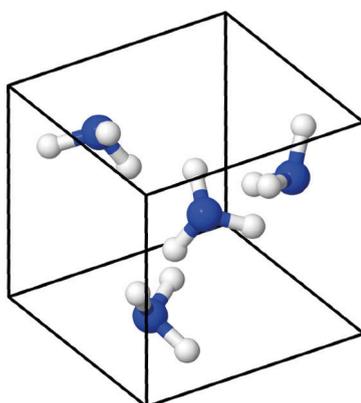


Figure 5.1: Unit cell of a cubic ammonia crystal.

interactions for all discussed properties. Throughout, we will illustrate certain concepts using a cubic deutero-ammonia crystal (see Fig. 5.1). A version of this chapter has been published in *WIREs Comput. Mol. Sci.* [13] (Copyright © 2016 John Wiley & Sons, Ltd.).

5.1 Static Lattice Energies

First, we start with the discussion of static lattice energies E_{latt} , which can be computed for any given molecular crystal structure with

$$E_{\text{latt}} = \frac{E_{\text{cryst}}}{Z} - E_{\text{gas}}, \quad (5.1)$$

where E_{cryst} is the energy of the molecular-crystal unit cell, E_{gas} is the minimum energy of the isolated molecule forming the crystal, and Z refers to the number of molecules within the unit cell. For molecular dimers benchmarks of interaction energies can easily be obtained by comparing with high-level coupled cluster [CCSD(T)] calculations at the basis-set limit using several databases [28, 29]. Unfortunately, such theoretical benchmark calculations cannot be performed for a variety of molecular crystals due to the computational cost of accurate coupled-cluster calculations. However, recent quantum Monte Carlo results [106] indicate the possibility of highly accurate and well balanced molecular crystal benchmark sets in the (near) future.

Currently, the quality of calculated lattice energies is often evaluated by comparison to experimental sublimation enthalpies. One very well studied system is the benzene crystal. The experimental lattice energy amounts to -55.3 ± 2.2 kJ/mol [55]. The highest-level first-principles result available was calculated with a fragment coupled-cluster approach by including two, three, and four-molecule interactions, yielding a lattice energy of $-55.9 \pm 0.76 \pm 0.1$ kJ/mol [55]. The fragment hybrid many-body interaction model (HMBI) based on CCSD(T) provides a lattice energy of -53.0 kJ/mol [107, 108] and the periodic local MP2 approach leads to a lattice energy of -56.6 kJ/mol [109]. This shows that all calculations using post-HF methods provide an excellent agreement with experimental data for the benzene crystal. In contrast, the widely used density functional B3LYP [110–112] without any dispersion correction yields a lattice energy of only -15.9 kJ/mol, which corresponds to an error of almost 40 kJ/mol [113]. Therefore, it is imperative to include a proper description of vdW interactions in DFT calculations. The PBE+MBD and PBE+XDM methods yield lattice energies of -55.0 and -49.5 kJ/mol, respectively [39, 40]. Furthermore, the lattice energy calculated with PBE-D3 with and without three-body terms amounts to -51.0 and -54.8 kJ/mol, respectively [114]. It can be seen that most vdW-inclusive DFT methods provide lattice energies for the benzene crystal that are in excellent agreement with the experimental value.

In 2012, Otero-de-la-Roza and Johnson assembled the so-called C21 benchmark set containing 21 molecular crystals ranging from dispersion-bound to hydrogen-bonded systems [39]. The experimental sublimation enthalpies were always back-corrected for

vibrational contributions, yielding benchmark values for 0 K lattice energies. This benchmark set was later extended and refined by Reilly and Tkatchenko, yielding the so-called X23 benchmark set [40]. For this benchmark set PBE+MBD yields systematically more accurate lattice energies than PBE+TS. The largest difference between TS and MBD is found for the vdW-bound systems. In addition, switching from the GGA functional PBE to the PBE0 hybrid functional leads again to a consistently better description of the lattice energy. The TS method consistently overestimates lattice energies, yielding a mean absolute error (MAE) of 10.0 kJ/mol for PBE0+TS. For PBE0+MBD the corresponding MAE amounts to only 3.8 kJ/mol, placing it within the 4.2 kJ/mol (1 kcal/mol) window of chemical accuracy. The D3 dispersion correction has also been tested for several functionals with the X23 benchmark set [114]. The best performance was achieved for TPSS-D3 without three-body terms, also yielding an MAE of 3.8 kJ/mol.

Nyman *et al.* [56] have recently studied the accuracy of several force fields widely used in CSP calculations for the X23 benchmark set. The resulting deviations in lattice energies are about 2-3 times larger than the best vdW-inclusive DFT methods. These deviations are reasonable in the context of using such force fields in early or intermediate stages of CSP calculations but not always sufficient for correctly predicting the rank ordering of polymorphs. Furthermore, it should be noted that such force-field approaches are often parameterized for a limited number of atom types and environments and hence not very transferable.

All of the above benchmarks are for absolute lattice energies. However, for CSP calculations we are mainly interested in relative stabilities. The C21 and X23 test sets include two polymorphs of oxalic acid. Experimentally [115, 116], the α polymorph is slightly more stable than the β form (by 0.2 kJ/mol). The DFA methods discussed above yield energy differences ranging from about -4 kJ/mol to 4 kJ/mol. For PBE0+TS the β form is more stable by about 1.5 kJ/mol, while PBE0+MBD predicts that the α form is more stable by about 1 kJ/mol, which is consistent with experiment. Moreover, the relative stability of three glycine polymorphs has been studied for several vdW-inclusive DFAs by Marom *et al.* [117]. Only PBE0+MBD was able to capture correctly the qualitative stability ranking, while PBE+TS, PBE0+TS, and PBE+MBD yielded a different qualitative picture. The resulting error in the calculated relative stabilities amounts to about 1 kJ/mol for PBE0+MBD. All these benchmarks suggest that high-level vdW-inclusive DFAs are necessary for accurate relative stabilities and that the stability ranking of polymorphs within an energy window of 1-2 kJ/mol remains a challenge. However, it is important to remember though that the experimental reference values for polymorph and absolute stabilities are also associated with an uncertainty, and that the often-used back correction of experimental sublimation enthalpies introduces an additional uncertainty in these benchmarks.

5.2 Geometries

As with lattice energies, there is currently no high-level first-principles benchmark set available for geometries of various molecular crystals. Therefore, we have to compare

the theoretically optimized structures with experimental crystal structures, usually measured by X-ray diffraction. However, all experimental structures are determined at finite temperatures, whereas theoretical geometry optimizations correspond to 0 K and also do not include zero-point effects. Most molecular crystals expand with increasing temperature and even the zero-point vibrations alone can lead to a volume expansion of about 3%, which will be shown later on for phase-I ammonia. One approach to combat this is to benchmark theoretical structures against the lowest-possible temperature structure available from experiment, minimizing the influence of thermal effects. The lattice vectors obtained with PBE+TS and PBE+MBD optimizations for a subset of the X23 test set show mean relative errors of -0.55% and -0.75% , respectively [40]. Moellmann and Grimme [114] have studied the structures of the C21 set by using PBE-D3 and TPSS-D3, yielding mean relative errors in the cell volume of -1.1% and -2.3% , respectively. Calculated structures have typically a smaller volume compared to experiment, mostly likely due to neglecting thermal expansion [40, 118]. Schatschneider *et al.* [118] have studied the structures of a large set of crystalline polycyclic aromatic hydrocarbons using PBE+TS, yielding on average an error of about $\pm 2\%$ for the lattice vectors. Most of the room temperature experimental densities agree within 5% with the calculated densities, while for lower temperatures the agreement improves to an average deviation of 2.3% [118]. The main reason for the observed overestimation of the densities is again the neglect of any temperature or zero-point effect in the geometry optimization. The most-recent blind test of molecular CSP methods [51] has also shown that several first-principles methods are able to predict structures in very good agreement with experimental geometries.

5.3 Temperature and Pressure

The lattice energies discussed so far contain only total energies (electronic energy and nuclear repulsion); from now on abbreviated with E_{tot} . These energies do not depend on temperature and pressure and are therefore only valid at a temperature of 0 K and a pressure of 0 bar. Strictly speaking, we are not even describing 0 K correctly, as vibrational zero-point energies are missing. While this might sometimes be a reasonable approximation for small isolated molecules in vacuum, temperature and pressure effects can be crucial for the accurate modeling of molecular crystals.

In order to include temperature effects we need to calculate the Helmholtz free energy $F(T, V)$, given by

$$F(T, V) = E_{\text{tot}}(V) + F_{\text{vib}}(T, V), \quad (5.2)$$

where $F_{\text{vib}}(T, V)$ is the vibrational free energy due to the nuclear motion on the Born-Oppenheimer energy surface, T describes the temperature, and V is the unit-cell volume of the molecular crystal. In principle, there are additional contributions to $F(T, V)$, such as the electronic free energy [119] and magnetic contributions. However, these effects can normally be neglected for insulating molecular crystals. Furthermore, we assume that we are always dealing with a perfectly periodic molecular crystal. The existence

of defects or disorder would also introduce additional terms in $F(T, V)$. The effect of pressure can be readily included by calculating the Gibbs free energy of the crystal:

$$G(p, T; V) = F(T; V) + pV, \quad (5.3)$$

where p is an external hydrostatic pressure acting on the unit cell.

In the following sections we will discuss the basics of approximations for including thermal effects with the aid of a simple model system — phase-I deutero-ammonia (ND_3). In addition, we also discuss how vdW interactions influence the results. All calculations for this model system were performed within the all-electron DFT code FHI-aims [74, 120–122], utilizing the *tight* species default settings and phonopy, a code for phonon calculations [123]. More computational details are provided at appropriate places throughout this chapter.

5.4 The Harmonic Approximation

As discussed in Chapter 4, the simplest way to calculate the Helmholtz free energy is the so-called harmonic approximation (HA). The harmonic Helmholtz free energy is then given by

$$F^{\text{HA}}(T) = E_{\text{tot}} + F_{\text{vib}}^{\text{HA}}(T), \quad (5.4)$$

with

$$F_{\text{vib}}^{\text{HA}}(T) = \int d\omega g(\omega) \frac{\hbar\omega}{2} + \int d\omega g(\omega) k_{\text{B}}T \ln \left[1 - \exp\left(-\frac{\hbar\omega}{k_{\text{B}}T}\right) \right], \quad (5.5)$$

where $g(\omega)$ is the phonon density of states (pDOS), i.e., the number of vibrational states at a certain frequency. The first integral in equation 5.5 describes zero-point vibrations, which are present in every quantum system even at a temperature of 0 K, while the second integral describes thermally induced vibrations and includes the vibrational entropy. Note that this term is particularly important for low-frequency vibrations, as the right part of the second integral is largest in magnitude for low frequencies (see Fig. 5.2). If a phonon calculation is performed by using the finite differences approach, there are several technical aspects one must consider. First, the size of the atomic displacements used must be tested carefully, as it is system dependent and also depends on the accuracy of the forces from the used electronic-structure code. The displacements must be large enough to not cause numerical errors in the forces but small enough to still be in the harmonic regime. In our ND_3 case, displacements between 0.001 to 0.01 Å yielded consistent results with a force accuracy of 10^{-4} eV/Å, therefore we have applied displacements of 0.005 Å for all calculations. Secondly, a large enough (super)cell has to be used so that the effect of one atomic displacement does not produce artifacts between periodic images. This was discussed for the X23 test set by Reilly and Tkatchenko [40] with the conclusion that for these systems the used cell should extend at least about 9-10 Å in each direction. However, this should be evaluated carefully for the studied system, especially for salts or molecular crystals involving heavier halogens. All of the calculations presented here were performed by using a $2 \times 2 \times 2$ supercell. For numerical

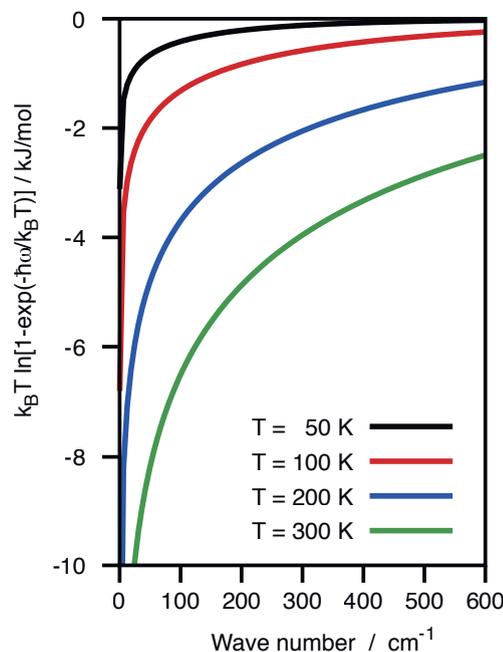


Figure 5.2: The temperature-dependent factor in the second integral of equation 5.5 vs. ω . Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

stability of the calculation it is important to ensure that the reciprocal-space sampling in the supercell is exactly the same as in the optimized unit cell, i.e., if the unit cell was optimized with a $n \times n \times n$ k -point grid, a $2 \times 2 \times 2$ supercell should have a $\frac{n}{2} \times \frac{n}{2} \times \frac{n}{2}$ k -point grid. In our ND_3 example we have used a $4 \times 4 \times 4$ k -grid for the optimization and hence a $2 \times 2 \times 2$ k -grid for the finite differences calculations using the mentioned supercell. Furthermore, the phonon density of states (pDOS) has to be evaluated with a dense q -point grid in reciprocal space (or by calculating large supercells) in order to obtain accurate free energies. We have used the program phonopy [123], which has interfaces to most of the popular periodic electronic-structure codes. Finally, the results have to be carefully analyzed for imaginary frequencies. The appearance of an imaginary frequency at the Γ point indicates that the crystal is not in a local minimum of its PES and this will have a large impact on the obtained free energies and low-frequency phonon modes. Due to the acoustic sum rule, the three acoustic modes have to be zero at the Γ point. However, small deviations are quite common due to the numeric nature of these calculations. Therefore, these three modes often have small imaginary frequencies, typically less than 1 cm^{-1} in magnitude.

5.4.1 Harmonic Helmholtz free energies

Using harmonic Helmholtz free energies, the relative stabilities of polymorphs can be determined for the desired thermodynamic conditions. It was shown by Marom *et al.* [117]

and Rivera *et al.* [124] that zero-point energies can influence the calculated relative stabilities of glycine polymorphs. The aspirin crystal has two polymorphs with degenerate lattice energies, form I and form II, but form I is experimentally much more abundant. Considering only harmonic zero-point energies, the free-energy difference between the two polymorphs remains below 1 kJ/mol [42]. However, when considering $F_{\text{vib}}^{\text{HA}}$ at 300 K, form I becomes more stable by 2.6 kJ/mol when calculated with PBE+MBD. In the case of PBE+TS, form II is more stable by 0.7 kJ/mol. This illustrates the importance of both many-body dispersion interactions and Helmholtz free energies for the prediction of polymorph stabilities.

5.4.2 Harmonic vibrational spectra

The pDOS provides information about all vibrational states within the molecular crystal, i.e., it contains not only modes accessible by infra-red (IR) and Raman spectroscopy (modes at the Γ point) but also out-of-phase phonon modes, which are long-range intermolecular modes. Fig. 5.3(a) shows the obtained pDOS for the ND_3 crystal at the respective optimized geometry of PBE, PBE+TS and PBE+MBD. The four peaks above 800 cm^{-1} correspond to the internal vibrations of the ammonia molecules inside the crystal (see Chapter 4). The experimentally determined internal frequencies [125] for ND_3 are shown as dashed lines in Fig. 5.3(a). It can be seen that these frequencies are nicely captured by all three methods. In this frequency range all the vibrations involve large energy changes, and therefore dispersion interactions play only a minor role for these vibrations. The peaks below 500 cm^{-1} correspond to phonons (or lattice vibrations), i.e., vibrations involving intermolecular motion. Most of these low-frequency modes correspond to intermolecular librations and translations. If the molecules in a crystal have freely-rotating functional groups, like methyl groups, these low-energy intramolecular rotations can also occur in this frequency range. The low-frequency vibrations of ND_3 are shown in detail in Fig. 5.3(b). It can be seen that there are qualitative differences, as well as shifts, in the peak positions of the pDOS between PBE and the vdW-inclusive methods. However, as our test system is a highly symmetric molecular crystal of a small, rigid molecules, the differences are small and stem mostly from differences in the optimized unit cells. For example, the unit-cell volume calculated with PBE is about 15% larger than with PBE+MBD, with the latter being closest to experiment.

In contrast, Fig. 5.4 shows the low-frequency pDOS for a cubic hexamethylbenzene crystal (HMB) calculated for the three discussed methods at the experimental unit-cell volume. HMB has more flexibility than ammonia due to its methyl groups and the crystal is mainly held together by dispersion interactions. In this case, we observe substantial differences in the pDOS between the different methods, even at the same volume. To understand how important these differences are, Fig. 5.3(c) also shows an experimental inelastic neutron scattering (INS) spectrum, measured at 15 K [127]. As an INS spectrum does not represent a generalized pDOS we cannot directly compare the intensities but we can directly compare the position of the peaks (see the gray lines). It can be seen that PBE+MBD reproduces the experimental peak positions quite well, whereas the peaks of PBE and PBE+TS seem to be slightly shifted to higher frequencies.

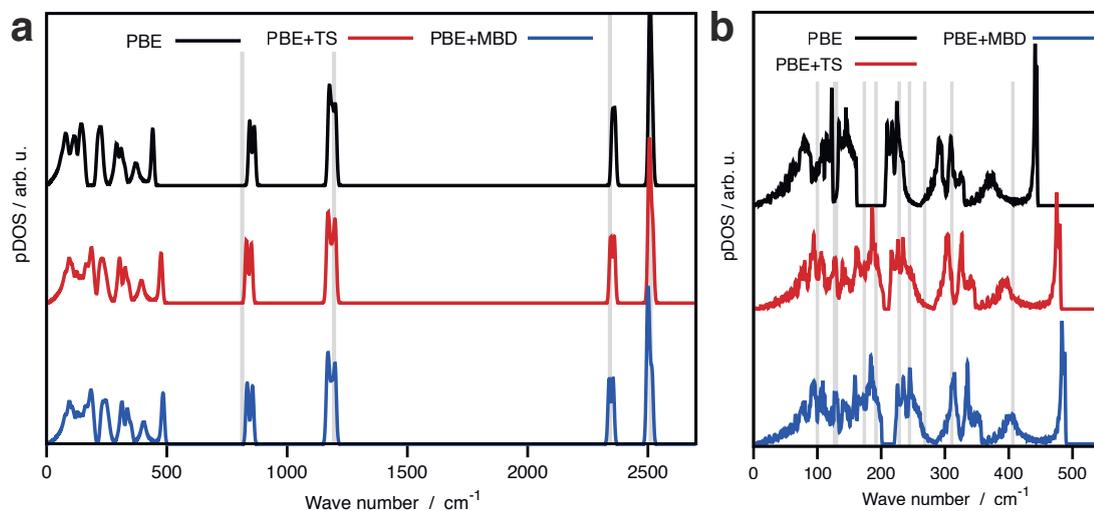


Figure 5.3: Full pDOS of ND₃ (a) and the low-frequency region of the ND₃ pDOS determined with PBE, PBE+TS and PBE+MBD. The gray lines in (a) show the location of the experimental internal modes as determined by Holt *et al.* [125] and in (b) they mark experimentally observed IR (61 K) and Raman (18 K) frequencies [126]. Adapted with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

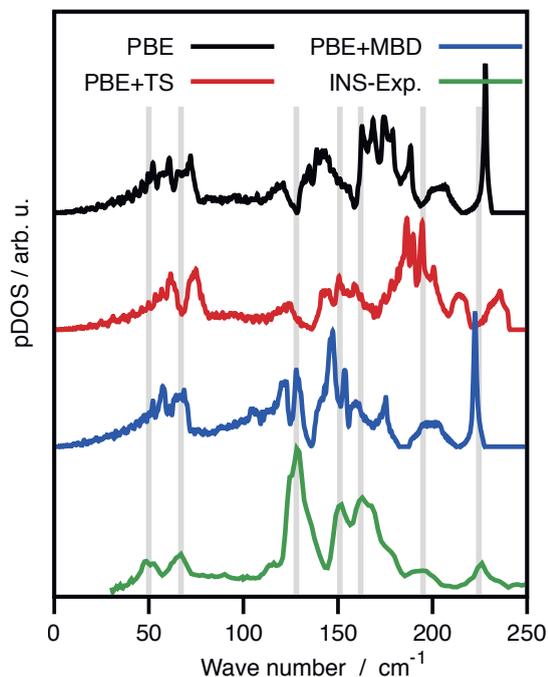


Figure 5.4: Low-frequency pDOS for HMB. The gray lines mark the peak maxima of the experimental INS spectrum measured at a temperature of 15 K by Ciezak *et al.* [127]. Adapted with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

These spectra are compared at the experimental volume to allow for a direct comparison in the absence of structural effects. The results are quite different if calculated at the respective minima of E_{tot} for each method. In this case, the PBE structure has a significantly larger volume and therefore all peaks in the pDOS are shifted to lower wave numbers. Comparing different methods and experiment can be very sensitive to the temperature of the experiment and the unit cell used in calculating the spectra (see Chapter 6).

Although differences between the pDOS calculated with different methods can seem subtle they can be very significant. Motta *et al.* [128] recently calculated the pDOS of a durene crystal and found that including dispersion (using PBE+TS) was essential for obtaining a good agreement with a low-temperature INS spectrum, allowing for the subsequent calculation of charge-transport properties. As noted above, the relative ordering of the two known polymorphs of aspirin changes significantly when calculating Helmholtz free energies with PBE+MBD or PBE+TS [42]. This can be traced to the pDOS calculated with the two methods. For PBE+TS the two forms have comparable spectra but in the MBD case, a peak is found at about 30 cm^{-1} for form I that is completely missing in the TS case and was also not found in form II [42]. It is this difference the pDOS of the two forms that leads to the changes in the calculated relative Helmholtz free energies and to PBE+MBD rationalizing the experimental observation of form I being more abundant.

In recent years, THz time-domain spectroscopy has emerged as very powerful tool to detect drugs and explosives, as well as to distinguish between different polymorphs of molecular crystals [129, 130]. THz spectroscopy detects only vibrational modes at the Γ point and possesses the same selection rules as infrared (IR) spectroscopy. It is typically used to study only the low-frequency phonon modes up to a few THz ($1 \text{ THz} = 33.4 \text{ cm}^{-1}$), which are highly correlated with the crystal-packing arrangement of molecules inside a molecular crystal.

When we compare the intermolecular modes at the Γ point for our ND_3 example with experimental measurements we observe, in general, deviations between 10 and 70 cm^{-1} . PBE always yields lower frequencies than the two vdW-inclusive methods. The reason is probably the overestimation of the unit-cell volume in the case of PBE. Furthermore, we are comparing calculated results without any consideration of thermal effects, while the experimental measurements correspond to temperatures between 18 and 61 K . Already at these temperatures, anharmonic effects due to thermal expansion and atomic motion lead to significant frequency shifts. Reilly *et al.* [131] extensively studied the lattice modes of ammonia and deuterio-ammonia by using the harmonic approximation as well as molecular-dynamics simulations. The observed difference between those methods amounts to $15\text{-}30 \text{ cm}^{-1}$ at 77 K .

The intensities of IR and THz spectra are proportional to the change of the dipole moment p

$$I_{\text{IR}} \propto \left(\frac{dp}{dQ} \right)^2, \quad (5.6)$$

where Q is a normal-mode coordinate. Calculation of IR intensities is much more involved for periodic systems compared to isolated molecules in vacuum because there

is no unique definition of the dipole moment for periodic systems and therefore the calculation requires a Berry-phase approach [132]. However, this option is often not available in electronic-structure codes. Therefore, the intensity of each mode can also be obtained by calculating the difference between dipole moments of the unit cell calculated from atomic charges, e.g., Hirshfeld or Mulliken charges. This unit-cell dipole is then calculated for the ground state and for a geometry that is displaced along the respective normal-mode coordinate. This method is usually referred to as the *difference-dipole method*. Allis *et al.* [133] have applied this approach for a variety of DFAs in order to calculate the THz spectrum of the explosive HMX. This study shows that the THz modes highly depend on the DFA used. In recent years, vdW-inclusive DFT has been used to study the THz spectra of a number of molecular crystals including naproxen [134], naphthalene [135], durene [135], purine [136], α -D-glucose [137], enantiomers of ibuprofen [4], isomers of benzenediols [138], and polymorphs of diclofenac acid [139] and 2,4,6-trinitrotoluene [140]. In general, vdW-inclusive DFT methods provide better THz spectra than traditional DFAs but major differences can sometimes be found between several vdW-inclusive methods. The best agreement between calculated THz spectra in the HA and experimental measurements is found at very low temperatures (< 10 K), where anharmonic effects are likely to be minimized. Nowadays, also low-frequency Raman spectra can be obtained experimentally for molecular crystals, providing additional information about low-frequency vibrations. In this case the intensity is related to the change in polarizability. Recently, it was shown for several molecular crystals that vdW-inclusive DFT can also provide low-frequency Raman spectra in good agreement with experiment [136, 141, 142]. In Chapter 6, we will discuss the applicability of PBE+TS and PBE+MBD for the description of low-frequency spectra of a purine crystal.

5.4.3 What is missing in the harmonic approximation?

The harmonic approximation enables us to calculate free energies and obtain vibrational spectra. But what exactly are we missing in this approximation? As previously discussed, we have omitted all higher-order terms in the Taylor expansion of the PES (see Eq. 4.1). So how valid is this approximation? To answer that question we have taken two representative modes of ND_3 and calculated the PBE+MBD energy for several displacements along the normal-mode coordinates (see Fig. 5.5). For a unitary displacement the norm of the eigenvectors is equal to one. The mode in (a) shows an intermolecular translation and the mode in (b) corresponds to an internal wagging motion. It can be seen for (a) that the HA provides good results for small displacements from the equilibrium geometry but there are significant differences from the *exact* result at larger displacements. We can see from the energy scale in (a) that only a few kJ/mol are required to get into the anharmonic regime of that phonon mode. This suggests that anharmonicity will be a serious issue for low-frequency phonon modes at high temperatures. Displacing along mode (b) requires much more energy compared to (a). In the case of (b) it can be seen that the minimum of the harmonic curve does not correspond

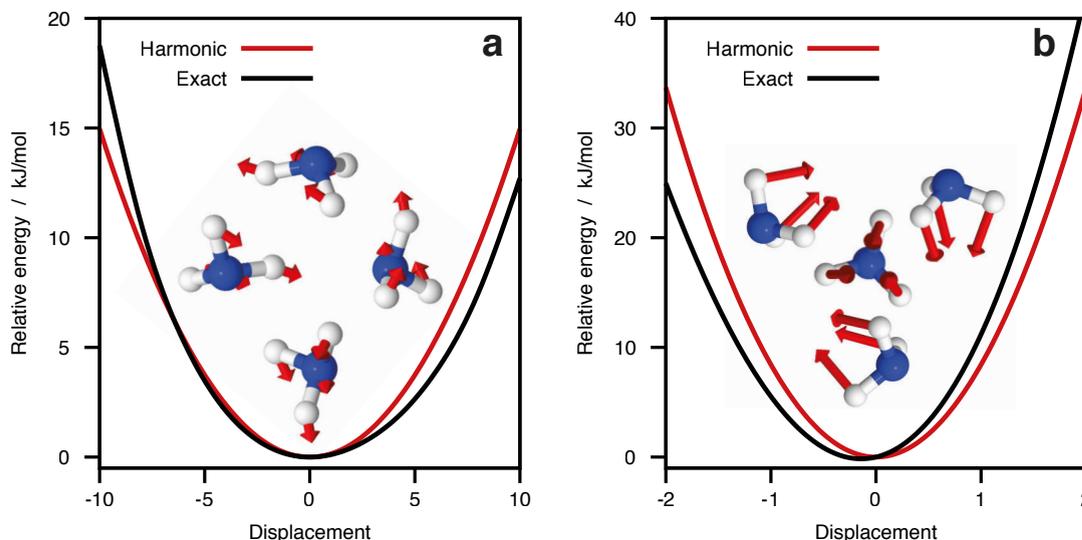


Figure 5.5: Harmonicity of two phonon modes in ND_3 calculated with PBE+MBD. The displacements are shown at a relative scale in which a displacement of 1 means that the norm of the eigenvectors is equal to one. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

to the minimum of the *exact* curve. Considering these anharmonic effects would lead to a shift in the phonon-mode frequency.

Furthermore, the modes are described in the HA as independent oscillators, which means that there are no interactions or coupling between them. In a system at finite temperature the motion of the atoms inside the crystal will certainly not be constrained to movements along normal-mode coordinates. Hence, in order to describe the atomic motion correctly, one would need at least a superposition of several normal modes. Another serious problem is that we have only approximated our PES around the minimum of the total energy E_{tot} . As a result, the pDOS corresponds to exactly this geometry and the free energy at finite temperatures also assumes that the structure of the system does not change at all with temperature. Hence, it is not possible to observe or model phase transitions in the HA. One effect often neglected is the thermal expansion of molecular crystals. The volume of the unit cell will generally increase with increasing temperature. Finally, it is not possible to describe charge transport with the HA since phonon lifetimes and the thermal conductivity are infinite.

5.5 The Quasi-Harmonic Approximation

A straightforward way to improve upon the HA is the so-called quasi-harmonic approximation (QHA) [143]. Therein, the HA is still applied but at several different unit-cell volumes. We start with the HA at the minimum of E_{tot} . Subsequently, the geometry of the unit cell is optimized for several fixed unit-cell volumes V around the equilibrium unit-cell volume and the HA is applied to each of these structures. In the QHA the free

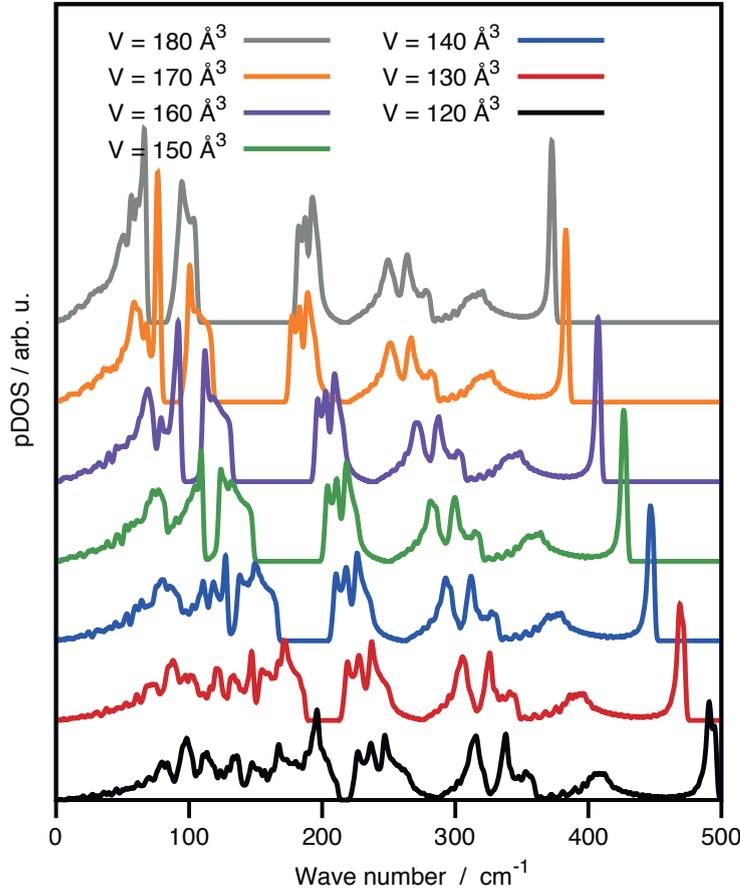


Figure 5.6: Phonon density of states of solid ND_3 calculated with PBE at several unit-cell volumes. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

energy now depends explicitly on the unit-cell volume

$$F^{\text{QHA}}(T, V) = E_{\text{tot}}(V) + F_{\text{vib}}^{\text{HA}}(T, V), \quad (5.7)$$

and therefore the pDOS also depends on the cell volume. Fig. 5.6 shows the dependence of the low-frequency pDOS on the unit-cell volume for the ND_3 crystal as described by PBE. It can be clearly seen that, in general, peaks shift to smaller frequencies with increasing volume and that the pDOS below 200 cm^{-1} changes significantly with increasing volume.

The result of applying the QHA is that we now know for several unit-cell volumes V the harmonic free energy as a function of the temperature T . Therefore, we can now calculate the unit-cell volume corresponding to a specific temperature by fitting our data

Table 5.1: Unit-cell volumes for ND₃ phase I for several methods calculated via optimization (opt), extracted from a quasi-harmonic approximation (n K) and optimization under thermal pressure (p_{th}). All volumes are given in Å³ while p_{th} is in GPa. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

Method	$V(\text{opt})$	$V(2\text{ K})$	$V(77\text{ K})$	$V(180\text{ K})$	$V(p_{\text{th}})$	p_{th}
PBE	142.6	149.9	152.2	163.9	162.6	-0.52
PBE+TS	123.0	126.2	127.2	131.2	133.3	-0.58
PBE+MBD	123.8	127.9	128.6	132.5	131.8	-0.39
Exp. [145]	-	128.6	130.6	134.6	-	-

to an equation of state (EOS). The Murnaghan EOS [144] is often used for this purpose:

$$F^{\text{QHA}}(V) = F_0 + \frac{B_0 V}{B'_0} \left[\frac{(V_0/V)^{B'_0}}{B'_0 - 1} + 1 \right] - \frac{B_0 V_0}{B'_0 - 1}, \quad (5.8)$$

where F_0 is the equilibrium free energy for a certain temperature, V_0 is the corresponding equilibrium volume, B_0 is the bulk modulus at equilibrium volume V_0 , and B'_0 is its pressure derivative.

5.5.1 Thermal expansion

Fig. 5.7 shows the EOS fits for the model ND₃ crystal, while the resulting unit-cell volumes are compared with the experimental data in Table 5.1. At 2 K the unit-cell volume was measured experimentally as 128.6 Å³. It can be clearly seen that a simple lattice optimization does not provide satisfactory results in terms of the unit-cell volume. The PBE functional overestimates V by about 11 % due to missing attractive interactions between the molecules and the two vdW-inclusive methods underestimate V by about 4 %. As discussed before, lattice optimizations do not take into account zero-point vibrations, which results in most cases in slightly too small unit-cell volumes for vdW-inclusive DFT [40]. The QHA includes zero-point vibrations and the DFA+vdW results agree at 2 K very well with experimental results; the error for PBE+TS and PBE+MBD amounts to 1.9 % and 0.5 %, respectively. Note, that PBE alone shows now an error of 17 %, which illustrates how neglecting thermal and zero-point effects can lead to spurious cancellation of some errors. It also reinforces the importance of dispersion interactions for the modeling of even a largely hydrogen-bonded system such as ammonia. Increasing the temperature to 180 K results in thermal expansion of about 5 % [145]. PBE without dispersion now shows an error of about 22 % in predicting the unit-cell volume, while PBE+TS and PBE+MBD provide a very reasonable description with errors being 2.5 % and 1.5 %, respectively. The PBE+MBD results agree consistently better with the experimental observations than the PBE+TS results.

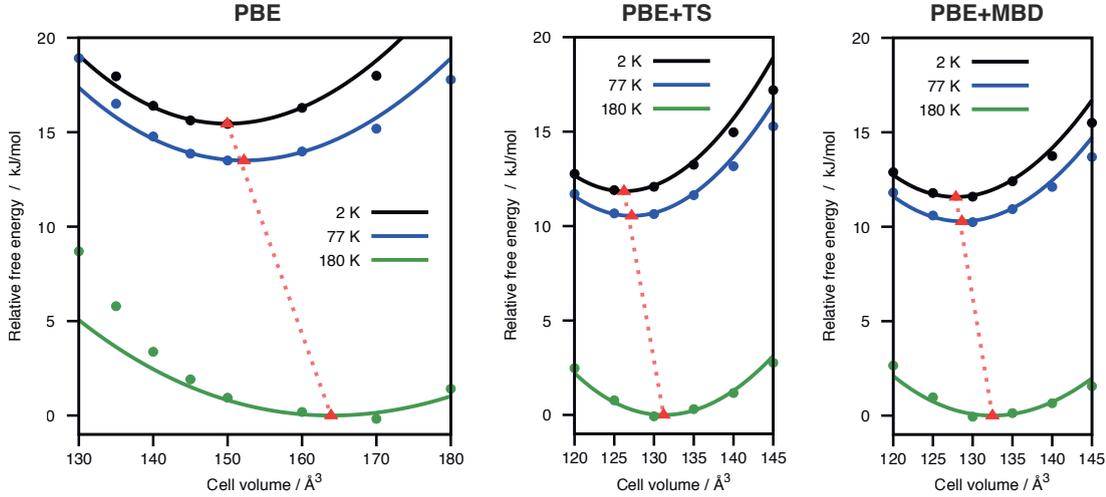


Figure 5.7: Quasi-harmonic approximation for solid ND_3 calculated with different methods at three different temperatures. The solid lines represent Murnaghan EOS fits and the red triangles mark the corresponding minima. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

It is also possible to include pressure effects by simply adding the pV term to F^{QHA} , which then becomes G^{QHA} .

$$G^{\text{QHA}}(p, T, V) = E_{\text{tot}}(V) + F_{\text{vib}}^{\text{HA}}(T, V) + pV \quad (5.9)$$

At ambient conditions pV is very small and is therefore neglected in our QHA calculations. If one wishes only to account for pressure effects without considering thermal effects, an external hydrostatic pressure can be added to the stress tensor. For example, Schatschneider *et al.* studied oligoacenes up to a pressure of 25 GPa using PBE+TS [146]. This hydrostatic external pressure also enables one to approximate the effect of thermal expansion. Describing the effect of $F_{\text{vib}}^{\text{ha}}(T, V)$ utilizing the pV term leads to

$$p_{\text{th}} = \frac{\partial F_{\text{vib}}^{\text{HA}}(T, V)}{\partial V}, \quad (5.10)$$

where p_{th} is a negative thermal pressure, which can be applied to the stress tensor during optimization to mimic thermal effects. In a very crude approximation p_{th} can be obtained by finite differences using the free energies for a certain temperature of the optimized unit cell and a slightly larger or smaller unit cell. In order to demonstrate this we have calculated the thermal pressure for the ND_3 model system for 180 K by considering only the optimized unit cell and the closest available cell from the QHA, which has a larger volume than the optimized cell. This thermal pressure was then applied during a lattice relaxation. The obtained values for p_{th} and the resulting volumes are shown in Table 5.1. It can be seen that the obtained volumes agree with the values from the QHA at 180 K within 2%. Using this thermal-pressure approach, Otero-de-la-Roza and Johnson found that for the C21 data set PBE+D2, PBE+TS, PBE+XDM, and vdW-DF2 yield

mean absolute percentage deviations between 1.3 and 2.8% for cell lengths and between 0.1 and 0.3% for cell angles [39].

Recently, Heit et al. studied the thermal expansion of crystalline carbon dioxide within the QHA using MP2 [147]. Up to a temperature of 195 K, they underestimated the unit-cell volume by only 2–3% compared to experiment. Erba *et al.* have recently reported the directional-dependent thermal expansion of the urea crystal studied within the QHA by using several DFAs and different dispersion corrections [148].

The unit-cell volume corresponding to a specific temperature is determined by an interplay between the 0 K total energies and the vibrational free energies. A measure for the actual expansion of the crystal with temperature is the volumetric thermal expansion coefficient α_V , which can be written as

$$\alpha_V = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p \quad (5.11)$$

The knowledge of α_V enables us to calculate also the heat capacity at constant pressure (C_p), which can be directly compared with experimental calorimetric measurements. C_p is given by

$$C_p = C_V + \alpha_V(T)^2 B(T) V(T) T, \quad (5.12)$$

with B being the bulk modulus and C_V can be calculated according to

$$C_V = k_B \int d\omega g(\omega) \frac{(\hbar\omega/k_B T)^2 \exp(\hbar\omega/k_B T)}{[\exp(\hbar\omega/k_B T) - 1]^2}. \quad (5.13)$$

Fig. 5.8 shows the calculated values for C_p and the linear thermal expansion coefficient $\alpha = \alpha_V/3$ for our ND₃ example compared to experimental values [149]. It can be seen that PBE overestimates C_p at low temperatures and underestimates it at higher temperatures, whereas α is constantly overestimated by about 100%. In contrast, the vdW-inclusive methods underestimate C_p and α by about 20%. The only exception is PBE+TS, which follows the experimental values at low temperature but starts underestimating α at about 100 K. This underestimation of thermal expansion has to be expected, since we are neglecting anharmonic effects due to internal atomic motion.

5.5.2 Elastic properties

Another type of response property that is highly temperature dependent are the elastic constants, which quantify the response of a crystal w.r.t. elastic deformation. The elastic constant matrix \mathbf{C} of a molecular crystal can be obtained by a Taylor expansion around the equilibrium geometry [150]:

$$E(V, \epsilon) = E_0 + V_0 \left(\sum_{i=1}^6 \sigma_i \epsilon_i + \frac{1}{2} \sum_{ij=1}^6 C_{ij} \epsilon_i \epsilon_j \right), \quad (5.14)$$

where ϵ is the strain applied to the unit cell and σ is the corresponding stress of the unit cell. The elastic constants can then be approximated by the second-order derivatives of

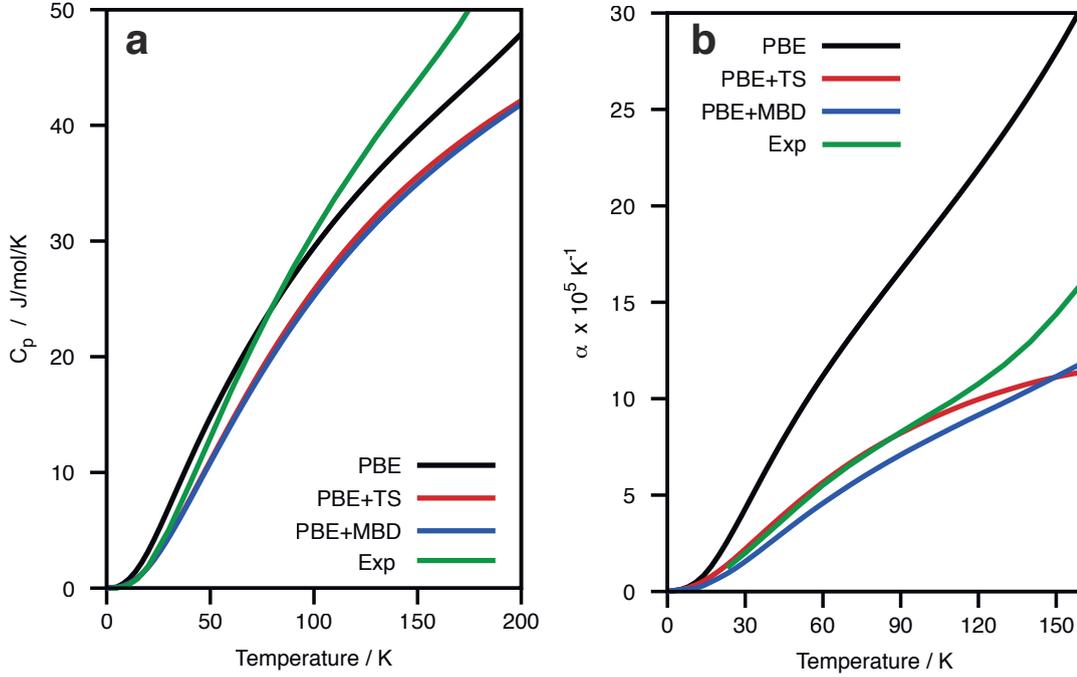


Figure 5.8: Heat capacity at constant pressure (a) and linear thermal expansion coefficient (b) of solid ND_3 obtained from the QHA compared to experimental values from Ref. 149. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

the energy with respect to the applied strain:

$$C_{ij} = \frac{1}{V_0} \frac{\partial^2 E}{\partial \epsilon_i \partial \epsilon_j}. \quad (5.15)$$

Furthermore, the elastic constants can also be obtained via the stress-strain relation

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\epsilon}, \quad (5.16)$$

where the elastic constant matrix is obtained by calculating the stress for several strained unit cells. Often, the elastic constants are calculated based on the equilibrium geometry and correspond therefore to 0 K. Thermal effects can also be included in a simple quasi-harmonic way by obtaining the unit-cell volume for the desired temperature from the QHA and calculating the elastic constants for this structure. We will illustrate this temperature dependence by using our ND_3 model crystal. This example has a cubic unit cell and therefore possesses only three unique elastic constants: C_{11} , C_{12} , and C_{44} . The first two constants are related to volumetric elasticity and C_{44} is related to shear deformation. The bulk modulus B is an inverse measurement for the compressibility of the molecular crystal and can be calculated for a cubic crystal as

$$B = \frac{C_{11} + 2C_{12}}{3}. \quad (5.17)$$

Table 5.2: Elastic constants, bulk modulus, and anisotropy ratio for solid ND₃ calculated at the minima of E_{tot} and at volumes corresponding to 194 K according to respective QHAs, compared to experimental values. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

Method	C_{11} [GPa]	C_{12} [GPa]	C_{44} [GPa]	B [GPa]	A
PBE(opt)	9.8±0.3	3.7±0.1	4.6±0.1	5.7	1.5
PBE+TS(opt)	15.6±0.4	6.8±0.3	8.3±0.1	9.8	1.9
PBE+MBD(opt)	15.6±0.2	6.4±0.1	8.7±0.1	9.5	1.9
PBE(194 K)	3.5±0.2	1.6±0.1	1.3±0.1	2.3	1.4
PBE+TS(194 K)	9.8±0.3	3.7±0.2	5.0±0.1	5.8	1.6
PBE+MBD(194 K)	11.0±0.3	4.6±0.2	5.8±0.1	6.7	1.8
Exp(95 K) [151]	10.0±0.5	5.6±1.0	5.6±0.4	7.1	2.5
Exp(194 K) [152]	9.5±0.8	5.5±0.9	4.9±0.4	6.8	2.4

We have used the above mentioned stress-strain relationship for the calculation of our elastic constants. Strains of $\pm 1\%$ and $\pm 2\%$ were applied to the unit cell and the resulting elastic constants were determined via a least-squares fit. Calculated elastic constants for PBE, PBE+TS and PBE+MBD are given in Table 5.2 and compared with experiment.

First, we compare the results for the optimized geometries with the experimental values obtained at 95 K. It can be seen that PBE underestimates the bulk modulus and all elastic constants while both vdW-inclusive methods overestimate the experimental values. Comparing the results obtained at the estimated unit-cell volumes at 194 K with the experimental results at 194 K, we can see that all of the elastic constants decrease with increasing temperature (and hence volume). The PBE bulk modulus is now almost three times smaller than the experimental value, while the vdW-inclusive methods agree very well with experiment. The deviation in the bulk modulus amounts to 1.0 GPa for PBE+TS and only 0.1 GPa for PBE+MBD.

Another important elastic property of a crystal is the so-called Young's modulus Y_i , which describes the tendency of deformation along a certain direction i . Y_i can be calculated according to

$$Y_i = \frac{\sigma_i}{\epsilon_i}, \quad (5.18)$$

where σ_i describes an uniaxial stress acting in direction i and ϵ_i is the resulting axial strain [153]. Fig. 5.9 shows spherical polar plots of the Young's modulus [154–156] for ND₃. The plot for PBE is very isotropic while PBE+MBD is more anisotropic. If we compare the plots of the optimized geometries with the experimental plot, one might think that PBE is describing the Young's modulus better than PBE+MBD, but when we compare the values for 194 K we can immediately see that vdW interactions have a large influence on elastic properties. We can also measure the anisotropy of the elastic

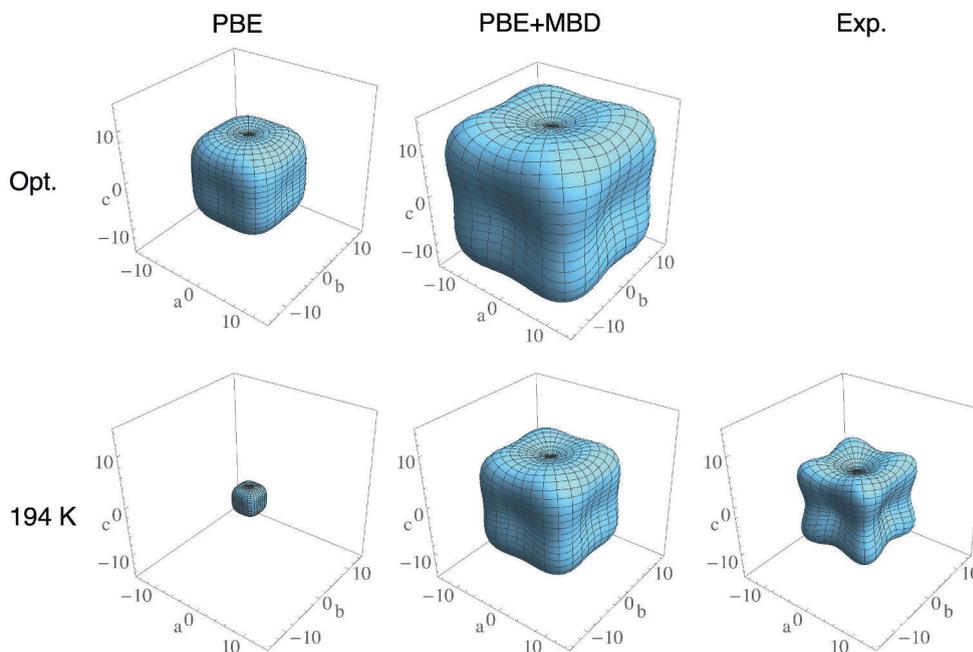


Figure 5.9: Spherical plots of the Young’s modulus of ND_3 (in GPa) obtained at the minima of E_{tot} and at volumes corresponding to 194 K from the QHA for PBE and PBE+MBD, as well as experimental values at 194 K of Ref. 152. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.

constants with the so-called anisotropy factor A , which is for cubic crystals given by

$$A = \frac{2C_{44}}{C_{11} - C_{12}}. \quad (5.19)$$

It can be seen that PBE+MBD shows the largest anisotropy among the theoretical values, but still underestimates the degree of anisotropy compared to experimental measurements, likely due to the lack of anharmonicity in calculating the elastic constants.

The use of DFAs for studying mechanical properties has grown in recent years. The elastic properties of urea have been studied by Erba *et al.* in a quasi-harmonic fashion yielding encouraging results for PBE, PBE0, and B3LYP when paired with the D3 dispersion correction [148]. The mechanical properties of the two aspirin polymorphs have also been studied using PBE+TS and PBE+MBD [42]. Despite their near degenerate lattice energies and a free-energy difference of 2.6 kJ/mol, the two forms show remarkably different elastic properties. As in the case of ND_3 , PBE+MBD also yields more anisotropic elastic response for the two polymorphs of aspirin and smaller elastic constants, agreeing better with experimental values.

5.6 Summary

The modeling and prediction of molecular crystals utilizing first-principles methods has significantly evolved in recent years. With the development of accurate vdW-inclusive DFT methods, in particular, first-principles methods can now be applied to realistic molecular crystals, leading to new insights and understanding of structure, stabilities, polymorphism and response properties such as phonons and elastic moduli. This ability of first-principles methods to model these properties in an accurate, balanced and transferable fashion will make methods such as vdW-inclusive DFT central tools for predicting and engineering molecular crystals in future. Furthermore, high-level first principles methods such as MP2 and beyond can already be applied to several molecular crystals and encouraging progress is being made to enable the applications for realistic systems in the future [54]. Also, the utilization of symmetry-adapted perturbation theory for periodic systems will enable new insights into the nature of intermolecular interactions in molecular crystals [157].

Current vdW-inclusive DFT approaches enable us to calculate lattice energies with a accuracy better than 4.2 kJ/mol. Key to achieving this accuracy is to go beyond a pairwise model of dispersion and include non-additive many-body contributions, for example, by using the MBD method [37, 88]. A key aspect of modeling molecular crystals is in understanding that all of their properties (including their structure) can be highly temperature dependent. Many first-principles studies of molecular crystals treat them in the harmonic limit. While this can give powerful insights, the harmonic approximation neglects anharmonic effects due to the expansion of the unit cell and thermal motion. Omitting these effects when calculating derived properties and quantities such as low-frequency vibrational spectra and elastic constants, can lead to large deviations between experiment and theory, as we have shown using our model system of phase-I deuterio-ammonia. By just considering fully optimized structures, the experimental unit cell-volume at 180 K of our deuterio-ammonia example is overestimated by PBE and underestimated by both vdW-inclusive approaches. An efficient way of accounting for the effect of thermal expansion using first-principles methods is provided by the QHA. Within this approximation, PBE+MBD is able to capture about 80 % of the thermal expansion, leading to an error in the unit cell volume of only 1.5 %. In contrast, pure PBE within the QHA results in an overestimation of the unit cell volume by 22 %, illustrating once again the crucial nature of vdW interactions for molecular-crystal properties. Accounting for thermal expansion is also crucial for the calculation of elastic properties. For instance, within the QHA, PBE+MBD is able to describe the bulk modulus at 194 K with an accuracy of 1.5 % while using the fully-optimized structure leads to an error of almost 40 %.

Chapter 6

Low-Frequency Spectra of a Purine Crystal

Low-frequency vibrational spectra enable the experimental distinction of molecular crystal polymorphs, enantiomers, and isomers [129, 130]. Typically, these low-frequency vibrations are measured via THz time-domain spectroscopy, while nowadays also accurate low-frequency Raman spectra can be obtained for molecular crystals. For a variety of different molecular-crystal systems, THz spectra have been computed using vdW-inclusive DFT [4, 134–139]. Given the non-local nature of the phonons within this frequency range, the results highly depend on the used DFA and dispersion model. These calculations are mostly performed within the harmonic approximation, neglecting any anharmonic effects due to thermal expansion and internal motion. Therefore, the calculations are usually compared to low-temperature experiments, for which anharmonic effects are minimal. However, real-life applications typically demand room-temperature spectra. Regularly, empirical scaling factors [136, 158] are used for the obtained vibrational frequencies in order to match the experimental observations. These scaling factors should effectively iron out the error of the underlying DFA and the neglect of anharmonic effects. While this is a convenient approach for improving the agreement with experimental spectra, such an approach is not really helpful when predicting the spectrum of a so far experimentally unknown molecular crystal. Recently, Ruggiero *et al.* [136] mea-

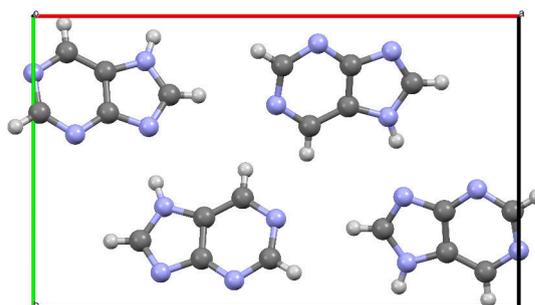


Figure 6.1: The unit cell of the studied orthorhombic purine crystal.

sured room-temperature THz and low-frequency Raman spectra for the purine molecular crystal (see Fig. 6.1). In addition, they also calculated the corresponding spectra utilizing the QHA with the PBE-D3 approach. After applying a scaling factor of 0.85 their calculated frequencies had a mean absolute error (MAE) of only 3 wave numbers w.r.t. the experimental measurements. In this chapter, we will discuss how well we can describe the low-frequency vibrations of purine with our vdW models without the usage of an empirical scaling factor. Therefore, we will also utilize the QHA and in addition estimate further anharmonic effects by utilizing Morse oscillators (see Chapter 4).

6.1 Computational Methods

First, the purine unit cell [159] was fully optimized with PBE+TS [34, 82] using light (l) species default settings for basis functions and integration grids within FHI-aims [74, 120–122, 160–162]. These calculations were performed with a $2 \times 4 \times 9$ k-grid and force components were converged to 10^{-3} eV/Å. In addition, lattice optimizations were also performed with external hydrostatic pressures of 0.4, 0.2, -0.2, -0.4, and -0.6 GPa in order to obtain optimized structures with different unit cell volumes for the QHA. In this case the sampled volume ranges from -2.5% to +4.5% w.r.t. the fully optimized unit-cell volume. Next, the unit-cell volume corresponding to a temperature of 300 K was determined via the QHA as described in Chapter 4. For the fully optimized structure as well as the thermally expanded structure corresponding to room temperature, the phonon modes at the Γ point were calculated within the harmonic approximation using phonopy [123] together with FHI-aims. Finite displacements of 0.005 Å were calculated for a $1 \times 2 \times 3$ supercell utilizing a $2 \times 2 \times 3$ k-grid. The relative infrared (THz) and Raman intensities for each mode were calculated by finite displacements along the normal-mode coordinates Q . The infrared (IR) intensity I_I and the Raman intensity I_R of a normal mode can be estimated with

$$I_I \propto \left(\frac{\partial p}{\partial Q} \right)^2, \quad \text{and} \quad I_R \propto \left(\frac{\partial \alpha}{\partial Q} \right)^2, \quad (6.1)$$

with p being the dipole moment and α being the polarizability. While the dipole is not uniquely defined in periodic systems, the change of the dipole moment using throughout the same unit-cell definition is a well-defined quantity. The values of α are obtained from the TS model (see Chapter 3). For the visualization of the vibrational spectra a Gaussian broadening of 2.5 cm^{-1} was applied.

Furthermore, we want to estimate also anharmonic effects beyond the QHA. For that we utilize Morse oscillators, which are discussed in detail in Chapter 4. Therefore, the thermally-expanded room-temperature structure is displaced along normal-mode coordinates so that the energy change amounts to $\pm 0.5 k_B T$ and $\pm k_B T$ according to the harmonic approximation, with k_B being the Boltzmann constant and the temperature $T = 300$ K. Next, the actual energies of all displaced structures are calculated using PBE+TS using the same settings as described above. Then, a Morse potential [104, 105]

is fitted for every mode. The Morse potential is given by

$$V(x) = D \left(1 - e^{-a(x-x_0)}\right)^2, \quad (6.2)$$

with x being the displacement amplitude, and the parameters D , a , and x_0 describe the well depth, the width of the potential, and the minimum of the potential, respectively. The corresponding frequency ω for the excitation from the ground state to the first vibrationally-excited state can be calculated according to

$$\omega = \omega_0 - 2\omega_0 x_e, \quad (6.3)$$

with

$$\omega_0 = \sqrt{\frac{2Da^2}{\mu}} \quad (6.4)$$

and

$$x_e = \frac{\hbar\omega_0}{4D}, \quad (6.5)$$

where μ describes the reduced mass. Replacing the harmonic oscillators by Morse oscillators leads to a frequency shift but the obtained intensities remain unchanged.

Furthermore, we also want to study what effect MBD interactions, exact exchange, and a larger basis set have on the obtained vibrational spectra. Therefore, all further on described calculations were performed with PBE+MBD and PBE0+MBD utilizing light (l) species default settings, and PBE+MBD using tight (t) species default settings. Other computational settings are identically to the above described calculations. However, we will rely for these calculations on the room-temperature unit-cell volume obtained with PBE+TS within the QHA. Therefore, lattice optimizations were only performed without external pressure and with a hydrostatic pressure of -0.4 GPa. The lattice constants corresponding to the above determined volume were obtained via linear regression from the two respective optimizations. Next, the geometries for the so obtained unit cells were optimized and the vibrational spectra utilizing Morse oscillators were calculated as described above. For MBD calculations, the polarizability used for the calculation of Raman intensities is obtained directly from the MBD approach.

6.2 Calculated Low-Frequency Spectra

First, we discuss the three different approximations considered using the PBE+TS method. The obtained spectra are shown in Fig. 6.2 and the relevant phonon frequencies are listed in Tab. 6.1. Within this table we compare the obtained IR and Raman-active phonon frequencies with experimental peak maxima. In almost all cases the corresponding experimental frequency can be easily identified. However, the second and third IR-active modes are in close proximity in our calculations. In the experiment, the third IR-active mode has a larger intensity than the second one. Therefore, we compare the calculated mode with larger intensity to the third experimental IR frequency.

Table 6.1: Obtained IR and Raman active PBE+TS phonon frequencies for the fully optimized structure within the harmonic approximation (Opt./HA), for the thermally expanded unit cell within the harmonic approximation (QHA/HA), and for the thermally expanded unit cell using Morse oscillators (QHA/Morse) compared with experimental measurements [136]. Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given.

	Exp.	Opt./HA	QHA/HA	QHA/Morse
IR	41.4	46.6	41.6	43.1
Raman	46.4	46.6	41.6	43.1
IR	47.1	59.6	52.3	49.7
IR	52.0	56.7	49.3	51.3
IR	59.3	79.9	70.3	67.0
Raman	66.2	79.9	80.0	74.8
IR	73.4	85.6	75.4	75.5
Raman	99.5	107.1	102.8	97.4
Raman	110.5	135.4	126.7	122.1
MAE	-	12.7	7.5	5.2
MAX	-	24.9	16.2	11.6

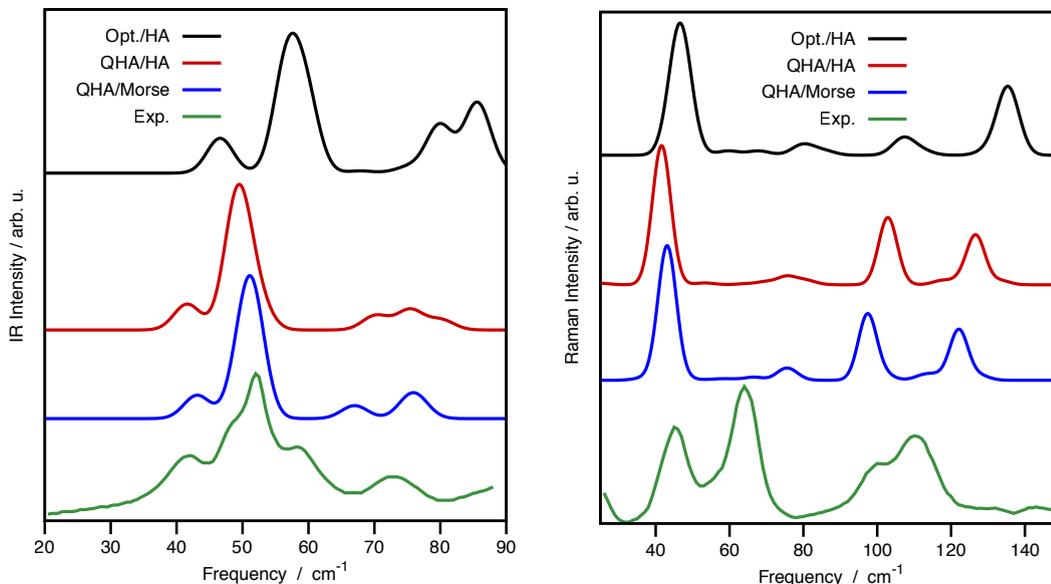


Figure 6.2: Obtained PBE+TS low-frequency THz/IR (left) and Raman (right) spectra for the fully optimized structure within the harmonic approximation (Opt./HA), for the thermally expanded unit cell within the harmonic approximation (QHA/HA), and for the thermally expanded unit cell using Morse oscillators (QHA/Morse) compared with experimental measurements (Exp.) [136]. The individual spectra are offset for clarity.

Let us start with the commonly used harmonic approximation on top of the fully relaxed structure. In this case we observe significant deviations from the experimental observations. The mean absolute error (MAE) w.r.t the experiment amounts to 12.7 cm^{-1} with the maximum absolute deviation (MAX) being 24.9 cm^{-1} . It can be seen that in this case all frequencies are significantly overestimated. Since we did not include any thermal expansion in this approach, we underestimate the room-temperature unit cell volume by about 3%. The largest error in the individual lattice constants amounts to about 2%. Next, we move on to the thermally expanded structure according to the QHA. In this case our obtained unit cell volume agrees very well with the experimentally measured one; the error amounts to only 0.3%. It can be seen that all peaks shift now to lower frequencies due to thermal expansion of the crystal. Now, the simulated spectrum is already in better agreement with experiment, with a MAE of 7.5 cm^{-1} and a MAX of 16.2 cm^{-1} . Replacing now the harmonic oscillators with Morse oscillators leads to a further improvement of the peak positions, with a MAE of 5.2 cm^{-1} and a MAX of 11.6 cm^{-1} . This result shows that it is possible to obtain accurate phonon frequencies for this system, even without the usage of empirical scaling factors.

The calculated IR intensities are in a relatively good agreement with the experimental spectrum but for the Raman intensities there are significant differences. For example the intensity of the second Raman-active mode is significantly underestimated. However, our main focus in this chapter lies on the accuracy of the phonon frequencies. Our Raman intensities here are approximated by finite differences for the polarizability as described by the TS model. Furthermore, the Morse oscillator leads to a shift in the frequency, while the intensities are the same as for the harmonic approximation. Therefore, we cannot expect to obtain completely accurate intensities given our used approximations. However, the obtained intensities correctly identify IR and Raman-active modes. Note that a more accurate description of the intensities can be obtained by using for example the Coupled-Perturbed Kohn-Sham method [163–165] as shown in Ref. 136.

Furthermore, we also analyze what impact MBD interactions, exact exchange, and a larger basis set have on the resulting spectra. All these calculations were performed at a unit cell volume of 532.5 \AA^3 , which is the room-temperature volume obtained with PBE+TS via the QHA. The obtained harmonic frequencies are listed in Table 6.2, while the Morse frequencies are given in Table 6.3. The vibrational spectra obtained via Morse oscillators are shown in Fig. 6.3. It can be seen that all shown methods yield relatively similar results. At the harmonic level, the MAE varies from 6.8 cm^{-1} (PBE+MBD/t) to 7.5 cm^{-1} (PBE+TS/l) and the maximal observed deviation varies from 15.8 to 18.4 cm^{-1} . At the Morse level the MAE varies from 4.8 cm^{-1} (PBE+MBD/t) to 5.2 cm^{-1} (PBE+TS/l) and the maximal observed deviation varies from 8.8 to 11.6 cm^{-1} . The applied Morse model improves consistently for all methods the MAE as well as the MAX. The improvement over the harmonic treatment amounts to about 2 cm^{-1} for the MAE and the maximum absolute deviation improves significantly by 30-45 %.

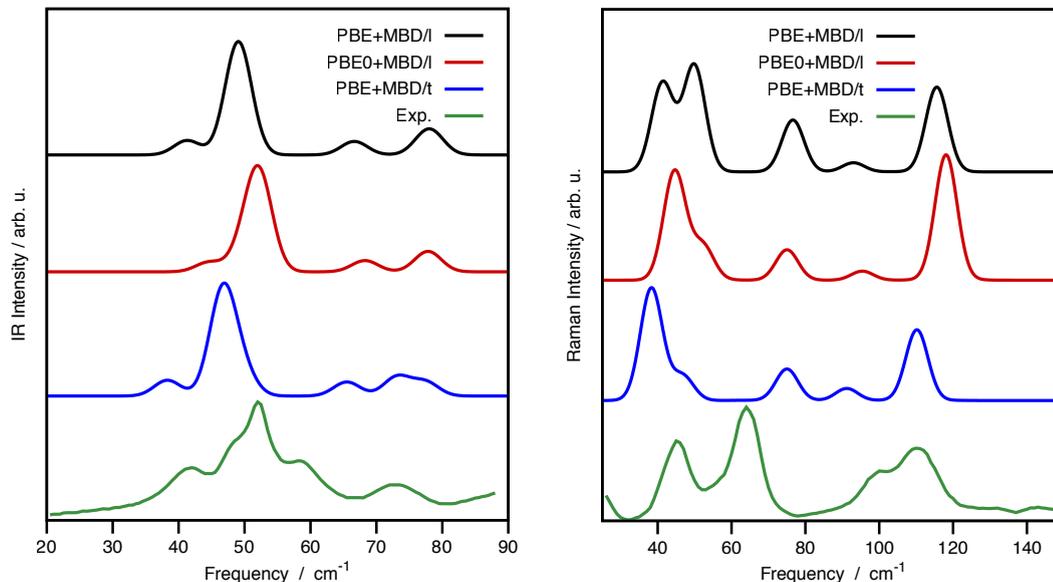


Figure 6.3: Obtained low-frequency THz/IR (left) and Raman (right) spectra for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t compared with experimental measurements (Exp.) [136]. All calculations were performed at a unit cell volume of 532.5 \AA^3 and Morse oscillators were used throughout. The individual spectra are offset for clarity.

Table 6.2: Obtained IR and Raman active PBE+TS phonon frequencies for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t within the harmonic approximation compared with experimental measurements [136]. All calculations were performed at a thermally expanded unit cell volume of 532.5 \AA^3 . Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given.

	Exp.	PBE+TS/l	PBE+MBD/l	PBE0+MBD/l	PBE+MBD/t
IR	41.4	41.6	41.6	43.4	42.2
Raman	46.4	41.6	41.6/49.5	43.4/50.9	42.2/50.2
IR	47.1	52.3	51.7	51.4	53.1
IR	52.0	49.3	49.5	50.9	50.2
IR	59.3	70.3	69.6	70.1	70.5
Raman	66.2	80.0	84.6	82.0	82.2
IR	73.4	75.4	78.4	77.9	78.2
Raman	99.5	102.8	97.0	97.4	97.5
Raman	110.5	126.7	121.6	122.6	120.1
MAE	-	7.5	7.4	6.9	6.8
MAX	-	16.2	18.4	15.8	16.0

Table 6.3: Obtained IR and Raman active PBE+TS phonon frequencies for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t using Morse oscillators compared with experimental measurements [136]. All calculations were performed at a thermally expanded unit cell volume of 532.5 \AA^3 . Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given..

	Exp.	PBE+TS/l	PBE+MBD/l	PBE0+MBD/l	PBE+MBD/t
IR	41.4	43.1	41.3	44.6	38.3
Raman	46.4	43.1	41.3/49.0	44.6/52.2	38.3/46.7
IR	47.1	49.7	49.4	49.8	49.7
IR	52.0	51.3	49.0	52.2	46.7
IR	59.3	67.0	66.6	68.2	65.5
Raman	66.2	74.8	76.6	75.0	75.0
IR	73.4	75.5	77.4	77.9	73.1
Raman	99.5	97.4	93.0	95.4	91.2
Raman	110.5	122.1	115.6	118.1	110.2
MAE	-	5.2	5.0	5.1	4.8
MAX	-	11.6	10.4	8.9	8.8

6.3 Summary

This chapter has shown that accurate phonon frequencies can be obtained with vdW-inclusive DFT even without the utilization of an empirical scaling factor. The harmonic approximation on top of fully-optimized structures is insufficient for the calculation of low-frequency vibrational spectra at room temperature. The QHA already significantly improves the description. Furthermore, our introduced Morse oscillator model is able to further improve the frequencies for all shown methods, leading to a MAE of about 5 cm^{-1} . This illustrates the usefulness of this model and also indicates that it should improve the description of vibrational free energies. All the studied methods yield rather similar MAEs when the Morse model is applied. Therefore, PBE+TS or PBE+MBD calculations on the light level will probably be sufficient for similar crystals. However, for flexible or more polar molecular crystals the usage of a larger basis set and/or a hybrid functional might be required.

Chapter 7

Time-Resolved THz Spectroscopy of an Aspirin Crystal

In the previous chapter we have discussed the calculation of low-frequency vibrational spectra for one molecular crystal. Nowadays, accurate THz spectra can be measured also in a two-dimensional (2D) time-dependent fashion, allowing for the observation of the THz response of a certain sample. Recently, such a 2D spectrum was measured for an aspirin crystal and we utilized our PBE+MBD framework to provide a qualitative interpretation of the experimental observations. This study has been published in *Physical Review Letters* [166]. The author of this thesis contributed to that paper all first-principles calculations using PBE+MBD, which are described in this chapter. The experimental THz spectra were measured by Giulia Folpini from the Max-Born-Institut für Nichtlineare Optik und Kurzzeitspektroskopie in Berlin.

7.1 Background

Two-dimensional (2D) spectroscopy in the THz regime ($\approx 1\text{-}25$ THz) allows for the study of ultrafast dynamics and to uncover couplings between different excitations [167, 168]. Herein, our focus lies on the low-end of the THz regime, where we observe mostly phonon excitations. 2D spectra are typically measured in the so-called *pump-probe* approach. First, a laser pulse is used to excite the sample, i.e., to create a non-equilibrium state. Then, a second pulse is applied after a specific time interval to measure the THz spectrum at that moment in time. Performing these measurements for a variety of different relaxation times provides information about the response of the system to the applied THz excitations.

Here, we discuss a 2D spectrum of aspirin (acetylsalicylic acid) form I (see Fig. 7.1). It is noted in passing that aspirin was long believed to exist only in this form. However, a second less common form was predicted [169] and later experimentally confirmed [170]. Reilly and Tkatchenko have studied low-frequency vibrations in both forms of aspirin applying the MBD method [42]. They found a coupling between low-frequency phonon modes and collective electronic fluctuations, which lead for form I to a shift of phonon

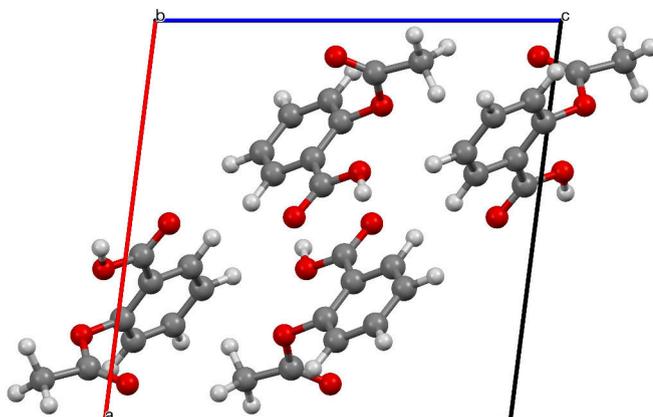


Figure 7.1: Unit cell of aspirin form I.

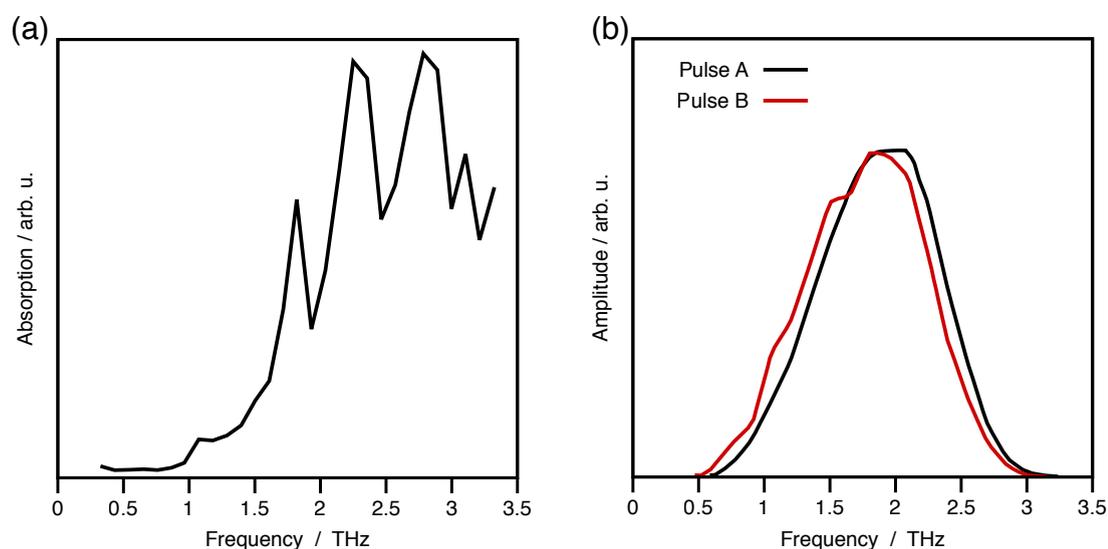


Figure 7.2: Measured linear absorption of the aspirin sample at $T = 80$ K (a) and the amplitude spectra of applied pulses A and B (b). Data from Ref. 166.

frequencies to smaller frequencies. The experimentally measured linear THz spectrum of aspirin at 80 K is shown in Fig. 7.2(a) [166]. The main peaks are located at 1.8, 2.3, and 2.8 THz, and a small feature is detected at 1.1 THz. These results are in good agreement with previous measurements [171]. The 2D THz spectra revealed that the peak originally located at 1.1 THz experiences a blueshift to 1.7 THz upon nonlinear absorption [166]. The frequency distribution of the two used pulses is shown in Fig. 7.2(b). In this chapter, we try to rationalize this peak shift based on PBE+MBD first-principles calculations.

7.2 Computational Methods

The PBE+MBD [35, 64, 82] method was used throughout utilizing the all-electron code FHI-aims [74, 120–122]. The unit cell of aspirin form I was optimized with the *light* species default settings for basis functions and integration grids using a $3 \times 6 \times 3$ k-grid. Force components were converged to 10^{-3} eV/Å. All phonon calculations were carried out within the harmonic approximation utilizing phonopy [123]. The finite displacement calculations were performed with displacements of 0.005 Å using a $1 \times 2 \times 1$ supercell, ensuring a length of at least 11 Å in each direction in order to avoid artifacts. Furthermore, a $3 \times 3 \times 3$ k-grid was used for the supercell calculations employing light species default settings and a force accuracy of 10^{-5} eV/Å.

In order to model the THz response of the crystal, we calculated the THz spectra for the optimized structure and several appropriately modulated structures. In the latter case, the optimized structure was displaced along a linear combination of normal modes, which corresponds to the amplitude spectra of the used THz pulses shown in Fig. 7.2. We used the average of both spectra and the magnitude of the modulation is always described in terms of the maximum displacement within this linear combination. For individual normal modes at the Γ point the displacement δ_i of atom i was calculated according to

$$\delta_i = \frac{A}{\sqrt{N m_i}} \mathbf{e}_i, \quad (7.1)$$

where A describes the displacement amplitude, N is the number of atoms in the cell, and m_i refers to the mass of atom i in atomic mass units (amu), as implemented in phonopy. The vector \mathbf{e}_i describes the part of the normalized eigenvector, which belongs to atom i . The magnitude of the used displacement amplitude A for each mode is determined by the spectra shown in Fig. 7.2. We studied five different modulated structures, for which the maximum displacement amplitudes (A_{\max}) varied between 0.1 and $4 \text{ \AA} \sqrt{\text{amu}}$. The intensities of the THz (far-IR spectra) were calculated as described in Chapter 6 by finite differences using a displacement amplitude $A = 1 \text{ \AA} \sqrt{\text{amu}}$ and a Gaussian broadening of 0.03 THz was applied.

7.3 THz Response of Aspirin Form I

The so obtained THz spectra for the optimized structure and all considered modulated structures are shown in Fig. 7.3. It can be seen that our calculated THz spectrum for the optimized geometry agrees quite well with the experimentally obtained linear absorption spectrum. All peak positions agree within around 0.2 THz. Such small deviations from the experimentally observed spectrum are expected since we did not consider any anharmonic effects for the phonon frequencies and neglected the thermal expansion of the unit cell. Our obtained unit cell volume differs from the experimentally measured unit cell at 100 K [172] by less than 2 %. We note here that no empirical scaling factor was applied to the frequencies and that our obtained peaks located at 1.1 and 1.3 THz will both contribute to the measured peak at 1.1 THz (see Fig. 7.2(a)). We have

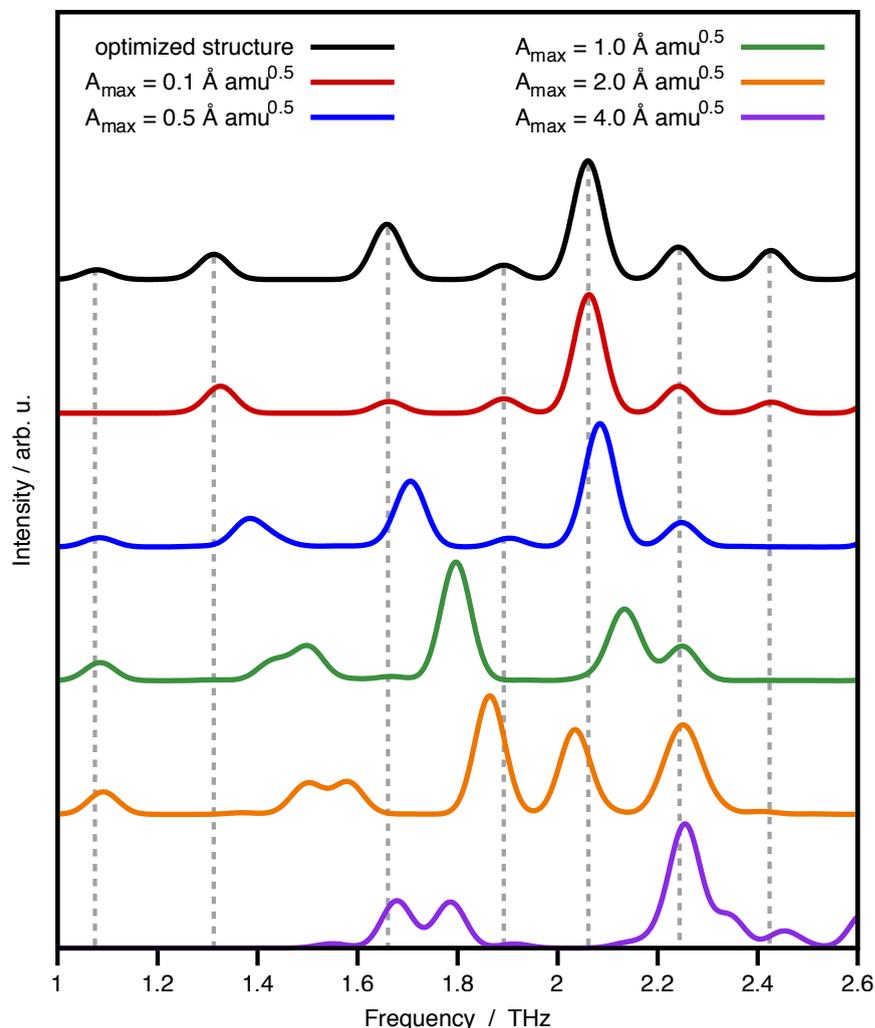


Figure 7.3: THz spectra for the optimized structure of aspirin form I and for all studied modulated structures. The shown spectra were normalized to the largest observed intensity within the plotted spectral range and offset for clarity. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.

visualized IR-active low-frequency modes of the optimized structure in Fig. 7.4. It can be seen that these modes are a mixture between the expected intermolecular translations and rotation at this frequency range, and concerted motions involving methyl groups. This was attributed to plasmon-phonon coupling [42]. The mode located at 1.3 THz involves particularly pronounced methyl group rotations.

Fig. 7.3 shows that peaks generally shift to larger frequencies (blue shift) with increasing modulation. For the peak of the optimized structure initially located at 1.3 THz, we observe a constant blue shift with increasing modulation, at a modulation with $A_{\max} = 4 \text{ \AA} \sqrt{\text{amu}}$ it is located around 1.7 THz. Furthermore, this mode experiences the largest blue shift among the studied modes. These findings are in qualitative agreement

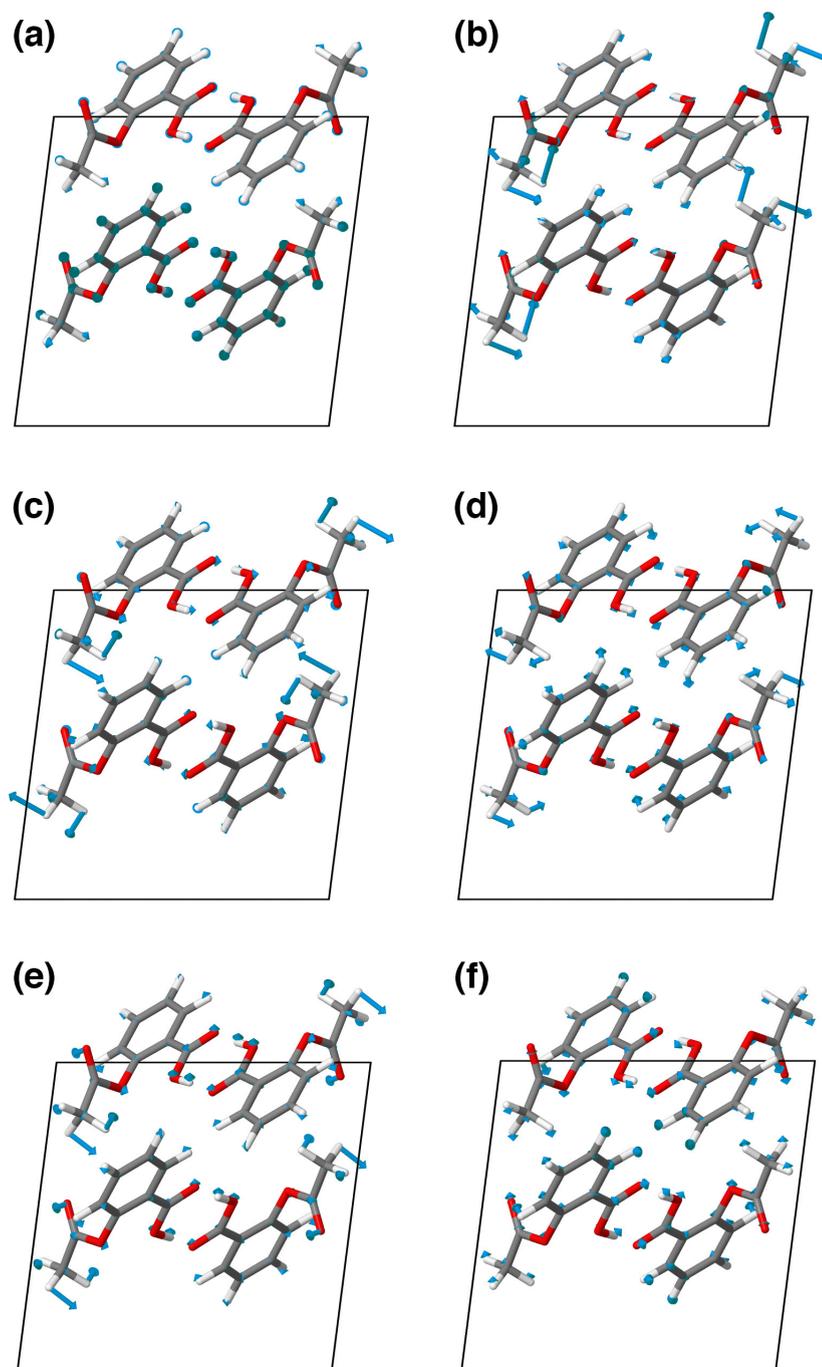


Figure 7.4: Lowest-frequency IR-active phonon modes of the optimized aspirin structure located at 1.08 THz (a), 1.31 THz (b), 1.66 THz (c), 1.89 THz (d), 2.06 THz (e), and 2.24 THz (f). The mass-weighted displacement vectors (blue) represent the direction and the relative magnitude of the atomic motion. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.

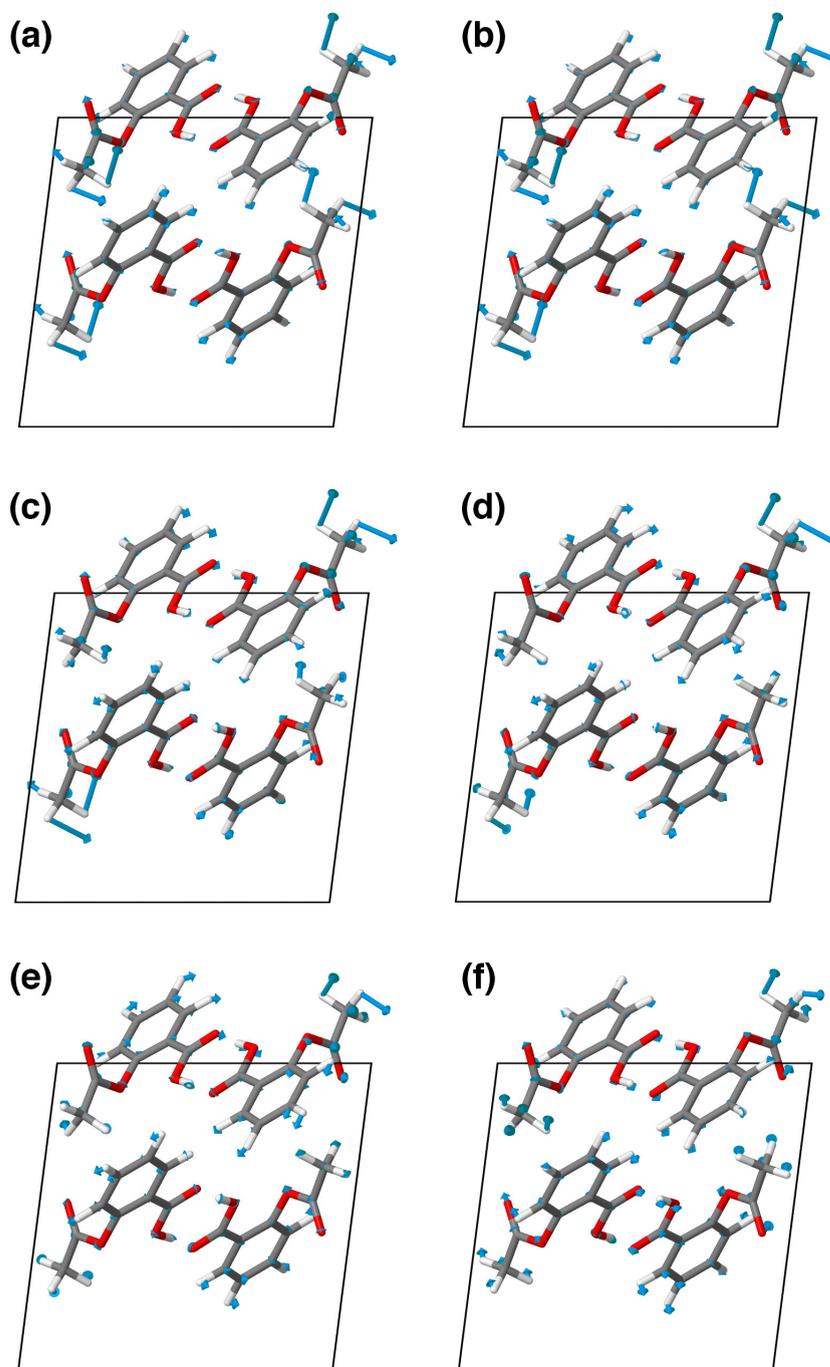


Figure 7.5: Representative phonon mode within the blue-shifted peak initially located at 1.3 THz for the optimized structure and all modulated structures: a) optimized structure (1.31 THz); b) $A_{\max} = 0.1 \text{ \AA}\sqrt{\text{amu}}$ (1.33 THz); c) $A_{\max} = 0.5 \text{ \AA}\sqrt{\text{amu}}$ (1.38 THz); d) $A_{\max} = 1.0 \text{ \AA}\sqrt{\text{amu}}$ (1.43 THz); e) $A_{\max} = 2.0 \text{ \AA}\sqrt{\text{amu}}$ (1.49 THz); f) $A_{\max} = 4.0 \text{ \AA}\sqrt{\text{amu}}$ (1.67 THz). The mass-weighted displacement vectors (blue) represent the direction and the relative magnitude of the atomic motion. For systems in which more than one mode contribute to this peak, one representative mode is shown. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.

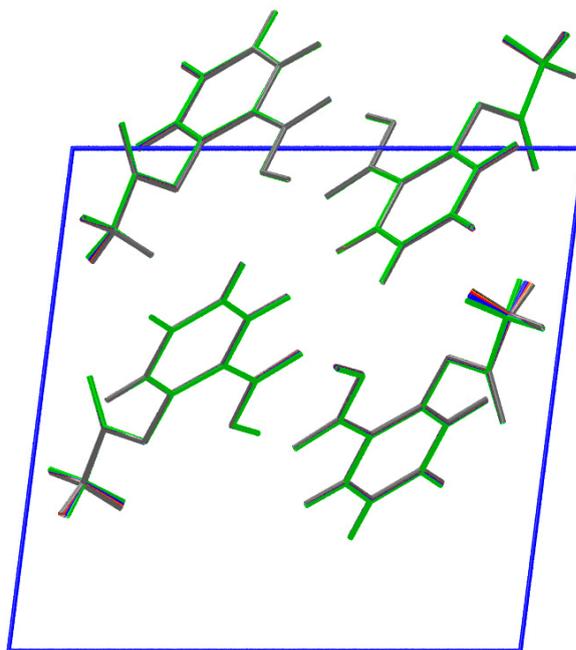


Figure 7.6: Overlay of the unit cell of the optimized structure (gray) with the distorted structures of a maximum displacement amplitude of 0.1 (black), 0.5 (orange), 1 (red), 2 (blue), and 4 $\text{\AA}\sqrt{\text{amu}}$ (green). Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.

with the experimental observations. We have visualized in Fig. 7.5 how the phonon mode of the discussed peak changes with modulation. It can be clearly seen that the observed hindered methyl-group rotations change their symmetry and decrease in magnitude with increasing modulation of the underlying structure. Finally, we analyze the change in the structure that leads to the observed modification of phonon modes and hence to the frequency shifts. Therefore, Fig. 7.6 shows an overlay of all modulated structures with the optimized unit cell. It can be seen that the main geometrical change upon modulation is the modification of methyl group torsion angles. For the modulations considered here the differences to the optimized structure vary from 0.2° and 11° .

7.4 Summary

In summary, we have shown that our applied PBE+MBD calculations are able to qualitatively describe the experimentally-observed THz response of the studied low-frequency mode. In case of the optimized structure this mode contains a combination of intermolecular translations/rotation with concerted methyl-group motions. The latter motions decrease with increasing modulation of the structure, leading to the observed frequency shift to higher wave numbers. While this obtained result is in qualitative agreement with the experimental observations, more sophisticated time-dependent calculations would be

necessary for studying the THz response in an accurate and reliable way. One could envision a model system based on quantum Drude oscillators [173–175] which couples the electronic and vibrational degrees of freedom in a similar fashion as Venkataram *et al.* [176].

Part III

Crystal Structure Prediction

Chapter 8

Sixth Blind Test of Organic Crystal Structure Prediction Methods

After having discussed the modeling of properties for already known crystal structures, we turn now to the discussion of organic crystal structure prediction. One important measure of the quality of an organic crystal structure prediction (CSP) procedure are the regular blind tests organized by the CCDC. The most recent sixth blind test included five quite diverse target systems and the results have been published in *Acta Cryst. B* [51]. This chapter discusses our contribution to this blind test (Submission 25: Johannes Hoja, Hsin-Yu Ko, Roberto Car, Robert A. DiStasio Jr., and Alexandre Tkatchenko). A part of the described calculations was performed by Hsin-Yu Ko from Princeton University.

8.1 Background

This sixth CCDC blind test features five target systems, which are shown in Fig. 8.1. This quite diverse set of systems includes a small rigid molecule, a flexible highly polymorphic molecule, a salt, a co-crystal, and a fairly large molecule. Within this blind test, participants were only provided with the chemical diagram of the involved molecules. Each participating group could submit for every system two separate lists of up to 100 structures, ranked according to their relative stabilities with rank 1 being the most stable structure. After the submission deadline, these structures were compared to the experimentally obtained crystal structures using geometrical criteria.

Our submission focuses on the accurate calculation of relative stabilities, which is the final part of a full CSP. Therefore, we rely for the crystallographic space sampling on an initial structure set provided by the group of Sarah L. Price at the University College London. On top of these structures we apply a hierarchical procedure which stepwise reduces the number of considered structures and improves the description of structures and stabilities utilizing vdW-inclusive DFT.

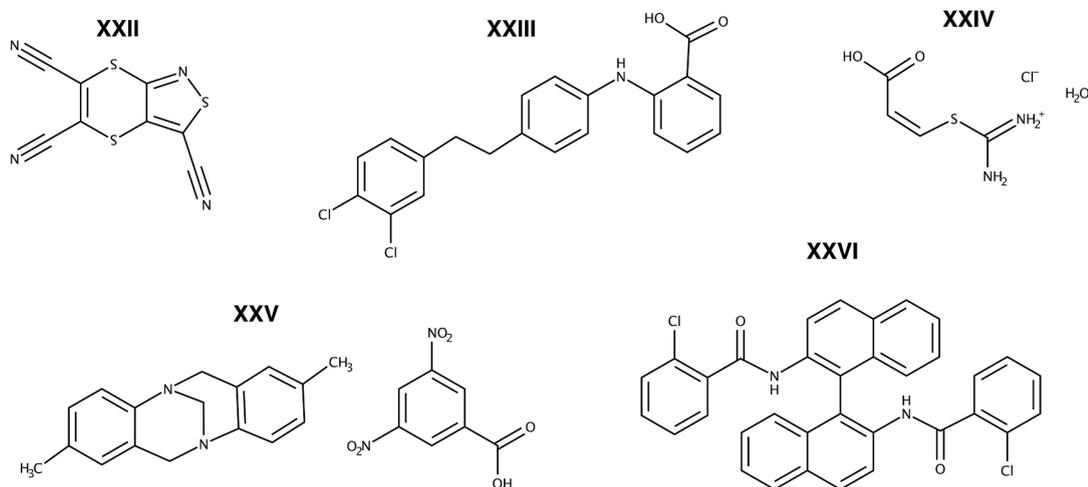


Figure 8.1: The five target systems of the sixth CCDC blind test.

8.2 Computational Methods

For every target system we start with an initial structure set of 1000 low-energy structures, which was provided by the Price group: Rebecca K. Hylton (XXII), Louise S. Price (XXIII), Rui Guo (XXIV), Rona E. Watson (XXV), and Luca Iuzzolino (XXVI). These initial structures were created by a quasi-random search within the CrystalPredictor [177, 178] software and energies were then evaluated using a distributed multipole analysis [179] and an atom-atom empirical potential for the description of exchange repulsion and dispersion interactions [180]. The initial structure set for system XXIV and XXV included also a further optimization with the CrystalOptimizer software [181]. For more details the reader is referred to the Supporting Information of Ref. 51 (Submission 18).

Starting with this sets of 1000 likely structures, we employed a hierarchical procedure, which is illustrated in Fig. 8.2. First, we calculated an initial energy ranking for all 1000 structures without any geometry optimization utilizing the PBE [82] functional supplemented by the pairwise TS [34] dispersion model. These and all subsequent electronic structure calculations were performed using the all-electron code FHI-aims [74, 120–122]. For this initial energy ranking we applied the so-called light species default settings (1) for integration grids and basis functions within FHI-aims. Subsequently, the structures were ranked according to their relative stability per molecule and geometry optimizations with fixed lattice parameters were performed for the 200 most-stable structures. This optimization was also done at the PBE+TS/light level and force components were converged to 10^{-3} eV/Å. Next, full lattice relaxations without any symmetry constraints were performed for the top 100 structures using the same settings as before. Then, the energies of these 100 structures were calculated on top of the PBE+TS-optimized structures using PBE+MBD [35, 64] with tight species default settings. For system XXIII duplicate structures were removed, yielding 31 structures. The relative stabilities at this stage constitute our first submitted lists.

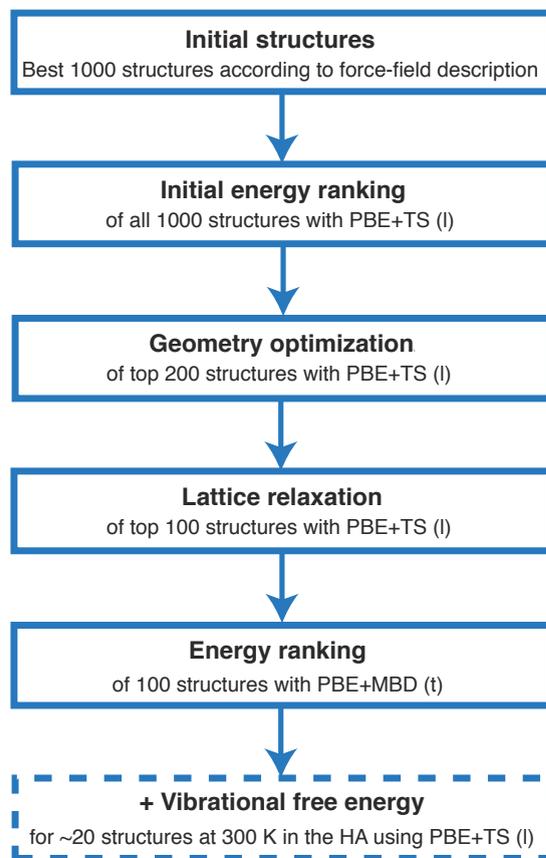


Figure 8.2: Illustration of our complete re-ranking procedure (applied for systems XXII and XXIII).

By now, it is also well-known that vibrational free energies can have a significant influence on the relative stabilities of molecular crystal structures [42]. Therefore, we calculated for about 20 structures for systems XXII and XXIII also vibrational free energies within the harmonic approximation at 300 K. These phonon calculations were carried out with the finite difference method using FHI-aims and phonopy [123]. All force calculations were performed at the PBE+TS level of theory using light species default settings. Finite displacements of 0.005\AA were used throughout and the phonon calculations were always performed on (super)cells extending at least 10\AA in every direction in order to minimize artifacts. The second list includes now stability rankings based on Helmholtz free energies, which are the sum of the static PBE+MBD energy and the vibrational free energy. Due to time reasons, an abbreviated version of this procedure was used for systems XXIV and XXV. For system XXIV, 100 lattice optimizations were performed directly after the initial ranking of the 1000 starting structures and force components were converged to 10^{-2} eV/\AA . Subsequently, PBE+MBD/tight energy calculations were carried out for the top 50 structures. For system XXV, lattice relaxations were performed for the top 20 structures after the initial stability ranking utilizing the

same settings as for system XXIV. Subsequently, PBE+MBD/tight energies were calculated for these 20 relaxed structures. Due to time limitations, system XXVI could not be calculated before the blind test submission deadline. However, the stability ranking for XXVI was calculated in a post analysis utilizing the same approach as for system XXIV.

8.3 Initial Structure Screening

Since full lattice relaxations are not feasible for 1000 structures on a DFT level, we performed a hierarchical screening in order to limit the number of structures. This initial screening is performed at the PBE+TS (light) level. We discuss here the impact of the geometry and lattice optimization on the relative stabilities of system XXIII. Fig. 8.3 illustrates these three steps in terms of the final 31 structures, which are highlighted in all rankings and grouped in intervals of 5 or 6 structures.

The initial ranking based on the force-field structures has a spread of 34 kJ/mol but all highlighted structures are already found within the top 12 kJ/mol. For these structures, a geometry optimization changes relative stabilities on average by 2.0 kJ/mol

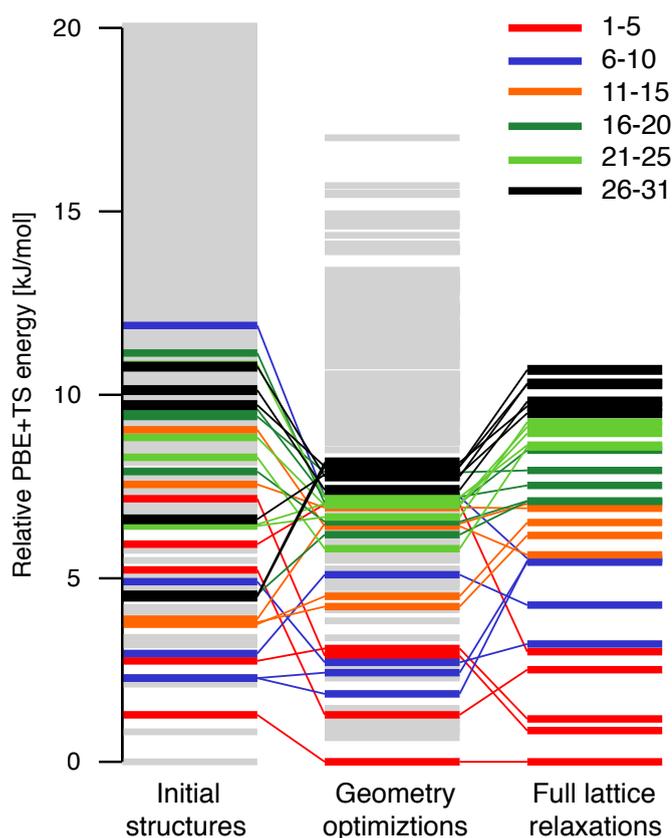


Figure 8.3: Impact of the applied geometry and lattice optimizations on PBE+TS energies used for the initial screening of structures.

with a maximum of 4.7 kJ/mol and a standard deviation of 1.3 kJ/mol. At this stage, the spread of all calculated structures reduces to about 17 kJ/mol. The subsequent lattice optimization further modifies relative energies on average by 1.7 kJ/mol with a maximum of 4.0 kJ/mol and a standard deviation of 1.0 kJ/mol. Therefore, performing full lattice relaxations for the top 8 kJ/mol is statistically sufficient to not miss the most stable structure according to the PBE+TS description. After the lattice optimizations our energy range amounts to 11.5 kJ/mol. While this hierarchical screening is sufficient to not miss important structures according to the PBE+TS description, our final energy ranking is performed using PBE+MBD. The difference between PBE+MBD/tight and PBE+TS/light relative stabilities amounts in this case on average to 3.2 kJ/mol with a standard deviation of 2.2 kJ/mol and an observed maximum of 7.5 kJ/mol. Therefore, it would be advisable to perform more PBE+TS lattice relaxations to ensure that the most stable structure according to PBE+MBD is indeed included in the final structure set.

8.4 Overall Results of the Blind Test

Our final stability rankings are visualized in Fig. 8.4. Experimentally confirmed structures are highlighted in red. Note that system XXVI was not part of our original blind test contribution but was added in a post analysis. Our blind test results are summarized in Table 8.1 along with two other selected submissions. For the full blind test results the reader is referred to Ref. 51.

We will discuss our overall blind test results in context of the submission by Price *et al.* since our submission attempted to provide a first-principle re-ranking of the provided force-field structures. In both submissions, 5 out of the 7 attempted predictions were successful. This illustrates that our applied hierarchical procedure was successful and did not lose any important structure in an intermediate step. Note that form C and E of system XXIII could not have been obtained by the used crystallographic space sampling since only $Z' = 1$ structures were considered. However, form A of system XXIII and the structure of the salt (XXIV) were not successfully predicted. Now let us discuss the obtained stability ranks. Rank 1 in the table indicates that the experimentally

Table 8.1: Blind test results of selected groups [51]. The numbers indicate the rank of the experimentally obtained structure within the submitted lists.

Submission	XXII		XXIII					XXIV		XXV		XXVI						
	L1	L2	A		B		C		D		E		L1	L2	L1	L2		
			L1	L2	L1	L2	L1	L2	L1	L2								
This work	3	1	-	-	2	5			14	2			-		1		(1)	
Price <i>et al.</i> [51]	6	2	-	-	1	2			85	44			-	-	1	1	2	1
Neumann <i>et al.</i> [51]	2		26	85	2	4	-	6	11	39	-	-	2		6		1	1

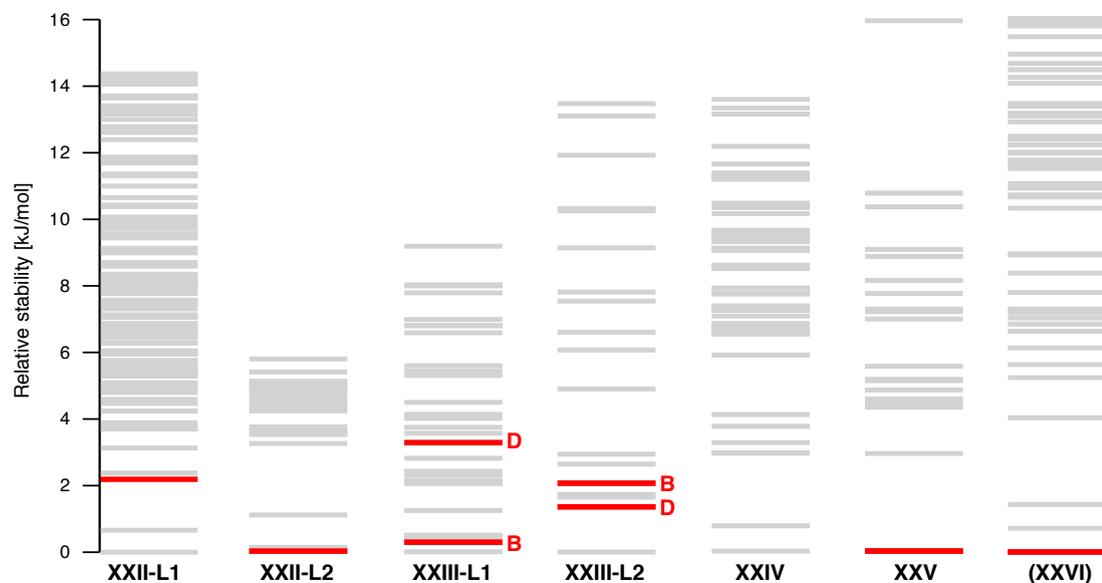


Figure 8.4: Results of our blind-test submission. Experimentally observed structures are shown in red while other structures are shown in gray. The first or only result of each system shows the PBE+MBD ranking while the second lists (L2) include in addition also vibrational free energies at room temperature. Note that the results of system XXVI were obtained after the blind test submission deadline.

obtained structure was indeed predicted as the most stable structure. It can be seen that our ranking — as well as the Price *et al.* ranking — provides a perfect rank (1) for systems XXV and XXVI. For system XXII our free energy ranking (list 2) yields rank 1 for the experimental structure, which is a slight improvement over the results obtained by Price *et al.* (rank 2). For system XXIII we observe a larger difference between the two approaches. In our free energy ranking form B and D are observed within the top 5 structures while the force-field approach used by Price *et al.* yields rank 2 and 44, respectively. This suggests that first-principle approaches can be crucial for flexible systems with a lot of polymorphs in a narrow energy window. Furthermore, it can be seen that our free energy rankings provide significantly better results than our static energies. This highlights again how important thermal effects are for stability rankings at finite temperatures. The accuracy of our PBE+TS-optimized structures in comparison to the experimentally determined unit cells will be discussed in the next chapter.

Finally, we briefly mention one other blind test submission — Neumann *et al.* While this approach does not yield perfect stability rankings, it was able to predict all but one structures of the blind test. This includes form A of system XXIII and in addition also form C, which is a $Z' = 2$ structure. Furthermore, this was the only submission within the blind test, which correctly predicted the salt structure (XXIV). Further details will be discussed in the following chapter.

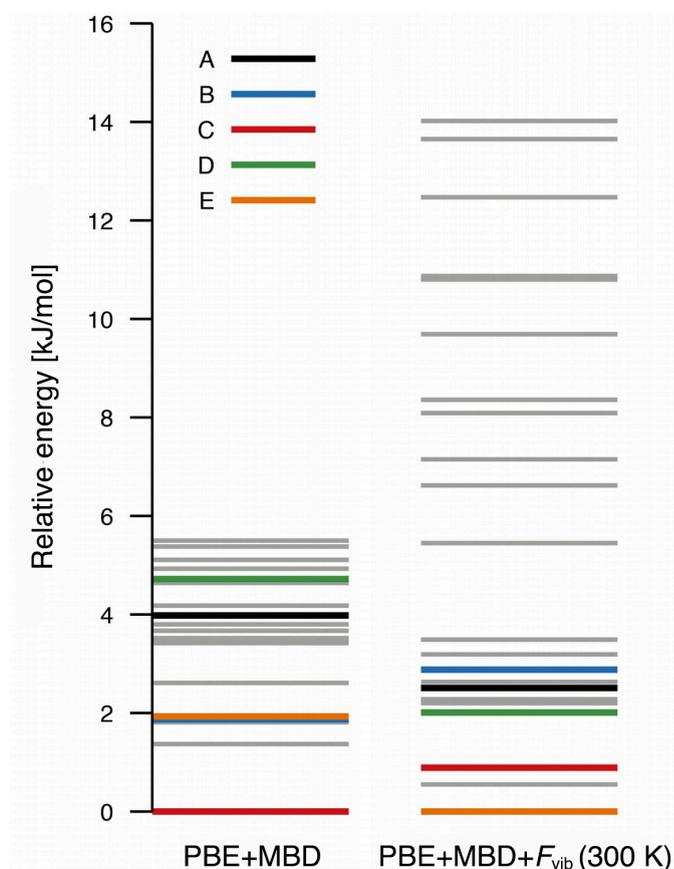


Figure 8.5: Relative stability rankings for system XXIII. Experimentally observed structures are shown in color. Reproduced from Ref. 51 (CC-BY).

8.5 Discussion

Now we provide a short discussion of the individual systems including a post-analysis. For system XXII our first list ranks the experimental structure as rank 3, which is located 2.2 kJ/mol away from the most stable structure. By including vibrational free energies, we are able to correctly rank the experimental structure as rank 1.

For system XXIII we were able to predict forms B and D with ranks 2 and 14 in list 1. After inclusion of thermal effects (list 2) both forms are located within the top 5 structures. Experimental evidence suggests that form D should be the most stable polymorph at room temperature. In our ranking we indeed observe that form D becomes more stable than form B after thermal effects are accounted for. Since we were not able to predict all forms, we performed a post analysis and calculated the relative stabilities for the missed structures using the experimental crystal structures as starting points for the lattice relaxations. The results are added to our obtained stability rankings and shown in Fig. 8.5. According to the PBE+MBD ranking (list 1) form C is the most stable structure and we find all experimentally observed structures within an energy window of about 5 kJ/mol. Adding vibrational free energies (list 2) significantly changes the relative stability ordering and form E becomes the most stable

form. Furthermore, we observe that the inclusion of vibrational free energies increases the spread of the relative stabilities from about 6 to 14 kJ/mol. In contrast, the energy window of the experimentally observed structures is narrowing, and we find now all five forms in an energy window of only 2.9 kJ/mol.

For system XXIV, the experimentally observed structure was not present in our initial structure set and could not be predicted. Our post analysis revealed that the experimentally obtained structure is by 6.4 kJ/mol more stable than our predicted rank 1 structure.

For systems XXV and XXVI, our static stability ranking based on PBE+MBD is able to predict the experimentally obtained structure as rank 1. In both cases the following structures are reasonably separated and therefore vibrational free energies will presumably not significantly impact the rank of the experimental structures. In the case of XXV the second most stable structure is already separated by about 3 kJ/mol. In the case of XXVI we observe only two additional structures within the top 4 kJ/mol, located at 0.7 and 1.5 kJ/mol.

8.6 Summary

In this chapter, we have presented the blind-test results of our hierarchical stability re-ranking procedure. We have shown that this procedure retained all important structures from the initial structure set and we have obtained excellent stability rankings for the experimentally confirmed structures accessible through the available set of initial structures. The results suggest that first-principles approaches are especially relevant for stability rankings of complicated systems involving flexible molecules with numerous low-energy structures. Furthermore, we found that the inclusion of vibrational free energies can lead to a significant stability re-ordering and hence to improved stability ranks.

Chapter 9

Recommended Stability-Ranking Procedure Based on DFT+MBD

After our quite successful blind-test contribution, we wanted to further evaluate and improve our hierarchical CSP re-ranking procedure. Therefore, we discuss and systematically analyze the results obtained by combining our re-ranking approach with the crystallographic space sampling provided by Neumann *et al.* for the blind test systems [51]. In this regard, we are especially interested how our approach deals with salts and less symmetric $Z' = 2$ structures.

The CSP approach by Neumann *et al.* was able to correctly predict all but one experimentally observed structures within the top 100 most stable structures, building on top of their successes in previous blind tests [49, 50, 182]. In this approach, initial structures are generated via a Monte Carlo parallel tempering algorithm utilizing a tailor-made force field [183] within the Grace software package. After this initial screening, a set of candidate structures is then re-optimized in a hierarchical and statistically-controlled process using vdW-inclusive DFT [51, 182, 184]. Beyond the robust sampling of the essential regions of crystallographic space, these initial energy rankings can be substantially improved upon by utilizing state-of-the-art first-principles methodologies as described below.

Herein, we present and discuss a modified version of our hierarchical stability ranking approach, which includes now also hybrid DFA calculations. Furthermore, we also consider thermal expansion via the QHA and estimate anharmonic free energies by using Morse oscillators for selected structures. A version of this chapter has been published in *Science Advances* [185] (CC BY-NC). A part of the lattice optimizations was performed by Hsin-Yu Ko from Princeton University.

9.1 Computational Methods

For each system in the latest blind test, we utilized the top 100 crystal structures from the submission of Neumann *et al.* (which are available in the *Supplementary Information* of Ref. 51) as initial structures for this study. The five blind test systems are illustrated

in Fig. 8.1. For systems with two submitted lists, we used the list which also included $Z' = 2$ structures, *i.e.*, structures which have two molecules in the asymmetric unit. Form E of system XXIII was the only experimental structure not present in this set and was therefore added for completeness. All calculations were performed using the all-electron code FHI-aims [74, 120–122, 160–162]. Throughout this chapter, we utilize two different accuracy levels in FHI-aims, which are denoted as *light* and *tight*. For the *light* level, we use the light species default setting available in FHI-aims for all numerical atom-centered basis functions and integration grids. The number of k -points (n) in each direction was determined by the smallest integer satisfying $n \times a \geq 25 \text{ \AA}$, with a being the unit cell length in a given direction. For the *tight* level, we use the tight species default settings in FHI-aims and the number of k -points is determined by the criterion that $n \times a \geq 30 \text{ \AA}$. Many-body dispersion interactions were evaluated at the MBD@rsSCS level with a reciprocal-space implementation that utilized the same k -point mesh as the DFT calculations [35, 64]. Convergence criteria of 10^{-6} eV, 10^{-5} electrons/ \AA^3 , 10^{-4} eV/ \AA , and 10^{-3} eV were used for the total energy, charge density, forces, and sum of eigenvalues, respectively.

First, we performed full lattice and geometry relaxations (without any symmetry constraints) using the PBE functional [82] in conjunction with the effective-pairwise TS dispersion correction [34], ensuring that the smallest force component is less than 0.005 eV/\AA . Duplicate structures were identified using Mercury [186]. Structures were considered similar if 20 out of 20 molecules within the crystals matched within 25% in terms of distances and within 25° in terms of angles, which are the same criteria used in Ref. 51 to identify matches. Two similar structures were considered to be identical if their PBE+TS energy (*light*) agreed to within 1 kJ/mol and their root-mean-square deviation of 20 molecules (RMSD_{20}) is smaller than 0.5 \AA . This ensures that we are removing real duplicate structures but still consider similar structures with sufficiently different stabilities due to for example slightly different torsion angles. Only the most stable structure among identical structures was retained throughout this protocol. These optimized structures were symmetrized using PLATON [187] and are provided in the Supporting Information of Ref. 185. All structures were named according to their rank in the initial ranking by Neumann *et al.* In order to determine if an experimental structure was found, we used the same settings for the crystal similarity search as described above.

Next, relative energetic stabilities were computed based on these PBE+TS optimized structures by using PBE+TS and PBE+MBD [35, 64] with *tight* settings. In order to ensure the convergence of the relative energies, we have created a benchmark set consisting of 8 small structures of system XXII and 4 small structures of system XXIV. For these structures, PBE+MBD energies were computed using *really tight* settings for the integration grids and tier 3 basis functions (see Table A.1 in Appendix A). When considering all possible relative energies between structures from the same system, the mean absolute deviation (MAD) for the *tight* settings amounts to only 0.1 kJ/mol with a maximal deviation of 0.3 kJ/mol (see Table 9.1). This illustrates the fact that *tight* settings provide already converged relative stabilities.

Since PBE0 [84] calculations with *tight* settings are not possible for all of the studied

Table 9.1: Convergence of relative stabilities with basis set and grid settings. This table shows the mean absolute deviation (MAD) and the maximum absolute deviation (MAX) for all possible relative energies (within a system) from Table A.1 w.r.t. PBE+MBD calculations with *really tight* settings for basis set and grids [185].

Method	MAD [kJ/mol]	MAX [kJ/mol]
PBE+MBD/light	0.9	3.2
PBE+MBD/tight	0.1	0.3

Table 9.2: This table shows the mean absolute deviation (MAD) and the maximum absolute deviation (MAX) for all possible relative energies (within a system) from Table A.1 w.r.t. PBE0+MBD calculations with *tight* settings for basis set and grids. The method PBE0+MBD without a label for basis set and grid settings corresponds to the described PBE0+MBD estimate [185].

Method	MAD [kJ/mol]	MAX [kJ/mol]
PBE+MBD/light	1.6	4.8
PBE+MBD/tight	0.8	1.8
PBE0+MBD/light	0.8	2.6
PBE0+MBD	0.4	0.8

systems due to the massive computational cost and memory requirements, we approximate the PBE0+MBD energies by adding the difference between PBE0+MBD and PBE+MBD evaluated at the *light* level to the PBE+MBD energies calculated at the *tight* level. For the previously mentioned benchmark set, this approximation has a MAD of only 0.4 kJ/mol with a maximum deviation of 0.8 kJ/mol, when compared to PBE0+MBD energies evaluated with *tight* settings (see Table 9.2). In contrast, PBE0+MBD energies at the *light* level yield a MAD of 0.8 kJ/mol with a maximum deviation of 2.6 kJ/mol. Therefore, our approximation provides relative energies that are in very good agreement with *tight* PBE0+MBD energies. PBE0+MBD energies were computed for all structures of systems XXII, XXIII, and XXIV. For the remaining systems, PBE0+MBD calculations are available for (at least) the structures located within the top 4.5 kJ/mol of the PBE+MBD rankings.

Vibrational free energies (F_{vib}) were computed at the PBE+TS level with *light* settings by utilizing phonopy [123] and the finite-difference method within the harmonic approximation. The vibrational free energy F_{vib} was calculated according to Eq. 4.5. The final stability rankings are always based on PBE0+MBD+ F_{vib} energies evaluated at temperatures correspond to the experimental crystal structure measurements. An illustration of the applied hierarchical procedure is shown in Fig. 9.1. For the finite-difference calculations, we used displacements of 0.005 Å and (whenever necessary) supercells that ensure cell lengths greater than 10 Å in every direction. Furthermore, the vibrational free energy was evaluated in reciprocal space, where the number of q -points (n) in each direction is determined by the smallest integer satisfying $n \times a \geq 50$ Å. All structures had

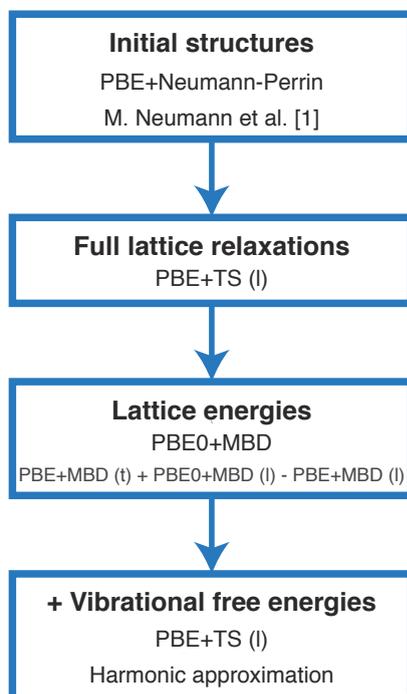


Figure 9.1: Illustration of the applied hierarchical CSP procedure.

no imaginary frequencies at the Γ -point and the magnitude of the three acoustic modes was smaller than 0.1 cm^{-1} in most cases and always smaller than 0.5 cm^{-1} . Vibrational free energies were calculated for (at least) all structures that are located within the top 3 kJ/mol according to the PBE0+MBD ranking. For system XXIII, vibrational free energies were calculated for all $Z' = 1$ structures and for all $Z' = 2$ structures containing up to 8 molecules per unit cell within the top 4.8 kJ/mol of the PBE0+MBD ranking.

For the QHA, we performed PBE+TS lattice and geometry optimizations of several structures from system XXIII using *light* settings with external hydrostatic pressures of 0.4, 0.2, -0.2, -0.4, and -0.6 GPa in order to obtain optimized structures with different unit cell volumes. A negative hydrostatic pressure constitutes a so-called *thermal pressure* [39]. Temperature effects in molecular crystals can significantly affect the unit cell volume. These thermal effects can approximately be accounted for by lattice optimizations under an appropriate thermal pressure, which leads to the volumetric expansion of the cell (see Refs. 13, 39). A minimization of the Gibbs free energy at a certain temperature w.r.t. the cell volume enables the calculation of the corresponding thermal pressure, which is defined as the derivative of the vibrational free energy w.r.t the volume. Then, harmonic vibrational free energies were computed for all of the obtained structures. Based on the *light* PBE+TS energies and harmonic vibrational free energies, the unit cell volume corresponding to 300 K was determined *via* the Murnaghan equation of state [144]. Based on these thermally-expanded structures, the stability rankings were calculated as described above.

For all thermally-expanded structures of system XXIII, we computed the anharmonic vibrational contributions to the free energies by replacing the harmonic oscillators by Morse oscillators. This is done for all phonon modes at the Γ -point of cells containing 4 molecules, *i.e.*, for Forms A, C, D, E, and Str. N70, this corresponds to the unit cell, while for Form B and Str. N18, N31, and N42, this corresponds to a $2 \times 1 \times 1$ supercell. The structures were displaced along all normal modes in both directions, corresponding to energy changes of $0.5 k_B T$ and $k_B T$ according to the harmonic approximation, where k_B is the Boltzmann constant and $T = 300$ K. The energies of all displaced structures were calculated with PBE+TS using *light* settings. To have a consistent sampling of the thermally-accessible energy window, we demanded that the largest observed energy change with respect to the optimized thermally-expanded structure always lies between $k_B T$ and $1.5 k_B T$. Therefore, the displacement amplitudes of a few low-frequency modes had to be reduced in order to sample the desired energy window. Next, we fitted a Morse potential [104, 105], given by

$$V(x) = D \left(1 - e^{-a(x-x_0)} \right)^2, \quad (9.1)$$

to the obtained data points for each mode. In this expression, x is the displacement amplitude, and the parameters D , a , and x_0 describe the well depth, the width of the potential, and the minimum of the potential, respectively. The energy of a vibrational mode in state v can be calculated by

$$E(v) = \hbar\omega_0 \left(v + \frac{1}{2} \right) - \frac{\hbar^2\omega_0^2}{4D} \left(v + \frac{1}{2} \right)^2, \quad (9.2)$$

with

$$\omega_0 = \sqrt{\frac{2a^2 D}{\mu}}, \quad (9.3)$$

where μ is the reduced mass. The anharmonic vibrational free energy (\tilde{F}_{vib}) at the Γ -point was computed according to

$$\tilde{F}_{\text{vib},\Gamma} = -k_B T \ln Q_{\text{vib}}, \quad (9.4)$$

with

$$Q_{\text{vib}} = \prod_i \sum_{\nu} \exp\left(\frac{-E_{i,\nu}}{k_B T}\right), \quad (9.5)$$

where i runs over phonon modes. This approach yields anharmonic vibrational free energies at the Γ -point for cells including 4 molecules. In order to account also for other q -points, we rely on the harmonic approximation and calculate the total vibrational free energies according to:

$$\tilde{F}_{\text{vib}} = F_{\text{vib,full}} + \tilde{F}_{\text{vib},\Gamma} - F_{\text{vib},\Gamma}, \quad (9.6)$$

where $F_{\text{vib,full}}$ is the fully converged harmonic vibrational free energy and $F_{\text{vib},\Gamma}$ is the harmonic vibrational free energy evaluated at the Γ point for the cells described above.

9.2 Energy Ranking Approach

As the foundation for the presented stability ranking approach, we utilize the top 100 molecular crystal structures (for every system of the latest blind test) from the above mentioned sampling approach of Neumann *et al.* using Grace (see Supplementary Information of Ref. 51). Form E of system XXIII is the only experimental structure that was not present in this set of initial structures and is included for completeness. We note in passing that this form was in fact generated by Neumann *et al.*, but was located just outside the energetic window considered for the $Z' = 2$ structures. In total, this set includes 501 structures (with unit cell sizes ranging from 15 to 992 atoms) and therefore provides a large-scale benchmark structural database under realistic CSP conditions.

Based on these initial molecular crystal structures, we have developed a robust hierarchical first-principles approach for energetically ranking all relevant polymorphs. This approach is directly applicable to pharmaceutically relevant systems and includes three important theoretical aspects that are commonly neglected in typical CSP protocols: (i) a sophisticated treatment of Pauli exchange-repulsion and electron correlation effects with hybrid functionals, (ii) inclusion of many-body dispersion interactions and dielectric screening effects, and (iii) an account of vibrational contributions to the free energy. In this regard, the hybrid PBE0 functional [84] in conjunction with the many-body dispersion (MBD) model [35, 64, 89, 188, 189] is able to predict absolute experimental lattice energies to within 1 kcal/mol [39, 40] and relative stabilities of several polymorphic systems to within 1 kJ/mol [40, 60, 88, 117]. Hence, the PBE0+MBD approach is used for all calculations of static lattice energies. Geometry and lattice optimizations, as well as vibrational free energies are computed with the PBE functional [82] in conjunction with the effective-pairwise TS dispersion model [34] (denoted as PBE+TS).

9.3 Polymorphic Energy Landscapes

The stability rankings obtained for the five blind-test systems are shown in Figs. 9.2 and 9.3. (all relative energies are available in Appendix C). Our proposed energy ranking, which includes all of the aforementioned theoretical contributions, is shown for every system in the last column and marked with PBE0+MBD+ F_{vib} . To illustrate the importance of each contribution to the stability ranking, Fig. 9.2 not only shows the final stability rankings, but also several intermediary steps, in which one or more of the three aforementioned theoretical contributions are not accounted for in the rankings. The first ranking considers only static lattice energies computed at the PBE+TS level, while the second ranking accounts for beyond-pairwise many-body dispersion interactions (PBE+MBD). In the third ranking, we include a more sophisticated treatment of Pauli exchange-repulsion *via* PBE0+MBD. In doing so, the deleterious effects of self-interaction error (a DFT artifact in which an electron interacts with itself) are significantly ameliorated, which leads to a substantial improvement in the description of electrostatic and charge-transfer effects. In the final ranking, we supplement the PBE0+MBD energies with harmonic vibrational free energy contributions (+ F_{vib}) at

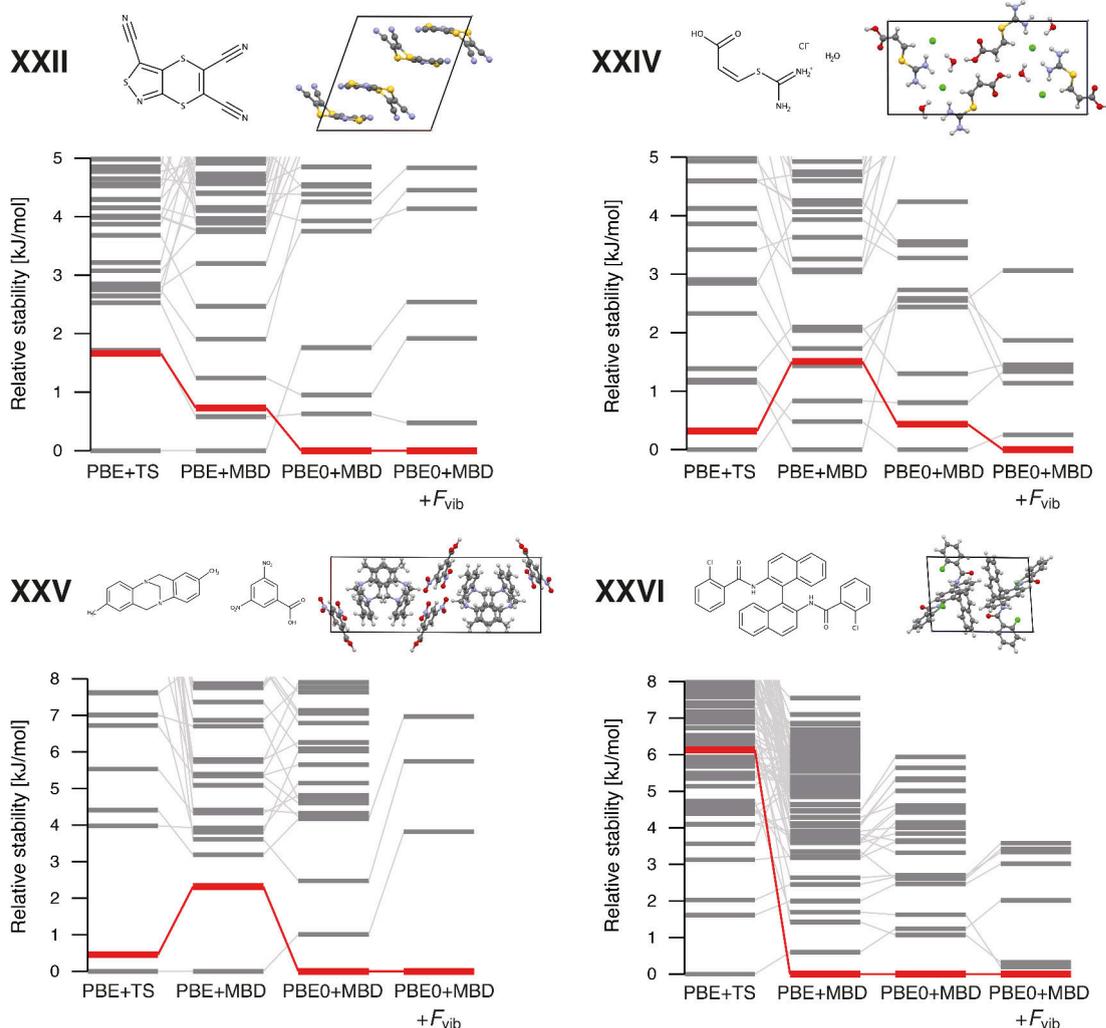


Figure 9.2: Relative stabilities for all steps in the present CSP stability ranking procedure for systems XXII, XXIV, XXV, and XXVI. For each ranking, the energy of the most stable crystal structure defines the zero-of-the-energy. Experimentally observed structures are highlighted in color while all other structures are in gray. The final ranking for each system corresponds to the Helmholtz free energies at the PBE0+MBD+ F_{vib} level, computed at 150 K (XXII), 240 K (XXIV), or 300 K (XXV, XXVI). The energies are normalized per molecule for XXII and XXVI and given per formula unit for XXIV and XXV. Adapted from Ref. 185 (CC BY-NC).

the PBE+TS level. This leads to our proposed PBE0+MBD+ F_{vib} final stability ranking based on Helmholtz free energies, which accounts for thermal entropic effects.

We first concentrate our discussion on systems XXII, XXIV, XXV, and XXVI. For all of these systems, our final stability ranking at the PBE0+MBD+ F_{vib} level predicts the experimental structure as the most stable form—the ideal outcome of any CSP protocol. As seen from the intermediate stability rankings, all of the three previously mentioned theoretical effects are required to obtain this result. For example,

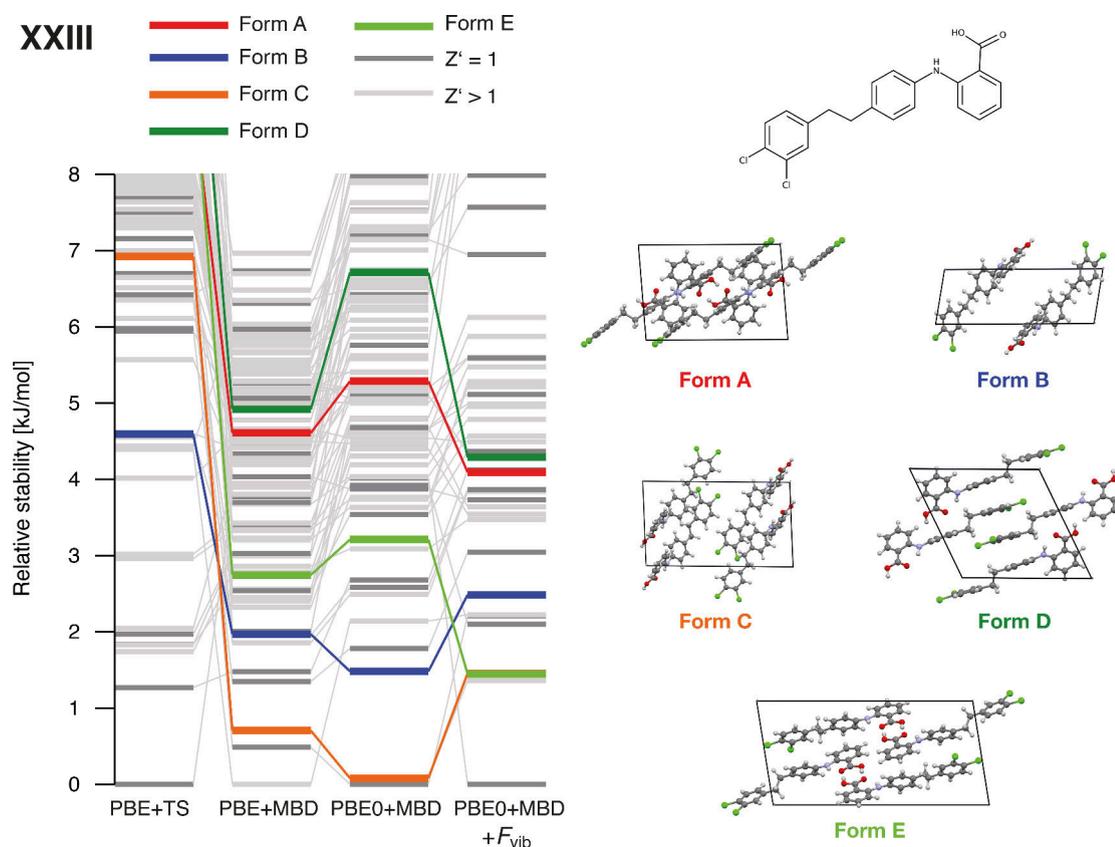


Figure 9.3: Relative stabilities for all steps in the present CSP stability ranking procedure for system XXIII. For each ranking, the energy of the most stable crystal structure defines the zero-of-the-energy. Experimentally observed structures are highlighted in color while all other structures are in gray. The final ranking for each system corresponds to the Helmholtz free energies at the PBE0+MBD+ F_{vib} level, computed at the 300 K. The energies are normalized per molecule. The right part of the figure shows the unit cells for all highlighted structures. Adapted from Ref. 185 (CC BY-NC).

Pauli exchange-repulsion (through the PBE0 functional) plays a crucial role for system XXII [190], while many-body dispersion effects are the most important factor for system XXVI. In addition, all structures with free energies that are within 1 kJ/mol of the experimental structure are essentially minor variations of the latter (see Section 9.5), which demonstrates the robustness of our CSP approach in dealing with pharmaceutically-relevant systems like salts, co-crystals, and molecular crystals involving large and flexible molecules.

Now we focus our discussion on the most challenging system in the blind test (XXIII). This system involves a conformationally flexible molecule and has five experimentally confirmed polymorphs [51]. The fact that this compound is also a former drug candidate [191] makes it an ideal testing ground for CSP of pharmaceutically-relevant molecules. Due to the flexibility of the involved molecule, various conformations are possible within the crystal, leading to a fairly complex polymorphic landscape with numerous crystal structures located within a very small energy window. As shown in

Fig. 9.3, the PBE+TS method is again insufficient for quantitative energy rankings and places all experimentally observed structures within the top 11 kJ/mol—an energy window containing 84 structures. Each refinement of the energetic rankings changes their relative stabilities, with all experimental structures observed within the top 4.3 kJ/mol (≈ 1 kcal/mol) in the final ranking with PBE0+MBD+ F_{vib} . At this level, all experimental structures were found within an energy interval of 3 kJ/mol, which is within the expected energy range associated with co-existing polymorphs [41]. We note here that our procedure finds one structure (Str. N70) that is ≈ 1.5 kJ/mol more stable than all experimentally observed structures, a remarkable finding that is discussed in more detail below.

The computational cost of the procedure presented herein depends heavily on the system size (and other system attributes) and is discussed below based on CPU timings obtained on 2.4 GHz Intel Xeon E5-2680 v4 cores. For the static lattice energies obtained with PBE0+MBD, a single-point energy evaluation (employing the settings described in the *Computational Methods* section) needs 3.5 CPU hours for the smallest unit cell (XXII-N44, 15 atoms) and approximately 750 CPU hours for the largest unit cell (XXV-N39, 896 atoms). For an average-sized system with 172 atoms in the unit cell (*e.g.*, form A of XXIII), a single PBE0+MBD energy evaluation is very reasonable and requires only 60 CPU hours. The computational cost associated with PBE+TS lattice and geometry optimizations depends on the number of optimization steps, but typically amounts to about $2\text{-}3 \times$ that of the corresponding PBE0+MBD single-point energy evaluation. The cost of the harmonic PBE+TS vibrational free energies depends on the size, shape, and symmetry of the unit cell, as these properties determine the required supercell size and the number of finite-difference displacements. For example, the time required for the PBE+TS vibrational free energy calculations ranges from 180 CPU hours (XXII-N2) to 45,000 CPU hours (XXIII-N3). In the case of structure XXII-N2, only 90 finite-difference displacements were required and the employed supercell consisted of 120 atoms, while the calculation involving structure XXIII-N3 required 516 finite displacements and a supercell containing nearly 1,400 atoms. The vibrational free energy calculation for the average-sized form A of system XXIII requires 258 finite displacements and a supercell equivalent to the unit cell, which results in a computation time of 750 CPU hours. Overall, we note that these computational resources are within reach of academic institutions and industrial laboratories.

9.4 Stability Ranking Results

For all systems with only one known polymorph, the systematic and hierarchical energy ranking protocol presented herein (PBE0+MBD+ F_{vib}) correctly produced the experimental structure as the most stable forms (RANK 1). This represents a significant improvement over the RANKS 2 (XXII), 2 (XXIV), 6 (XXV), and 1 (XXVI) obtained by the unrefined results of Neumann *et al.*, which again stresses the critical importance of an energy ranking protocol based on state-of-the-art first-principles based methodologies.

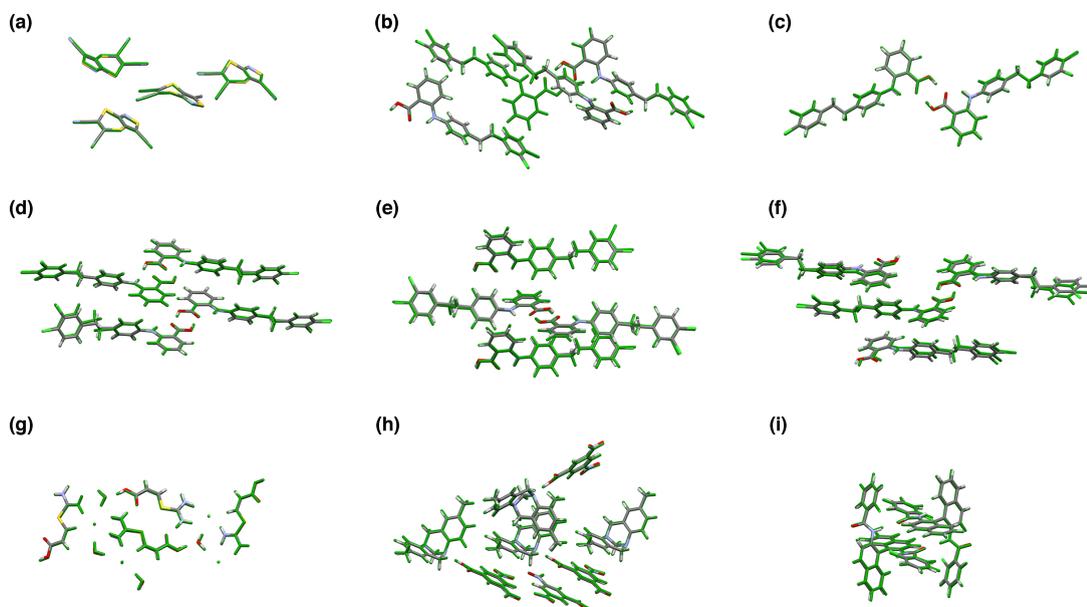


Figure 9.4: Overlay between the experimentally determined structures (element-specific colors) and the corresponding PBE+TS optimized structures for systems (green): (a) XXII, (b) XXIII-A, (c) XXIII-B, (d) XXIII-C, (e) XXIII-D, (f) XXIII-E, (g) XXIV, (h) XXV, and (i) XXVI. These overlays are shown for the molecules constituting the respective unit cell. Reproduced from Ref. 185 (CC BY-NC).

For system XXIII, all experimental structures were found within the top 18 structures, with the two $Z' = 2$ structures assigned RANK 3 (Form E) and 4 (Form C). When only considering the $Z' = 1$ structures, we find all three experimental structures among the top 10 structures, as compared to the top 26 in the initial ranking by Neumann *et al.* Moreover, all of our predicted structures agree to within 0.5 Å of the experimental structures as quantified by the root-mean-square deviation (RMSD) measure of a cluster consisting of 20 molecules. These so-called RMSD₂₀ values also agree to within 0.05 Å with the RMSD₂₀ values of the initial structures obtained by Neumann *et al.* Overlays of the predicted and experimental structures are provided in Fig. 9.4 and additional information about the accuracy of the structures is provided in Appendix B.

9.5 Discussion of Most-Stable Unobserved Structures

In this section, we briefly discuss several structures, which have not been observed in experiment but are very close in terms of stability to experimentally confirmed structures. Let us first discuss the four systems with only one known experimental structure (XXII, XXIV, XXV, XXVI). We always discuss the relative stabilities in terms of our recommended PBE0+MBD+ F_{vib} ranking. For system XXII we find one structure (XXII-N3), which is only 0.5 kJ/mol less stable than the experimental structures. However, this structure is very similar to the experimental one. In our similarity search we find that

15 out of 20 molecules match with the experimental structure. For system XXIV the first three non-observed polymorphs (XXIV-N3, XXIV-N10, XXIV-N53) are found within 1.3 kJ/mol of the experimental structure. All of them are a partial match w.r.t. the experimental structure, i.e., at least 10 molecules out of 20 match with the experimental structure. In system XXV there are no structures within 3.8 kJ/mol. For system XXVI we find two structures which are within 0.3 kJ/mol of the experimental structure. Also in this case, both of them are very similar to the experimental structure. Structure XXVI-N5 is a partial match (16/20 molecules) and structure XXVI-N4 is in fact a complete match but has a large RMSD of 0.9 Å. Therefore, all structures which are within 1.3 kJ/mol of these experimental structures have several structural features in common with the experimental structure.

Finally, let us discuss the former drug candidate XXIII. For this system we find two structures which are described as more stable than any experimental structure (XXIII-N70 and XXIII-N5). Structure XXIII-N70 is a partial match for form A (13/20 molecules) and will be discussed in more detail below. Structure XXIII-N5 is a structure with two molecules in the asymmetric unit ($Z' = 2$) and is virtually identical with form C. However, given the used settings for identical structures, it could not be excluded from the list since only 17 out of 20 molecules are considered a match. In addition, we find several $Z' = 1$ and $Z' = 2$ structures which are structurally sufficiently different and thermodynamically stable enough to potentially crystallize in experiment.

Our most stable structure (XXIII-N70) and form A (XXIII-N85) are structurally very similar, but only form A is observed experimentally despite the fact that structure N70 is predicted to be 4.1 kJ/mol more stable. Even our anharmonic free energy estimate will not significantly change the relative stability. This energy difference is too large to be simply attributed to a computational error. If form A crystallizes instead of structure N70, form A must be dynamically favored. The vibrational free energy stabilizes structure N70 by 2.8 kJ/mol compared to form A. Even at the PBE0+MBD level, structure N70 is still more stable than form A, but only by 1.3 kJ/mol. The vibrational free energy difference is calculated assuming crystals of infinite size. Free energy may not yet stabilize structure N70 over form A during the nucleation phase, when crystallites are small. However, since even without the vibrational free energy contribution structure N70 is more stable than form A, this observation cannot explain why structure N70 does not crystallize at all. Since the only notable difference between the two structures is the stacking of the sheets (see Fig. 9.5, it may be concluded that under the explored crystallization conditions crystal growth perpendicular to the sheets is significantly faster for the form A stacking than for the N70 stacking, potentially related to 2D nucleation on the surface. Extremely slow crystal growth may be the key to obtaining structure N70 rather than form A. Since structure N70 and form A share the same molecular conformation, the same hydrogen bonding and much of their surface chemistry, solvents favoring form A over forms B, C, D and E should also favor structure N70 over the other forms. Extremely long slurring or extremely low melting starting from form A may be alternative ways to obtain structure N70. The main challenge will be to avoid conversion to forms B, C, D or E before conversion to structure N70 is achieved.

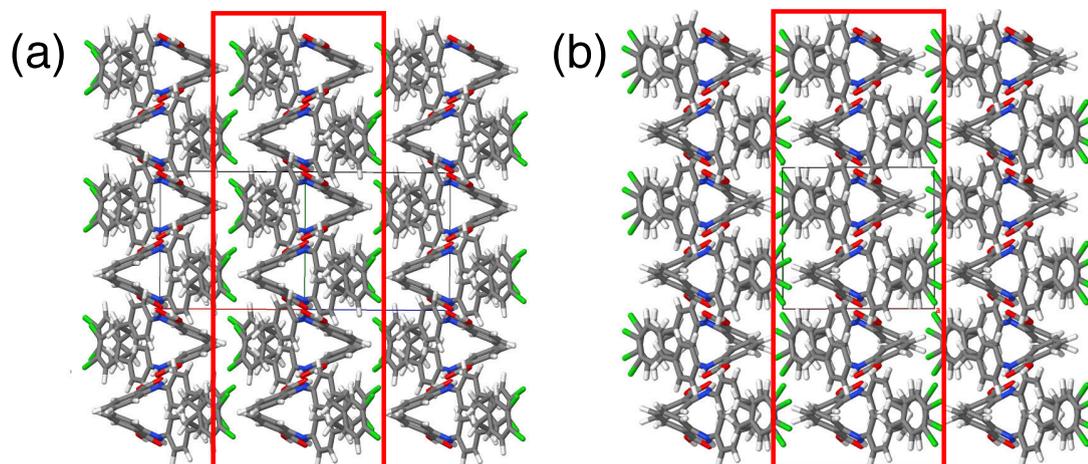


Figure 9.5: Structural differences between structure N70 (a) and form A (b) of system XXIII. Both structures share a sheet structure (red boxes) in which the molecules are arranged according to the same pattern, but the sheets are stacked differently.

9.6 Beyond the Harmonic Approximation

While our presented approach yields very good stability rankings, vibrational free energies were only calculated within the harmonic approximation on top of fully relaxed structures. Therefore, the geometry and lattice optimizations did not include temperature (thermal expansion) effects and the calculated vibrational free energies lack anharmonic effects. The missing thermal expansion can be seen in our obtained unit cell volumes. Comparing the obtained PBE+TS unit cell volumes with the experimental volumes measured at 300 K (all experimental structures except XXII and XXIV) shows that we underestimate these unit cell volumes by 3.6% on average. With five known polymorphs, system XXIII is the most experimentally studied system and it exhibits the most complicated polymorphic energy landscape among the systems investigated in this work. As such, we specifically address how thermal expansion and anharmonicity affect a small set of XXIII structures, from which we estimate their effect on relative stabilities in general. This set includes all experimentally observed structures of system XXIII (forms A, B, C, D, E) as well as the first four $Z' = 1$ structures (Str. N70, N31, N18, N42), which have yet to be experimentally observed.

The effects of thermal expansion can be calculated in the so-called quasi-harmonic approximation (QHA) [143], in which vibrational free energies are calculated within the harmonic approximation for several unit-cell volumes. The unit-cell volume corresponding to a certain temperature is then determined by the minimum of the Helmholtz free energy at that temperature, which is evaluated by fitting the energy-volume curves to the Murnaghan [144] equation-of-state. It has been shown for several molecular crystals that the QHA is capable of capturing a majority of the thermal expansion [13, 147, 148, 192]. Here, we calculated the unit-cell volumes corresponding to a temperature of 300 K using PBE+TS. With this approach, we are now able to predict room-temperature unit cell volumes to within 1.0% on average. As such, the QHA provides a simple but

effective way of including thermal effects in molecular crystal structures using first-principles based methodologies. A detailed comparison of these thermally expanded structures with experiment is available in Appendix B. Stability rankings calculated with PBE0+MBD+ F_{vib} on top of these thermally-expanded structures are shown in Fig. 9.6. We note in passing that these relative Helmholtz free energies can also be interpreted as relative Gibbs free energies since the additional pV term has only a negligible effect at ambient pressure. The largest observed change in relative stabilities stemming from the use of thermally-expanded structures (as compared to fully optimized 0 K structures) amounts to 1.4 kJ/mol at the PBE0+MBD+ F_{vib} level, but is only 0.4 kJ/mol on average. Therefore, we observe some re-ordering of stability rankings, but the general picture and the energy interval remain essentially the same.

In addition to thermal expansion, the vibrational contributions to the free energy also contain anharmonic effects. Here, we estimate these anharmonic effects by replacing the harmonic oscillators obtained *via* the harmonic approximation by Morse oscillators [104, 105]. The Morse oscillator models a particle in an anharmonic potential, for which dissociation is possible. It is the next logical step after the harmonic oscillator since it is also one of the few quantum-mechanical model systems, for which an analytic solution of the Schrödinger equation is known [66]. The Morse oscillator provides a more realistic picture than the harmonic oscillator since it has a finite number of non-equispaced energy levels. This model has been used to describe the spectra of diatomic molecules by improving upon harmonic vibrational frequencies for the hydroxyl groups in methanol, phenol, thymol, and the water dimer [66, 193, 194]. Here, we create four displaced structures per vibrational mode and use the corresponding PBE+TS energies to fit the parameters of the Morse potential and hence determine the Morse vibrational free energies. In this work, the Morse oscillators are independent of each other, *i.e.*, we do not account for coupling between vibrational modes. The corresponding free energy stability rankings with such an anharmonic treatment of the vibrational free energy are denoted by PBE0+MBD+ \tilde{F}_{vib} and shown in Fig. 9.6. At this level, all experimental structures are found within an energy window of only 1.5 kJ/mol, which is well within the expected energy range for co-existing polymorphs. We note in passing that Brandenburg and Grimme have also studied the experimental structures of system XXIII utilizing a semi-empirical tight-binding approach within the QHA; however, their values lie within a much larger energy window of ≈ 8 kJ/mol [58].

Quite interestingly, the unobserved polymorph of XXIII (Str. N70) is significantly more stable than any of the experimentally determined crystal structures, even after accounting for thermal expansion in the underlying crystal structures as well as anharmonic vibrational free energy contributions. In this regard, this polymorph is actually further stabilized by vibrational entropy and shares many structural features with form A. The most notable difference is the stacking pattern of the molecular sheets as discussed above. As such, we hypothesize that Form A might be kinetically favored over Str. N70 and this hitherto unobserved polymorph could potentially be crystallized by slowly melting Form A or introducing surfactants during the crystallization procedure. In addition, from a thermodynamic standpoint, Str. N18, N31, and N42 might also be observed experimentally, although Str. N42 involves a twisted molecular conformation

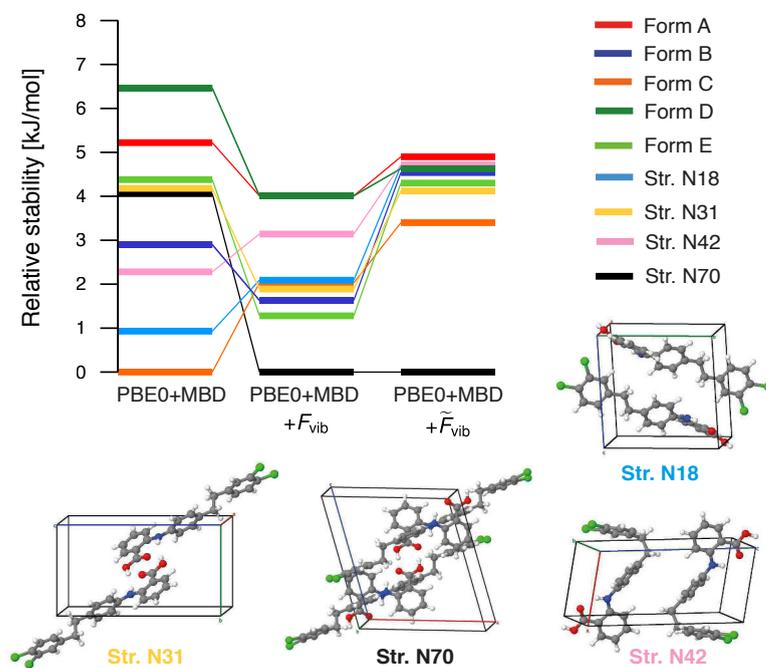


Figure 9.6: Stability rankings for thermally-expanded structures. Energetic rankings for all experimentally observed (Form A, B, C, D, E) and theoretically-predicted (N18, N31, N42, N70) structures for system XXIII. All energies were evaluated using thermally-expanded PBE+TS structures optimized at 300 K with the QHA. The last two rankings include harmonic (F_{vib}) and Morse anharmonic (\tilde{F}_{vib}) vibrational free energy contributions. Reproduced from Ref. 185 (CC BY-NC).

which might not be easily accessible in solution. Experimental evidence [51] suggests that Form A should be the most stable structure at low temperatures and Form D the most stable structure at room temperature. Indeed, we observe that Form D is stabilized by thermal effects and predicted to be more stable than Form A at the PBE0+MBD+ \tilde{F}_{vib} level. In addition, inclusion of anharmonic vibrational free energies brings all of the experimentally determined structures closer together, *i.e.*, all of the $Z' = 1$ structures are now within 0.4 kJ/mol.

In terms of the computational cost, the QHA adds 4,700 CPU hours to the 950 CPU hours required for the initial PBE0+MBD+ F_{vib} calculation in the average-sized form A of system XXIII. For comparison, the corresponding Morse free energy calculation needs an additional 4,300 CPU hours when all Γ -point modes are taken into account. Therefore, the QHA and subsequent Morse free energy calculation increases the computation time by approximately a factor of ten.

9.7 Summary

We have introduced a robust and computationally feasible procedure that yields accurate and reliable descriptions of the structures and stabilities for the thermodynamically

relevant polymorphs of molecular crystals. The diverse set of systems studied in this work includes complex molecular crystals such as a salt, a co-crystal, and crystals consisting of flexible large molecules of pharmaceutical interest. Our approach explicitly accounts for all relevant enthalpic and entropic effects, including sophisticated treatments of Pauli exchange-repulsion, many-body dispersion interactions, and vibrational free energies at finite temperatures, all of which are directly obtained from quantum-mechanical calculations. The approach presented herein takes us one step closer to obtaining an enhanced fundamental understanding of polymorphic energy landscapes and to routinely employing computational molecular crystal structure prediction in conjunction with experimental polymorph screening. Such a joint theoretical-experimental procedure offers a comprehensive and sustainable solution to the grand challenges associated with molecular crystals polymorphs, whose very existence offers us the promise of novel and hitherto unexplored pharmaceutical agents on one hand, and quite devastating public health and economic repercussions on the other.

Chapter 10

Benchmarking the DFT+MBD CSP Procedure

In the previous chapter we have discussed the results obtained with our recommended CSP procedure. However, due to the hierarchical nature, several approximations have to be made. Therefore, we evaluate and discuss in this chapter the influence of several settings and approximations made within our procedure. The results presented in this chapter have been published in *Faraday Discussions* [195] and the following text is adapted from Ref. 195 with permission from The Royal Society of Chemistry.

10.1 Computational Methods

For the benchmark of our stability-ranking procedure described in Chapter 9 we turn again to the systems of the latest CCDC blind test [51] (see Fig. 8.1) and utilize all experimentally confirmed structures and in addition a selected structure set for systems XXII and XXIII (see Fig. 10.1). All calculations were performed by using the all-electron code FHI-aims [74, 120–122, 160–162]. We use in general two accuracy levels, which are labeled *light* and *tight*. The light level (l) corresponds to the light species default settings in FHI-aims in terms of integration grids and basis functions used to describe the Kohn-Sham orbitals. The number of used k -points (n) in each direction amounts to the smallest integer number satisfying the relation $n \times a \geq x$, with a describing the unit cell length in this direction and x being 25 Å. The tight level (t) corresponds to the tight species default settings within FHI-aims and the number of k -points is determined as for the light settings but with $x = 30$ Å. The following convergence criteria are used for all calculations: 10^{-6} eV for the total energy, 10^{-5} electrons/Å³ for the charge density, 10^{-4} eV/Å for the forces, and 10^{-3} eV for the sum of the eigenvalues. The employed naming scheme for structures corresponds to the labels used in the previous chapter [185].

Full lattice relaxations and geometry optimizations were performed by using the following methods: i) PBE+TS [34, 82] (pairwise Tkatchenko-Scheffler dispersion model), ii) PBE+MBD [35, 64] (many-body dispersion model at the MBD@rsSCS level), and iii) PBE0+MBD [84]. These optimizations were performed using the light settings and the

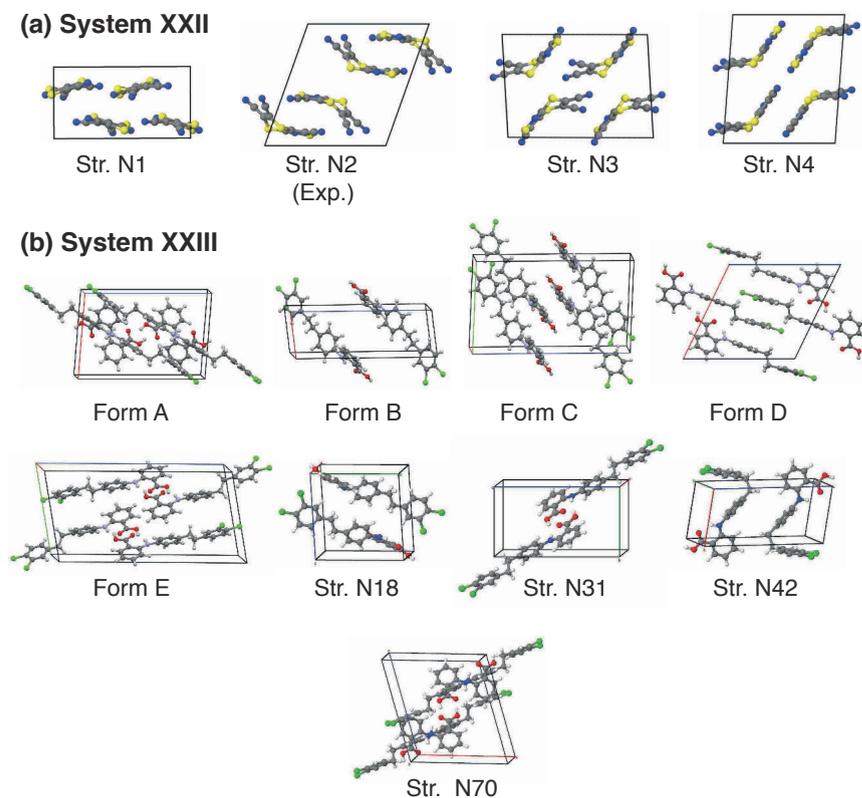


Figure 10.1: Visualization of the unit cell for all explicitly discussed structures. The coordinates of all optimized structures are available in the Supporting Information of Ref. 195. Adapted from Ref. 195 with permission from The Royal Society of Chemistry.

force components were converged to $0.005 \text{ eV}/\text{\AA}$. In addition, the PBE+MBD optimizations were also carried out utilizing the tight settings. The PBE+TS structures [185] were used as starting point for all other optimizations. The optimized structures were symmetrized by using PLATON [187] and are provided in the Supporting Information of Ref. 195. The agreement of the optimized structures w.r.t. experimental structures is measured by the root-mean-square-deviation (RMSD) of 20 molecules within the crystal (RMSD_{20}) utilizing Mercury [186].

The relative static lattice energies based on DFT total energies were evaluated at the PBE0+MBD level. These energies are obtained by calculating the difference between PBE0+MBD and PBE+MBD energies at the light level and adding this difference to PBE+MBD energies obtained at the tight level. This provides a good approximation of converged PBE0+MBD energies [185] (see Chapter 9) and reduces the computation time by about a factor of 10 compared to tight settings. All presented stability rankings are normalized per chemical unit (per molecule for XXII, XXIII, XXVI; per formula unit for XXV and XXIV).

Vibrational free energies were calculated utilizing PBE+TS, PBE+MBD, and PBE0+MBD within the harmonic approximation using phonopy [123] on top of structures optimized at the same level. In general, vibrational free energies were calculated by using

light settings. In addition, several structures were also calculated with PBE+MBD using tight settings. The temperature used to evaluate the free energies corresponds always to the temperature at which the experimental crystal structures were measured, i.e., 150 K for system XXII and 300 K for system XXIII. Displacements of 0.005 Å were used throughout for the finite displacements. Unless stated otherwise, we used supercells which extend at least 10 Å in every direction and the number of q -points for the reciprocal space sampling was calculated in the same way as the k -grid mentioned above, but using $x = 50$ Å.

Thermally expanded structures corresponding to 300 K were calculated for several structures of system XXIII within the QHA utilizing the PBE+TS method with light settings [185] (see Chapter 9). In addition, the room temperature structure for XXIII-A was calculated within the QHA utilizing entirely PBE+MBD with light settings for energies and vibrational free energies. Furthermore, we also calculated the vibrational free energies using PBE+MBD for the experimentally determined structures of system XXIII. For that, we fixed the lattice vectors to the experimental values and optimized only the atomic positions within the unit cells.

Finally, Morse free energies were calculated for several structures of system XXIII using PBE+TS and PBE+MBD with light settings as described in Chapter 9. This was done for the optimized structures as well as for thermally expanded structures (QHA structure at 300 K for PBE+TS and experimental unit cell for PBE+MBD). The method used for calculating the energy of the displaced structures corresponds always to the method used for the harmonic vibrational free energies.

10.2 Static Lattice Energies

First, we further analyze the different stability rankings we have discussed in the previous chapter. Therein, PBE+TS, PBE+MBD, and PBE0+MBD energies were calculated on top of PBE+TS-optimized structures for all systems of the latest CSP blind test [51]. We will use throughout the labeling scheme of the blind test for all systems. This very diverse set of systems includes a crystal consisting of a small and rigid molecule (XXII), a highly polymorphic system comprised of a flexible molecule (XXIII), a salt (XXIV), a co-crystal (XXV), and a crystal involving a large molecule (XXVI) (see Fig. 8.1). The accurate description of all these systems using a consistent CSP method is challenging due to the varying system sizes and their different chemical nature.

Let us start by discussing the importance of many-body dispersion (MBD) interactions and exact exchange for relative static lattice energies. Fig. 10.2 shows a correlation plot between PBE+TS and PBE+MBD energies for all systems. The most stable structure according to the PBE+MBD energies is always set to zero and all relative energies are normalized per chemical unit (per molecule for XXII, XXIII, XXVI; per formula unit for XXV and XXIV). In addition, perfect correlation is indicated with the black line and intervals of 1 kJ/mol and 1 kcal/mol around this line are marked with the dotted lines. It can be seen that for no system all data points are within the ± 1 kcal/mol energy window. This immediately illustrates the crucial nature of MBD interactions given the

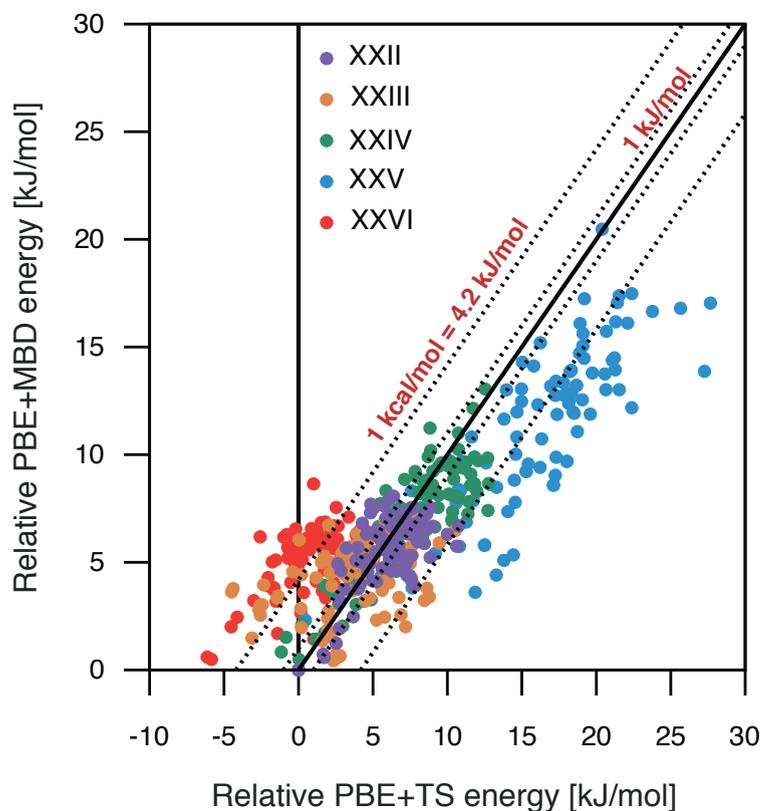


Figure 10.2: Correlation between PBE+TS and PBE+MBD energies calculated on top of PBE+TS-optimized structures[185]. The most stable PBE+MBD structure is always set to zero and energies are normalized per chemical unit. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

energy window for co-existing polymorphs. Furthermore, it can be seen that even re-fitting a pairwise approach would not yield a significant improvement of the correlation, given the large spread of the data points.

In order to numerically quantify the impact of MBD and exact exchange effects, we calculated and compared for every system all possible relative energies between polymorphs. To illustrate the effect of MBD interactions we discuss the deviation between the PBE+TS and PBE+MBD relative energies in terms of the mean absolute deviation (MAD) and the maximum absolute deviation (MAX). These values are shown for every system in Table 10.1. For system XXIII, the absolute deviations in relative energies are also illustrated as Box plots in Fig. 10.3.

It can be seen that the average impact of MBD interactions on relative stabilities varies between 1.8 and 3.2 kJ/mol. As expected, the least impact is found for the two smallest systems involving rigid molecules. The largest impact is found for the former drug candidate XXIII, which involves a quite flexible molecule, and the co-crystal XXV. Note that the maximal observed change in relative energies is for all systems larger than 7.7 kJ/mol. Therefore, MBD interactions clearly cannot be neglected for any of the studied systems.

Table 10.1: Statistics for the impact of several interactions on all possible relative energies expressed in terms of the mean absolute deviation (MAD) in kJ/mol and the maximum absolute deviation (MAX) in kJ/mol for all studied systems [185]; PBE+MBD vs. PBE+TS (MBD), PBE0+MBD vs. PBE+MBD (PBE0), PBE0+MBD+ F_{vib} vs. PBE0+MBD (F_{vib}). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

System	MBD		PBE0		F_{vib}	
	MAD	MAX	MAD	MAX	MAD	MAX
XXII	1.8	7.7	1.1	4.6	0.5	1.0
XXIII	3.2	13.4	1.1	4.2	1.9	8.4
XXIV	1.9	7.8	2.0	8.2	0.9	1.8
XXV	3.2	15.2	1.4	5.1	1.6	3.2
XXVI	2.4	8.7	0.7	2.3	1.0	2.3

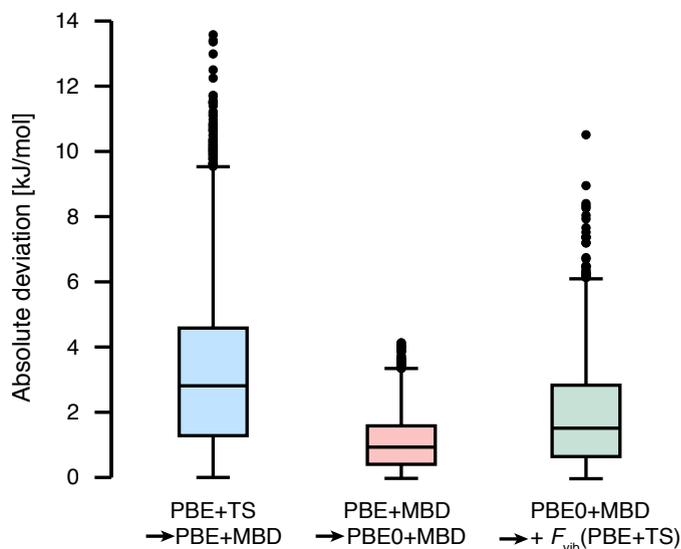


Figure 10.3: Box plots of all possible absolute deviations in relative energies for system XXIII.

Since MBD captures many-body interactions up to infinite order, we now investigate the convergence of the relative energies. We can decompose the MBD energy into orders n (n -atom contributions) according to Eq. 3.37. We illustrate this effect by using the five experimentally observed forms of system XXIII and in addition also 4 low-energy structures identified in Ref. 185 (see Fig. 10.4). These structures are visualized in Fig. 10.1. The plot shows the relative PBE+TS and PBE+MBD energies, with the MBD interactions evaluated up to the shown n -atom contribution.

Already the 2-atom (pairwise) MBD energy is closer to the final PBE+MBD result than the PBE+TS result. This is due to the fact that MBD includes in addition to

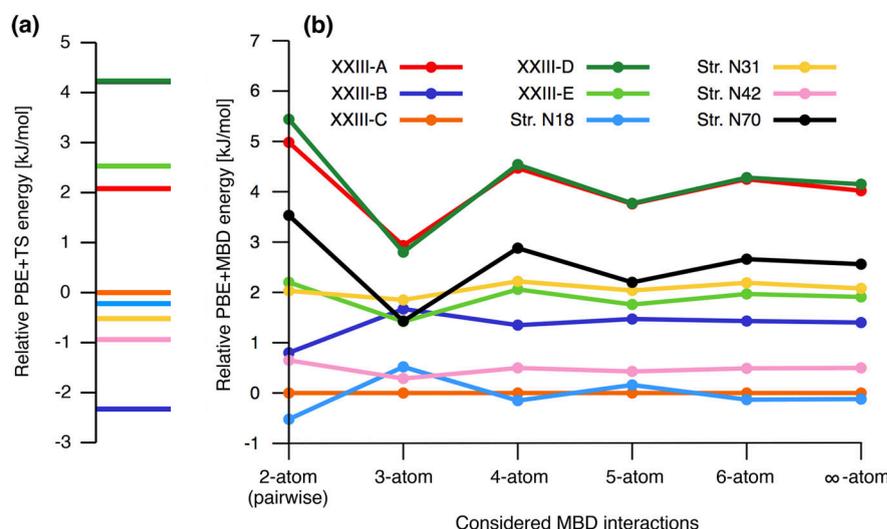


Figure 10.4: Relative PBE+TS (a) and PBE+MBD (b) energies for several structures of system XXIII. The energy of form C serves as reference and is set to zero. Part (b) shows the relative PBE+MBD energies when MBD interactions are included up to the shown n -atom contribution. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

many-body dispersion effects also dielectric screening effects. When we now include 3-atom interactions, the relative energies change up to 3.7 kJ/mol. However, the relative energies are by far not converged yet. Accounting in addition for 4-atom contributions still changes relative energies up to 2.4 kJ/mol. Due to the oscillatory behavior only the step from 5-atom to 6-atom contribution finally changes relative energies by less than 1 kJ/mol.

Now let us move on to the effect of exact exchange, i.e., switching from PBE to the hybrid functional PBE0 (see Table 10.1). This step improves for example the description of electrostatics and charge transfer. In general, we can see that this effect is less pronounced than the MBD effects with MADs varying between 0.7 and 2.0 kJ/mol. However, since the maximal changes are between 2.3 and 8.2 kJ/mol this effect is significant enough to change stability rankings. Exact exchange has as expected the largest effect on the salt, since this system has the largest electrostatic interactions. In addition, PBE0 is for example crucial for correctly identifying the experimental structure of system XXII as the most stable structure [185, 190].

10.3 Geometries

All stabilities in our hierarchical CSP procedure discussed in Chapter 9 [185] have been calculated on top of PBE+TS-optimized structures. Therefore, one natural step to improve this procedure would be to include MBD interactions also in the lattice relaxations and geometry optimizations. However, this would lead on average to an increase of force calculation times of about 50 %. Therefore, we evaluate in this section the effect of MBD

Table 10.2: Errors of optimized structures w.r.t. experimental measurements [51]; mean relative deviation (MRD), mean absolute relative deviation (MARD) as well as maximum absolute deviation (MAX) in percent. The labels l and t indicate light and tight species default settings within FHI-aims, respectively. For the cell angles we only take into account values which are not exactly 90 degrees due to symmetry. Adapted from Ref. 195 with permission from The Royal Society of Chemistry.

Method	Quantity	MRD	MARD	MAX
PBE+TS/ l	Lengths [%]	-0.8	1.4	4.0
	Angles [%]	0.8	1.5	4.3
	Volume [%]	-2.4	3.1	4.7
PBE+MBD/ l	Lengths [%]	-0.3	1.4	4.7
	Angles [%]	0.8	1.6	3.2
	Volume [%]	-0.9	2.2	3.8
PBE0+MBD/ l	Lengths [%]	-1.3	1.6	3.8
	Angles [%]	0.8	1.7	4.3
	Volume [%]	-3.7	3.8	5.9
PBE+MBD/ t	Lengths [%]	-0.3	1.4	4.1
	Angles [%]	0.7	1.4	2.7
	Volume [%]	-1.0	2.6	5.2

and also exact exchange effects on structures and corresponding stability rankings. We have re-optimized all structures from Ref. 185 corresponding to experimentally observed polymorphs using PBE+MBD and PBE0+MBD with light settings. In addition, the structures were also optimized with PBE+MBD using tight settings to study the impact of the used basis functions and integration grids. All optimized geometries are available in the Supporting Information of Ref. 195 and the errors w.r.t. the experimental measurements are shown in Table 10.2 (see also Appendix B). Let us start by discussing the accuracy of the PBE+TS-optimized structures. It can be seen that on average we can describe cell lengths and angles within 1 % of the experimental value and the mean absolute relative deviation (MARD) amounts to 1.4 and 1.5 %, respectively. Furthermore, all errors are less than 5 %. However, we underestimate the unit cell volume on average by 2.4 %. This is to be expected since molecular crystals tend to expand with temperature and our cell relaxation does not include any thermal effects. The unit cell volume can be improved by the utilizing the QHA, which will be discussed later on. Note that while we are underestimating most unit cell volumes, we are overestimating the volume of systems XXII and XXIV. In addition, the quality of the molecular structure is assessed in terms of the RMSD₂₀. On average our RMSD₂₀ w.r.t the experimental structures amounts to 0.25 Å with a maximum of 0.48 Å (XXIII-E). The possibly disordered nature of form E [51] is most likely the reason for the observed relatively large RMSD₂₀. Overall, our PBE+TS-optimizations yield already structures of good quality.

Now let us turn to the PBE+MBD (light) optimizations. The description of cell lengths and angles is comparable with the PBE+TS results, but the errors in the cell

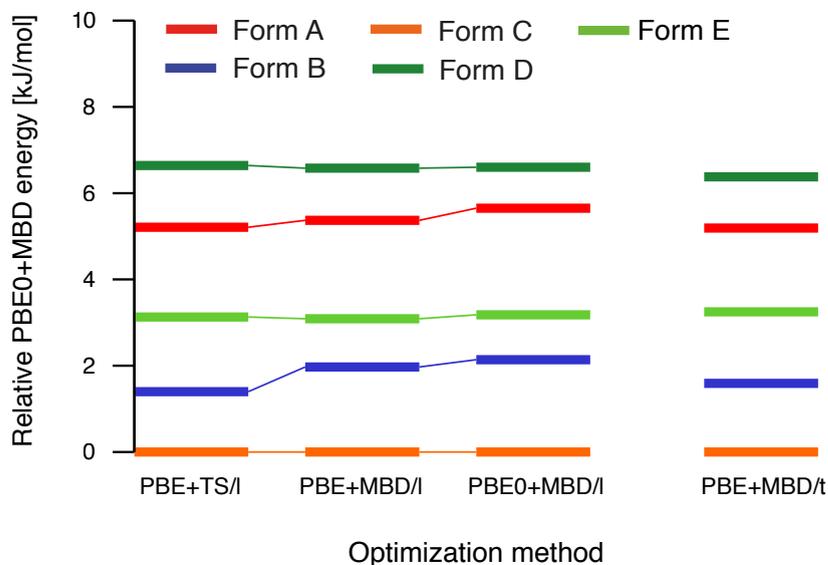


Figure 10.5: Relative PBE0+MBD energies of the five experimental polymorphs of system XXIII for different optimization methods. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

volume are now much smaller. PBE+MBD consistently yields larger unit cell volumes than PBE+TS. This is due to the fact that the pairwise dispersion approach tends to overestimate intermolecular interactions [40]. In terms of the RMSD_{20} one can only see a minor improvement with the average and the maximum amounting to 0.24 and 0.45 Å, respectively. When we move on to PBE0+MBD, we again see a similar picture for cell lengths and angles. However, the exact exchange leads to significantly smaller unit cell volumes. Now, the unit cell volume of XXII is only slightly overestimated by 0.6 %. On average, unit cell volumes are underestimated by 3.7 %. Again, this underestimation originates due to the missing thermal expansion in our structures and has to be expected. In terms of the RMSD_{20} the PBE0+MBD structures yield very similar results to the previously discussed methods with an average of 0.23 Å and a maximum of 0.47 Å. Finally, let us discuss the PBE+MBD optimizations with tight settings. It can be seen that the cell lengths and angles are only slightly better compared to the light settings and also the unit cell volumes are not changing significantly. Also the RMSD_{20} values are again similar with an average of 0.23 Å and a maximum of 0.46 Å. Therefore, we can conclude that light settings provide already a sufficient description of the crystal structure.

Overall, all discussed methods provide quite similar structures, especially in terms of the cell lengths and angles, as well as the RMSD_{20} . The only major difference is the cell volume, which in general follows the following trend: $\text{PBE0+MBD} < \text{PBE+TS} < \text{PBE+MBD}$.

The main focus of all CSP approaches is the stability ranking or in other words the polymorphic energy landscape. Therefore, we discuss now the effect of the optimization method on the relative stabilities for the five experimentally obtained structures

of system XXIII. The final static lattice energies within our CSP procedure are always calculated on the PBE0+MBD level. Hence, Fig. 10.5 shows the relative PBE0+MBD energies calculated on top of the differently optimized structures. We again discuss the differences with the mean absolute deviation (MAD) and the maximum absolute deviation (MAX) evaluated based on all possible relative energies. Changing the lattice relaxations from PBE+TS (light) to PBE+MBD (light) yields a MAD of only 0.3 kJ/mol with a MAX of 0.6 kJ/mol. PBE0+MBD optimizations yield 0.4 and 0.8 kJ/mol, respectively. Utilizing tight settings for PBE+MBD optimizations instead of light settings yield a MAD of 0.3 kJ/mol with a MAX of 0.5 kJ/mol. This illustrates that the PBE0+MBD energy ranking is only mildly affected by improving the optimization method, with all changes in relative energies being less than 0.8 kJ/mol. In addition, Mortazavi *et al.* have shown that computationally very efficient density functional tight binding optimizations could be used for a pre-screening when the stability is subsequently evaluated with PBE+MBD on top of these structures [196].

10.4 Vibrational Free Energies

In order to describe relative stabilities at finite temperatures, it is necessary to include thermal effects. The simplest way to achieve this, is to add to the electronic energies (and the nuclear repulsion) the vibrational free energy. This leads then to Helmholtz free energies, which include entropic effects. The vibrational free energy $F_{\text{vib}}(T)$ can be calculated within the harmonic approximation according to Eq. 4.5. However, accurate calculations are not as straightforward as simple energy calculations. For accurate results it is imperative that there are no imaginary modes present at the Γ -point. Such modes indicate that this structure does not correspond to a local minimum of the potential energy surface and also drastically change the vibrational free energy. The relative free energies are essentially determined by the low-frequency modes, since they have the largest impact on the entropic contributions. This is illustrated in Fig. 10.6. The left part of the figure shows the low-frequency pDOS for three XXIII structures. The right part of the figure illustrates the corresponding cumulative vibrational free energy for low frequencies, i.e. only phonon modes up to the plotted wave number are considered. The vibrational free energy of structure N70 always defines the zero of the energy and the final relative vibrational free energies after taking into account all modes are indicated with the three lines outside of the right plot. It can be seen that the relative vibrational free energies are essentially determined by low-frequency modes up to about 200 wave numbers. These low-frequency modes contain intermolecular motions and therefore vdW interactions play an important role for them.

Furthermore, calculating the vibrational free energy for the Γ point of the unit cell is insufficient for accurate results, which we will illustrate using our set of 9 structures for system XXIII. Fig. 10.7 shows relative stabilities in terms of the Helmholtz free energies at 300 K, for which the vibrational free energy was calculated at the PBE+TS level. Since we are using finite displacements, we have to avoid artifacts between displacements in periodic images. It was shown that at least 9-10 Å are necessary in every direction in

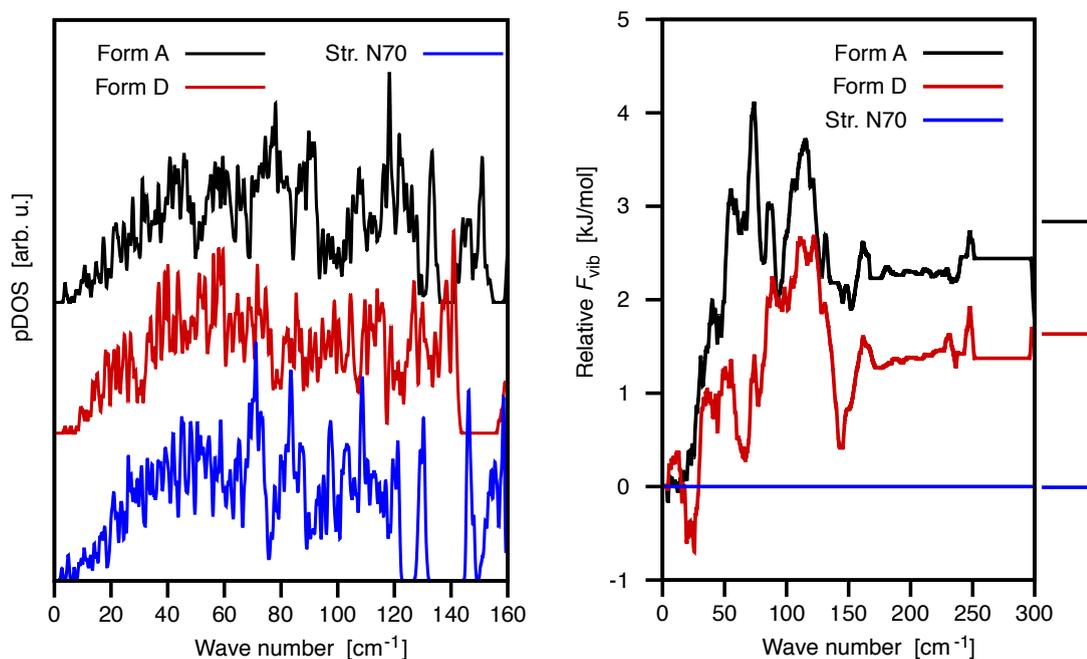


Figure 10.6: Low frequency part of the phonon density of states (pDOS) for three structures of system XXIII (left) and the corresponding cumulative relative vibrational free energies w.r.t the wave number (right). The vibrational free energy of structure N70 is always set to zero and the vibrational free energy is evaluated based on the pDOS up to the shown wave number. The right plot shows the low-frequency part up to 300 wave numbers. The final relative vibrational free energies after taking into account all modes are indicated with the three lines outside of the right plot. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

order to minimize artifacts [40]. Therefore, we have used for all phonon calculations in our CSP procedure [185] super cells, which extend at least 10 \AA in every direction. All values of part (a) in Fig. 10.7 were calculated with these cells but the vibrational free energy was evaluated differently. At first, the vibrational free energy was evaluated by only using the modes at the Γ -point of the unit cell (uc), followed by the Γ point of the used supercell (sc). In all other cases the vibrational free energy was evaluated with a q -grid in reciprocal space utilizing the q -point interpolation available in phonopy [123]. The number (n) of necessary q -points in each direction was calculated according to $n \times a \geq x$, with a being the cell length in the respective direction and the used values of x are given in the plot. It can be seen that both Γ -point plots provide a completely different stability ranking compared to the converged reciprocal space sampling. This is mainly due to the fact that the three acoustic modes have zero frequency at the Γ point but contribute significantly for other q -points. Note that for forms A, D, and structure N70 the supercell is identical to the unit cell, since in these unit cells all cell lengths are already larger than 10 \AA . Convergence to 0.1 kJ/mol is reached at $x = 50 \text{ \AA}$.

Part (b) of Fig. 10.7 illustrates the importance of the used supercell for the finite difference calculations. Here, the vibrational free energy is always evaluated with a

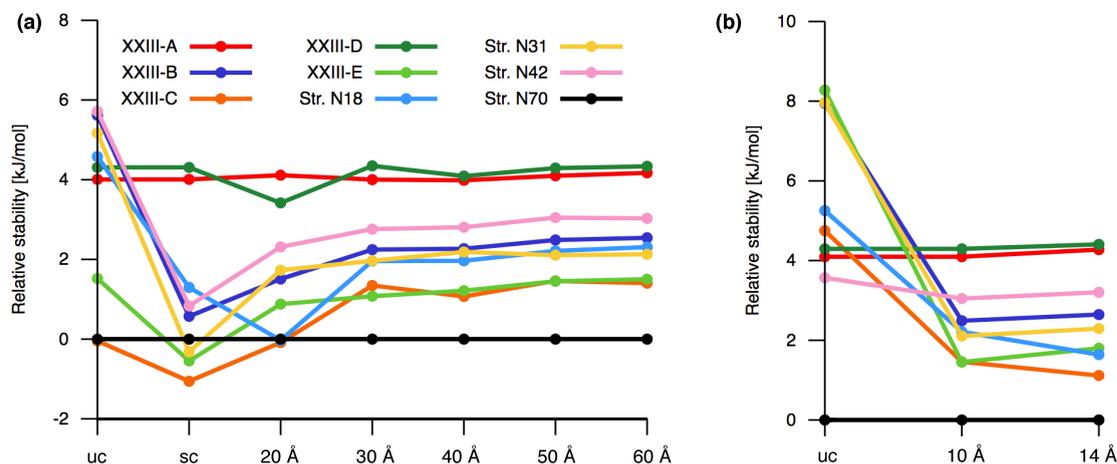


Figure 10.7: Convergence of vibrational free energies. Part (a) shows the relative stabilities of 9 structures from system XXIII calculated with PBE0+MBD+ F_{vib} (PBE+TS) at 300 K, where supercells with cell lengths of at least 10 Å were used for the finite difference calculation. The vibrational free energy was evaluated using only the Γ point modes of the unit cell (uc), the Γ -point modes of the used supercell (sc), or with a q -point mesh with the shown values used for x (see Methods). In part (b) a converged q -point mesh is used throughout but the finite difference calculations were performed by either using just the unit cell or supercells with minimum cell lengths of 10 and 14 Å, respectively. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

q -point grid corresponding to $x = 50$ Å but we used different cell sizes for the finite difference calculations. First, the calculations were performed by only using the unit cells (uc), followed by the smallest supercells which extend in every direction at least 10 Å and 14 Å, respectively. The MAD when comparing all relative energies between the unit cell and the 10 Å supercell approach amounts to 3.3 kJ/mol with a maximum of 6.8 kJ/mol. Moving from the 10 Å to the 14 Å supercells leads to a MAD of only 0.3 kJ/mol with a maximum energy change of 0.9 kJ/mol. This illustrates how crucial these supercells are for accurate vibrational free energies. In addition using larger supercells than 10 Å has only a minor effect on the relative stabilities and does not merit the immense increase in computation time. In general, the larger the supercell, the more points in reciprocal space are evaluated exactly, which increases the quality of the q -point interpolation used in phonopy. For most of the unit cell calculations one can see in the phonon band structure that the three acoustic modes are poorly described away from Γ . This is illustrated for structure N18 in Fig. 10.8. For the corresponding 10 Å supercell we find that one acoustic mode is still not properly resolved leading to imaginary frequencies away from the Γ point, while for the 14 Å supercell all issues with imaginary modes are resolved.

Now let us discuss the impact of the vibrational free energies on stability rankings (see Table 10.1). The average impact in terms of the MAD varied for the five systems

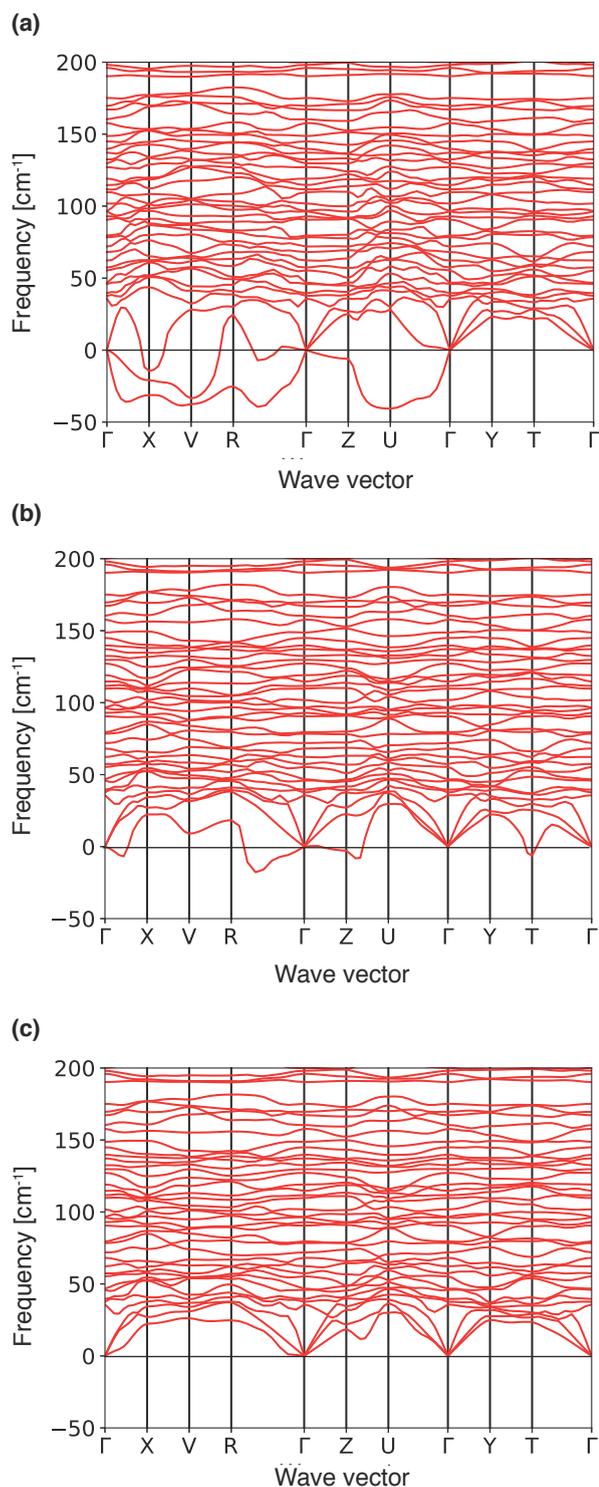


Figure 10.8: Low-frequency phonon band structure plot along several high-symmetry lines for structure XXIII-N18 calculated using the unit cell (a), a supercell with a minimum length of 10 Å (b), and a supercell with a minimum length of 14 Å (c) in every direction. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

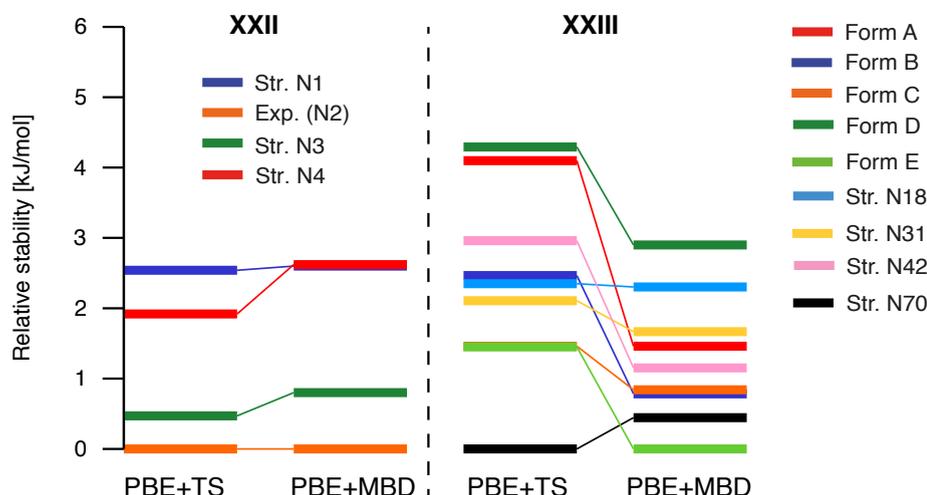


Figure 10.9: Comparison of PBE+TS and PBE+MBD vibrational free energies for several structures of systems XXII and XXIII. The relative stability consists of the PBE0+MBD energy plus the vibrational free energy evaluated with the respective method. The vibrational free energies were calculated at the temperature of the respective experimental crystal structure measurements (150 K for XXII and 300 K for XXIII). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

from 0.5 to 1.9 kJ/mol. While this effect is not that pronounced for systems involving rigid molecules, it can be crucial for the stability ranking of systems involving flexible molecules. For our highly polymorphic system XXIII vibrational free energies can change relative stabilities up to 8.4 kJ/mol.

All so far discussed vibrational free energies have been calculated with PBE+TS. Therefore, we now evaluate what impact MBD interactions and exact exchange have on the vibrational free energies. Let us begin by discussing the effect of MBD based on the four most stable structures of system XXII and 9 structures of system XXIII (see Fig. 10.9). The total energy is always evaluated with PBE0+MBD while the vibrational free energy is calculated by using PBE+TS and PBE+MBD, respectively. For system XXII we observe only minor changes in the relative stabilities; relative energies change on average by 0.4 kJ/mol and the maximum change amounts to 0.7 kJ/mol. In contrast, the effect of MBD interaction on the vibrational free energies is much stronger for system XXIII with an average change of 1.2 kJ/mol and a maximum modification of 3.1 kJ/mol. This is consistent with Ref. 42, which shows that MBD interactions change the relative free energy between two polymorphs of aspirin by about 3 kJ/mol. In general, MBD effects are more pronounced for cells with little symmetry involving especially flexible molecules.

Furthermore, let us discuss the influence of exact exchange on vibrational free energies. Therefore, two structures from system XXII, for which exact exchange significantly modified the relative stability, were selected as well as 3 structures from system XXIII, for which no supercell is necessary for the accurate evaluation of the vibrational free

Table 10.3: Relative free energies in kJ/mol. The total energy was always calculated with PBE0+MBD and the structure optimization as well as the vibrational free energy was performed with the labeled method. The vibrational free energies were always calculated at the experimental measuring temperature of the respective crystal structures (150 K for XXII and 300 K for XXIII). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

Method	ΔXXII^a	$\Delta\text{XXIII-1}^b$	$\Delta\text{XXIII-2}^c$	$\Delta\text{XXIII-3}^d$
PBE+TS/l	2.5	0.2	4.3	4.1
PBE+MBD/l	2.6	1.5	2.4	0.9
PBE0+MBD/l	2.0	-0.8	0.3	1.1
PBE+MBD/t	2.2	1.0	2.4	1.3

$$^a \Delta\text{XXII} = F(\text{XXII} - \text{N1}) - F(\text{XXII} - \text{N2}), \quad ^b \Delta\text{XXIII-1} = F(\text{XXIII} - \text{D}) - F(\text{XXIII} - \text{A}), \\ ^c \Delta\text{XXIII-2} = F(\text{XXIII} - \text{D}) - F(\text{XXIII} - \text{N70}), \quad ^d \Delta\text{XXIII-3} = F(\text{XXIII} - \text{A}) - F(\text{XXIII} - \text{N70}).$$

energy. This benchmark set provides 4 relative free energies (see Table 10.3). The structure optimization and the calculation of the vibrational free energy is done with the respectively mentioned method, while the total energy is evaluated with PBE0+MBD. In addition, the PBE+MBD calculations were also performed by using the tight settings in order to evaluate if a larger basis set would significantly change the result. The difference between PBE0+MBD and PBE+MBD (light) amounts to 1.3 kJ/mol in terms of the MAD and the maximal observed change is 2.3 kJ/mol. The MAD between PBE+TS and PBE0+MBD amounts to 2.1 kJ/mol with a maximum of 3.9 kJ/mol. Hence, exact exchange still modifies relative vibrational free energies, however, this effect is not large enough to justify the massive additional computational cost for production calculations. In addition, increasing the basis set yields only a minor modification of the relative stabilities with a maximum change of 0.5 kJ/mol.

10.5 Computational Cost

The computational cost of single-point energy calculations, lattice optimizations, and vibrational free energies highly depends on the size and shape of the involved unit cells. Here we discuss the computation times for form A of system XXIII. This monoclinic unit cell contains 172 atoms, which is only slightly less than the average unit cell size studied in Ref. 185. Table 10.4 shows the relative computation times for one single-point energy and one force calculation for all discussed methods. The values are normalized to the PBE+TS single-point energy calculation with light settings (l), which needs 2.1 CPU hours on Intel Xeon E5-2680 v4 cores (2.4 GHz). It can be seen that the calculation of the many-body dispersion energy (MBD) adds only a negligible cost on top of the PBE+TS calculation. However, PBE+MBD forces need about 50 % more computation time than PBE+TS forces when computed at the light level. PBE0 calculations are at the light level about 17 times more expensive than PBE calculations. Furthermore, utilizing tight settings (t) increases the computation time of a PBE+MBD energy calculation by a factor of about 10 compared to light settings. When utilizing tight settings, PBE+TS

Table 10.4: Relative computation times for one single energy and force calculation of structure XXIII-A (172 atoms in the unit cell). All computation times were normalized to the PBE+TS/l energy calculation, which amounts to 2.1 CPU hours on Intel Xeon E5-2680 v4 cores (2.4 GHz). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

Method	Energy	Force
PBE+TS/l	1.0	1.4
PBE+MBD/l	1.0	2.1
PBE0+MBD/l	16.9	18.2
PBE+MBD/t	9.8	13.2

and PBE+MBD force calculations are essentially equally expensive since almost all computation time is spent on the underlying DFT calculation.

The number of needed optimization steps in a geometry and lattice optimization highly depends on the starting structure, i.e. how similar the initial molecular conformation and the unit cell shape are to the optimized structure. Typically, 50 - 100 steps are necessary to converge a structure with used settings. Therefore, the computation time amounts to about 50 - 100 times the time needed for one force calculation. The calculation of the harmonic vibrational free energy requires for XXIII-A 258 finite displacements and therefore also 258 force calculations. In this particular case the unit cell is sufficient for the finite differences calculation and no supercell is needed. Therefore, the calculation of the vibrational free energy of XXIII-A using PBE+TS with light settings requires about 750 CPU hours.

10.6 Beyond the Harmonic Approximation

So far we have calculated all free energies on top of the fully optimized structures. However, molecular crystals expand with increasing temperature. This structural change could lead to a modification of the relative stabilities. The simplest approach of including the effect of thermal expansion on a first-principles level is the QHA [13, 143, 147, 148, 197–199]. In this approach the vibrational free energy is calculated within the harmonic approximation for several unit cell volumes, and the volume at a certain temperature is determined by an equation-of-state (EOS) fit. We have calculated the cells corresponding to 300 K within the QHA for 9 structures from system XXIII [185]. Note that in this case we use PBE+TS for structures, vibrational free energies, as well as for the total energies. This approach yields unit cell volumes which agree on average within 1 % with the experimental structures (see Appendix B). The largest error was found for form E, where the unit cell volume is underestimated by 1.7 %.

The thermal expansion for form A is illustrated in Fig. 10.10 (a) at the PBE+TS level. The dots correspond to the harmonic free energies calculated at various volumes and at three different temperatures. The solid curves correspond to the Murnaghan equation-of-state fit, and the respective minimum is always marked with the red triangle.

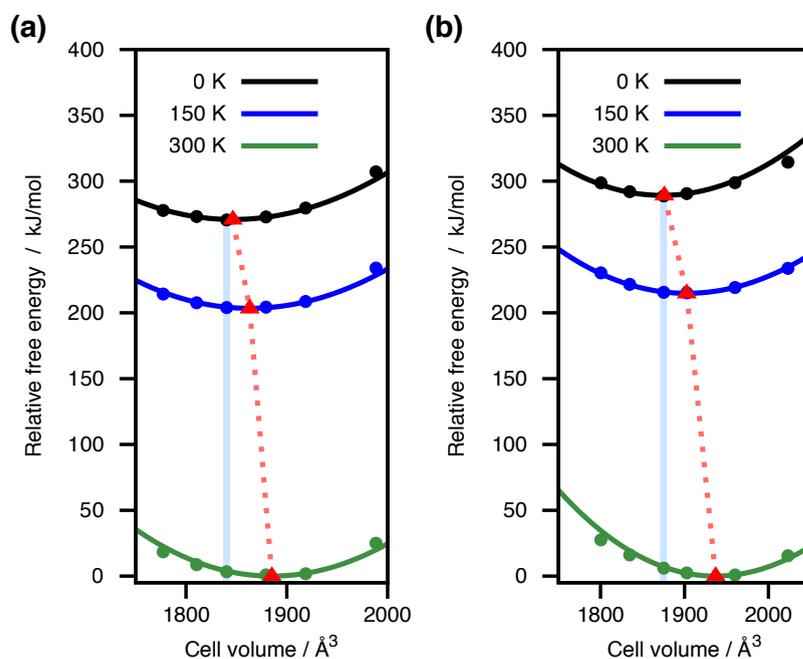


Figure 10.10: Illustration of the QHA thermal expansion for form A of system XXIII as described by PBE+TS (a) and PBE+MBD (b). The dots correspond to the harmonic free energies calculated at various volumes and at three different temperatures. The solid curves correspond to the Murnaghan equation-of-state fit, and the respective minimum is always marked with the red triangle. The light blue line indicates the volume of the fully optimized structure. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

In addition, the light blue line indicates the volume of the fully optimized structure. Already the inclusion of zero-point energy (0 K) leads to a volume increase of 0.4 %. At 150 and 300 K the thermal volumetric expansion in comparison with the optimized structure amounts to 1.3 and 2.5 %, respectively. The resulting volume at 300 K agrees within 0.7 % with the experimentally measured volume at that temperature. The QHA is sometimes performed by isotropically scaling the cell [58]. Since we perform lattice relaxations using a thermal pressures, we get an estimate for the directional thermal expansion. In this example the thermal expansion is quite different in all directions and amounts for the cell lengths a , b , and c to 0.8 %, 1.4 %, and 0.2 %, respectively. In addition, we have also performed the QHA for form A utilizing exclusively PBE+MBD (see Fig. 10.10 (b)). In this case the thermal expansion amounts at 300 K to 3.3 %, which overestimates the experimental cell volume by 1.9 %. However, when we pair the PBE+MBD phonons with PBE0+MBD total energies, the resulting volume agrees in this case within 0.2 % with the experimental volume. PBE+MBD prefers in general larger volumes than PBE+TS which leads here to an overestimation of the cell volume at 300 K. In contrast, the effect of exact exchange significantly reduces the cell volume. Overall, the QHA entirely done at the PBE+TS level provides a good description of the unit cell volume at room temperature.

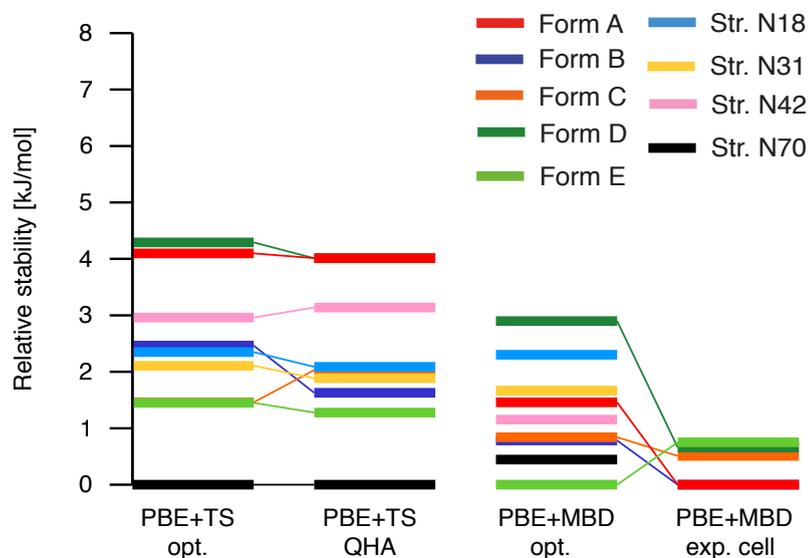


Figure 10.11: Relative stabilities for 9 structures of system XXIII for optimized and thermally expanded structures. The relative stability consists of the PBE0+MBD total energy plus the vibrational free energy evaluated with the respective method. The vibrational free energy was calculated in the harmonic approximation at 300 K. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

Next, let us focus on how the QHA affects the relative stabilities. Therefore, the left part of Fig. 10.11 compares the relative stabilities of the optimized XXIII structures with the thermally expanded ones. The total energies are again calculated with PBE0+MBD and the vibrational free energies with PBE+TS. It can be seen that the QHA only mildly modifies the relative stabilities, with an average change amounting to 0.4 kJ/mol and a maximum of 1.4 kJ/mol. At this level we find all so far experimentally observed structures within an energy window of only 2.6 kJ/mol, which in general corresponds quite well with the expected energy window of co-existing polymorphs. In comparison, Brandenburg and Grimme [58] also studies the experimentally confirmed polymorphs of system XXIII with the QHA utilizing a minimal basis set Hartree Fock method [200] (HF-3c) as well as density functional tight binding [201] (DFTB) calculations for vibrational free energies, and TPSS-D3 [32, 202] for total energies. In their case the spread between the experimental structures is much larger, amounting to 8 kJ/mol. Experimental evidence [51] suggests that form D should be the most stable one at room temperature. However, in our ranking form E is the most stable one and form D is in fact the second least-stable one. In addition, we find one structure (N70), which is described as more stable than any experimental structure.

Harmonic PBE+MBD vibrational free energies at the optimized structures describe form D as the least stable and form E as the most stable form, with a spread of 3 kJ/mol. Therefore, we investigated if we could reproduce with PBE+MBD phonons the expected stability ranking, assuming that we have the knowledge of the exact unit cell. These calculations were performed using the experimental cell vectors and the results

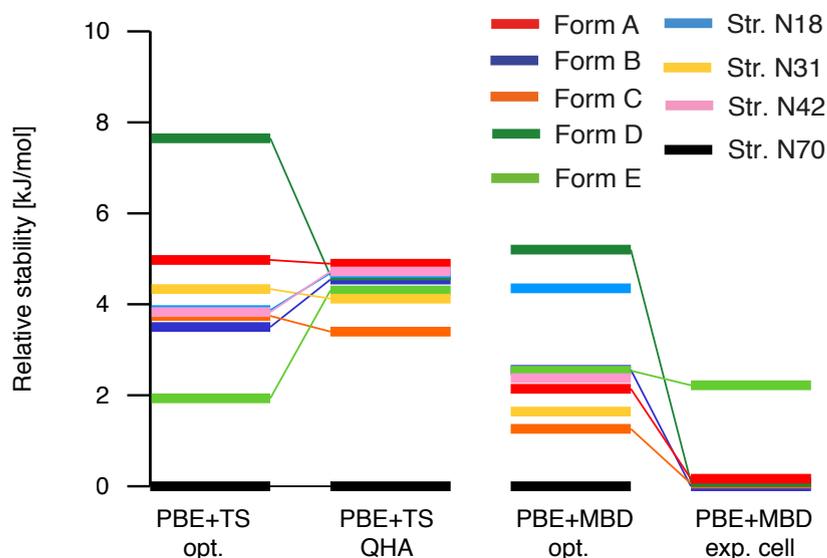


Figure 10.12: Relative stabilities for 9 structures of system XXIII for optimized and thermally expanded structures. The relative stability consists of the PBE0+MBD total energy plus the vibrational free energy evaluated with the respective method. The vibrational free energy was calculated using Morse oscillators at 300 K. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.

are shown in the right part of Fig. 10.11. In this case all experimental structures differ by less than 0.8 kJ/mol, which means that given our expected accuracy all structures are in fact degenerate.

All so far discussed vibrational free energies have been determined within the harmonic approximation. However, in reality these modes are not entirely harmonic. Therefore, we have estimated anharmonic vibrational free energies by replacing the harmonic oscillators by Morse oscillators [104, 105]. These anharmonic free energies were calculated for all optimized and thermally expanded structures with PBE+TS and PBE+MBD, respectively. The resulting stability rankings are shown in Fig. 10.12.

It can be seen that in case of optimized structures, form D is always described as the least stable structure and that the spread of the relative energies is larger than for harmonic vibrational free energies. Let us now focus on the PBE+TS QHA results. Here, form D is highly stabilized by the thermal expansion and all experimental structures are getting closer in terms of energies. Now, all experimental structures are found within an energy window of only 1.5 kJ/mol. Form C is described as the most stable experimental structure. However, we underestimate the unit cell volume of form C much more than for form D. Since a further increase in the unit cell volume would lead to higher energies, we can expect further destabilization of form C if we would have even more accurate unit cells. Interestingly, form N70 is still described as the most stable structure (see Chapter 9 for a discussion). Overall, the result of the QHA with PBE+TS anharmonic free energies already agrees well with the experimental expectation in terms of the stability ranking. Given our expected accuracy, all experimentally confirmed polymorphs are

essentially degenerate. The difference between this stability ranking (performed on top of thermally expanded structures utilizing Morse oscillators) and harmonic free energies evaluated at optimized structures amounts on average to 1.1 kJ/mol and the maximal observed change is 2.9 kJ/mol.

Finally, let us discuss the anharmonic free energies calculated with the experimental cell vectors with PBE+MBD. In this case all experimental structures except form E are found within an energy window of only about 0.2 kJ/mol. This picture agrees very well with the available experimental observations. However, form E is described as 2.2 kJ/mol less stable than the other polymorphs. This large difference occurs probably because the optimized PBE+MBD structure has a quite large RMSD w.r.t. the experimental structure (0.45 Å), most likely due to potential disorder in the experimental form E. We note that this approach for calculating anharmonic free energies does not take into account the interactions between phonon modes. More accurate anharmonic free energies could be obtained by using for example the vibrational self-consistent field approach [203–205] or path-integral molecular dynamics [206].

10.7 Summary

We have presented in this chapter an in-depth discussion how crucial various commonly neglected effects are for the calculation of relative stabilities of molecular crystal polymorphs. Many-body dispersion effects, exact exchange, and vibrational free energies all contribute significantly to the obtained relative stabilities. Therefore, none of the studied effects can safely be neglected, especially for systems involving flexible molecules. We have shown that our hierarchical CSP approach presented in Chapter 9 [185], which yields excellent results for the systems of the latest blind test, is already quite robust. Increasing the basis set for structures and vibrational free energies does not significantly change the resulting stability ranking. In addition, the evaluation of vibrational free energies at the PBE+TS level instead of PBE0+MBD results in an average absolute error of about 2 kJ/mol. Many-body dispersion interactions affect the vibrational free energies on average by about 1 kJ/mol but can change relative stabilities in system XXIII up to 3 kJ/mol. In the case of aspirin polymorphs, these interactions are crucial to obtain the correct stability [42]. Given the recent development and efficient implementation of analytical nuclear gradients for MBD [89, 207], geometry optimizations and vibrational free energies could be performed with PBE+MBD instead of PBE+TS in case of complicated polymorphic energy landscapes. This would increase the computation time for forces on average by about 50 %.

Furthermore, we have studied the effect of thermal expansion within the QHA. The resulting thermally expanded structures obtained at the PBE+TS level provide accurate unit cell volumes. However, the relative stabilities remain very similar to the ones of the fully optimized structures. In addition, we also estimated anharmonic free energies by using Morse oscillators. These results are for system XXIII in good agreement with the available experimental information on relative stabilities. Including the QHA as well as the anharmonic free energy estimates change relative stabilities on average by about 1

kJ/mol. Therefore, the harmonic approximation on top of fully optimized structures will often provide sufficient results in case of already well-separated energy landscapes.

Chapter 11

Relative Stabilities of the Five Coumarin Polymorphs

After this discussion of the five blind-test systems, we apply the DFT+MBD framework now to another highly polymorphic system — coumarin. The availability of experimental relative stabilities provides us with the opportunity to benchmark our calculated stabilities now also on a quantitative level. The molecular structure of a coumarin molecule is visualized in Fig. 11.1. This rather simple organic compound is for example used in perfumes and as a precursor for the synthesis of drugs. In 1929, two forms of coumarin were found by Bernauer [208] via crystallization from the melt. In 1934, Kofler and Geyr [209] identified one monoclinic and one orthorhombic form of coumarin. In 1939, Lindpaintner [210] obtained three polymorphs with different melting points.

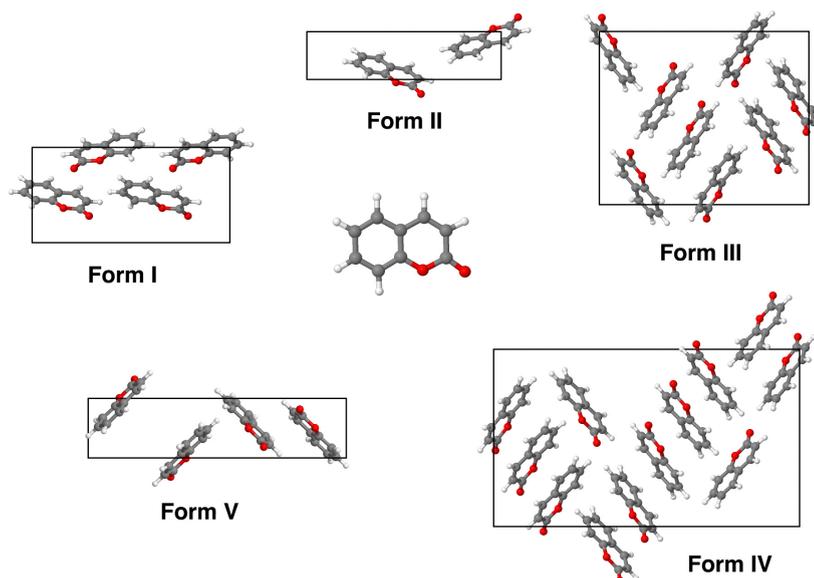


Figure 11.1: Illustration of a coumarin molecule and the unit cells of the five coumarin polymorphs.

While coumarin was clearly studied for quite some time, only one solid form (form I) was solved [211–213]. In Ref. 60 the crystal structures of four additional coumarin polymorphs are obtained via X-ray powder diffraction in synergy with computational CSP. The resulting five coumarin polymorphs are illustrated in Fig. 11.1. In this chapter, we show the polymorphic energy landscape obtained with PBE+TS and PBE+MBD using 50 low-energy structures from the two utilized CSP approaches [60]. Furthermore, we discuss the relative stabilities of the five coumarin polymorphs obtained with our DFT+MBD framework. The results presented in this chapter have been published in *Chemical Science* [60]. The author of this thesis contributed to that publication all DFT+TS and DFT+MBD calculations within FHI-aims, which are described below.

11.1 Computational Methods

In Ref. 60 two different CSP procedures were employed. In the first case, the crystal structure search was carried out by using an evolutionary algorithm within the USPEX code [214–217] and the structure relaxations were performed within DMACRYS [180]. The employed distributed multipole model was derived from MP2 charge densities and the FIT [218] empirical potential was used to describe exchange repulsion and dispersion interactions. The second CSP approach utilized a random structure search via the UPACK software [219] based on a rigid molecular structure determined with PBE0. The corresponding energies were evaluated with a modified OPLS force field [220]. The results of both CSP approaches were combined to create a set of 50 low-energy structures, which were subsequently relaxed with the vdW-DF2 functional [221]. This structure set served as starting point for our own calculations.

First, we calculated the PBE+TS [34, 82] and PBE+MBD [35, 64] energies on top of the vdW-DF2-optimized structures. All electronic structure calculations were carried out using the all-electron code FHI-aims [74]. Tight species default settings were used for integration grids and basis functions and the k-point mesh satisfied in every direction $a \times n \geq 30 \text{ \AA}$, with a being the cell length in a particular direction and n being the corresponding number of k-points.

Next, we studied the five experimentally observed polymorphs in more detail. Therefore, full lattice relaxations were performed using PBE+MBD with light species default settings in FHI-aims, converging force components to 10^{-3} eV/\AA . Subsequently, the energies of the five polymorphs were calculated with PBE+MBD using tight species default settings and with the PBE0+MBD approach utilized in the previous two chapters. Therein, the PBE0+MBD energy is calculated by adding the energy difference between PBE0+MBD/light and PBE+MBD/light to the obtained PBE+MBD/tight energies (see Chapter 9). Furthermore, the vibrational free energies were calculated for all five structures in the harmonic approximation using PBE+MBD with light species default settings. Finite displacements of 0.005 \AA were used and the vibrational free energies were evaluated using phonopy [123]. Supercells with at least 10 \AA in every direction were used for the finite-displacements calculations in order to minimize artifacts. In

addition, all these calculations were also carried out for forms I and V using the experimental cell vectors at 300 K and on top of the optimized structures utilizing tight species default settings.

11.2 Polymorphic Energy Landscapes

First, we discuss the polymorphic energy landscapes formed by the 50 low-energy structures obtained from the employed CSP procedures. Fig. 11.2 shows the obtained PBE+TS and PBE+MBD landscapes. In the PBE+TS case we find all 50 structures within an energy interval of 10 kJ/mol with form I being the most stable structures. All five experimentally observed polymorphs are located within the top 2.3 kJ/mol. This energy interval contains also 17 other structures, from which 5 can be considered polytypes of form I and II. These results are quite similar to the preceding vdW-DF2 energies [60], for which dispersion interactions are also included in a pairwise fashion. In this case, all experimentally confirmed polymorphs are found within 2.5 kJ/mol and this energy interval contains 15 other structures (see Ref. 60 for more details).

In the PBE+MBD case all 50 structures are within 7 kJ/mol and all five experimentally obtained polymorphs are located within only 1.8 kJ/mol. These top 1.8 kJ/mol contain except forms I-V only 4 other structures, where three of them can be considered polytypes of form I. Hence, the PBE+MBD energies show a much better separation between observed and non-observed structures than the pairwise approaches. This further highlights the importance of MBD interactions for polymorph stabilities. However,

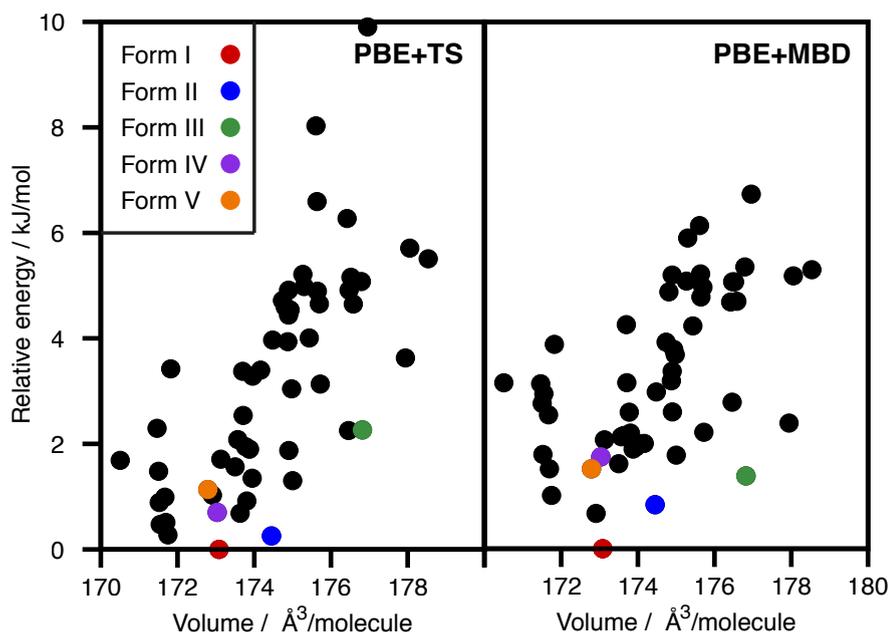


Figure 11.2: PBE+TS and PBE+MBD stability vs. volume for 50 vdW-DF2-optimized structures of coumarin.

note that thermodynamic stability rankings can always include non-observed low-energy structures which are kinetically instable or have not been crystallized yet.

11.3 Relative Stabilities of Forms I–V

Now let us discuss in detail the five experimentally observed polymorphs. Experimental evidence [60] suggests that the relative stability ordering at and above room temperature should be I>II>III>IV>V. This is based on observed transition between the forms and melting point differences with and without a Canada balsam additive. Form I was up to now the only characterized form and there is no evidence of a low-temperature phase transition, suggesting that form I is very likely the most stable polymorph. Due to the structural similarity of form III and IV they are expected to be energetically very similar. Form V is only stable for a short period of time under ambient conditions, suggesting that it is the least stable form. The experimentally determined relative Gibbs free energies around the melting points are listed in Tab. 11.1. It can be seen that all five polymorph lie in a very narrow energy interval of 0.8 kJ/mol. For more information the reader is referred to Ref. 60.

The calculated PBE0+MBD relative energies on top of the PBE+MBD-optimized structures are shown in Tab. 11.1. It can be seen that we observe the same stability ordering as in experiment, but the relative stabilities differ by up to 1.5 kJ/mol from the experimental values. In Ref. 60 a variety of different vdW-inclusive DFT methods were studied, but only methods including MBD interactions yielded this qualitative ordering.

However, since we have established during the previous chapters that vibrational free energies can be of significant importance for the relative stabilities of polymorphs, we have added vibrational free energies calculated at the PBE+MBD level to the static PBE0+MBD results. The resulting Helmholtz free energies are plotted up to 350 K in Fig. 11.3 and listed in Tab. 11.1 for 300 K and for the average melting point (340 K). It can be seen that the relative stabilities vary significantly with temperature. Especially, form V is significantly stabilized by temperature effects w.r.t. form I. At 300

Table 11.1: Relative stabilities in kJ/mol calculated using the PBE+MBD/light-optimized structures with different methods compared to the experimental estimates around 340 K (Exp.) [60]. The Helmholtz free energies F at 300 and 340 K are the sum of the PBE0+MBD static energy and the vibrational free energy calculated with PBE+MBD/light.

Polymorph	$E(\text{PBE+MBD})$	$E(\text{PBE0+MBD})$	$F(300 \text{ K})$	$F(340 \text{ K})$	Exp.
Form I	0.0	0.0	0.0	0.0	0.0
Form II	0.2	0.3	0.7	0.8	0.2
Form III	0.7	1.2	0.9	0.9	0.2
Form IV	1.1	1.8	0.9	0.8	0.3
Form V	1.1	2.2	0.2	0.0	0.8

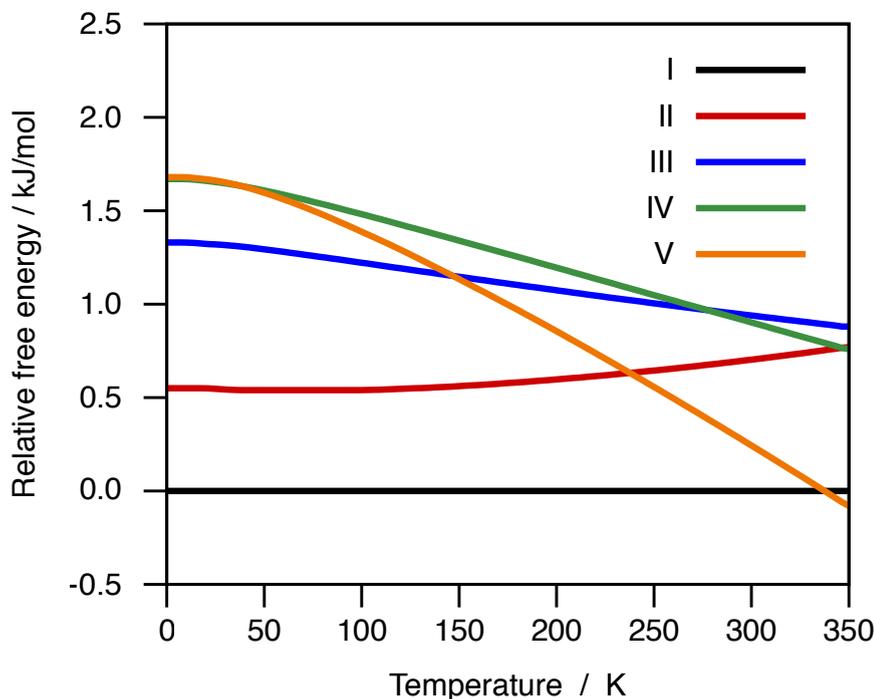


Figure 11.3: Relative Helmholtz free energies vs. temperature for the five coumarin polymorphs. The shown free energies are the sum of the PBE0+MBD static energy and the vibrational free energy calculated with PBE+MBD/light. The energy of form I defines at every temperature the zero of the energy [60].

K we observe the following stability ordering: $I > V > II > IV \approx III$. Now, form I remains the most stable polymorph and forms III and IV have as expected essentially the same stability. In contrast to experiment, form V is now described as the second most stable structure. To rationalize this stabilization, the low-frequency phonon density of states (pDOS) is plotted in Fig. 11.4 for forms I and V. The main difference in the free energy originates from the fact that form V has a larger pDOS at very low frequencies ($< 50 \text{ cm}^{-1}$) than form I. In order to determine if this stabilization of form V could be related to the light settings utilized in the phonon calculations, the vibrational free energy was also calculated using tight settings. The resulting qualitative trend between form I and V remains the same, yielding at 300 K an energy difference of only 0.3 kJ/mol in favor of form I. Furthermore, we also calculated the Helmholtz free energies of forms I and V using the experimentally measured cell vectors at 300 K. Also in this case, form V is stabilized by temperature effects and the energy difference to form I amounts at 300 K again to 0.3 kJ/mol. Therefore, thermal expansion has only a negligible effect on these relative stabilities. In Ref. 60 also full molecular dynamic simulations were performed with an OPLS force field, which revealed that form V is in fact destabilized w.r.t. form I due to anharmonic effects.

Finally, let us compare our calculated relative stabilities to the experimental Gibbs free energies. We utilize for this comparison the Helmholtz free energies obtained at the

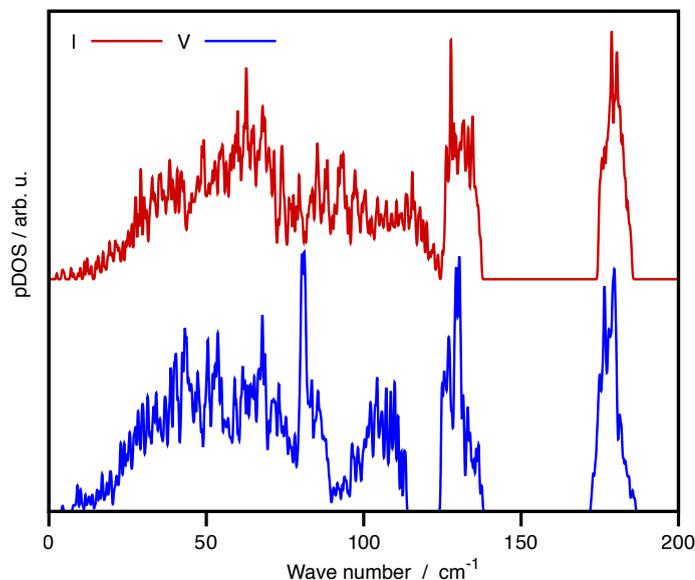


Figure 11.4: Low-frequency phonon density of states for coumarin forms I and V according to PBE+MBD.

average melting point of 340 K (see Tab. 11.1). Note that our obtained Helmholtz free energies can also be interpreted as Gibbs free energies since the additional pV term has only a negligible impact on the relative stabilities at ambient conditions. It can be seen that all five polymorphs are found within 0.9 kJ/mol. Although we observe a slightly different stability ordering, the largest error w.r.t. the experimental relative stabilities amounts to only 0.8 kJ/mol.

11.4 Summary

In this chapter we have studied the relative stability of coumarin polymorphs using the DFT+MBD framework. Many-body dispersion interactions lead in this case to a better separation between observed and non-observed structures in the polymorphic energy landscape than pairwise vdW approaches. The calculated relative free energies consisting of PBE0+MBD static energies and PBE+MBD vibrational free energies agree within 1 kJ/mol with the experimentally obtained relative Gibbs free energies and satisfy most observations in terms of the stability ordering of the polymorphs.

Chapter 12

Conclusion

A lot of progress has been made during the last few years in the modeling of molecular crystals from first principles. Nowadays, such methods can be applied to practically relevant molecular crystals, which leads to new insights and understanding in terms of structure and stabilities, as well as response properties like phonons, vibrational spectra, and elastic moduli. The accuracy and transferability of first-principles approaches will make them a central tool for the prediction and the engineering of molecular crystals in the future. However, one very important aspect is the understanding that many properties of molecular crystals — including their structure — can highly depend on the temperature. Therefore, accounting for thermal effects is important but significantly increases the computational cost. Hence, first-principles approaches are often limited to the harmonic approximation, which enables powerful insights but neglects any anharmonic effects due to the thermal expansion of the unit cell and internal motion.

The aim of this thesis was to accurately calculate relative stabilities of molecular-crystal polymorphs and related properties at finite temperatures on a first-principles level, which is still applicable to larger practically relevant molecular crystals. It was shown by Reilly et al. [40] that accurate lattice energies can be obtained for a set of small molecular crystals by using vdW-inclusive DFT. In the case of the PBE0 functional paired with the MBD vdW model, the MAE w.r.t. back-corrected experimental sublimation enthalpies amounts to less than 1 kcal/mol. Therefore, the MBD and the related pairwise TS vdW models were used throughout this thesis.

First, we have discussed state-of-the-art modeling approaches on the example of a simple cubic ammonia crystal. We have seen that by just considering fully optimized structures, the experimental unit-cell volume at a temperature of 180 K is overestimated by PBE and underestimated by both vdW-inclusive approaches. A very efficient way to account for the thermal expansion of a molecular crystal is the QHA. Within this approximation, PBE+MBD is able to capture the majority of the thermal expansion, leading to an excellent description of the unit-cell volume. In contrast, utilizing pure PBE within the QHA results in a significant overestimation of the unit cell volume, illustrating once again the crucial nature of vdW interactions for molecular-crystal properties. Furthermore, we have shown that accounting for the thermal expansion is

also crucial for the accurate calculation of elastic properties, such as the bulk modulus at finite temperatures.

Next, we have discussed the modeling of low-frequency vibrational spectra for an orthorhombic purine crystal. Also in this non-cubic case, the QHA provides an excellent description of the unit-cell volume at room temperature. In Chapter 6, we have shown that the QHA also significantly improves the description of vibrational frequencies in the THz regime when compared to experimental measurements at room temperature. However, we have so far neglected any additional anharmonic effects. Therefore, we have estimated further anharmonicities by replacing the harmonic oscillators with Morse oscillators. For the considered purine crystal this approach leads to a further improvement in terms of vibrational frequencies compared to the QHA.

The second part of this thesis dealt with the complex topic of organic CSP, which is imperative for the development of new pharmaceuticals and crystal engineering, since a variety of molecular-crystal properties are highly structure dependent. While first-principles calculations cannot routinely be applied for a complete CSP procedure, accurate calculations are essential for the final stability ranking. Therefore, we have developed a hierarchical CSP protocol intended for such a final stability ranking. Chapter 8 describes our contribution to the latest molecular CSP blind test organized by the Cambridge Crystallographic Data Centre. Therein, we have hierarchically refined and re-ranked an existing set of initial structures obtained via a quasi-random search. Our final stabilities were calculated using PBE+MBD and in some cases also vibrational free energies were considered. Excellent stability rankings were obtained for experimental structures available in the initial structure set. In addition, we observed that including vibrational free energies indeed improve the stability ranks.

Next, we have extended our CSP procedure to also include an enhanced description of the Pauli exchange repulsion via utilization of a hybrid density functional. To further validate our procedure, we have consistently applied it to the five blind-test systems utilizing a significantly different set of initial structures. In Chapter 9 the used initial structures were generated by a Monte Carlo parallel tempering algorithm and include for two systems also challenging $Z' = 2$ structures. Our most reliable level of theory in this recommended hierarchical computational approach (PBE0+MBD+ F_{vib}) provides excellent stability rankings for this very diverse set of molecular crystals. Even for the challenging polymorphic system XXIII, all five so far known polymorphs are located within the top 4.3 kJ/mol of our stability ranking. We stress that all considered effects are necessary for obtaining accurate results.

Moreover, we have also calculated in Chapter 9 anharmonic free energies for a set of XXIII structures to obtain further insight into this challenging polymorphic energy landscape. Therefore, we have utilized the QHA and the above mentioned Morse oscillators to estimate anharmonic effects. This leads to an average change of relative stabilities by 1 kJ/mol with an observed maximum of about 3 kJ/mol. This suggests that such an analysis might be important for highly polymorphic systems but these anharmonic effects would not lead to significant changes for an already well separated energy landscape. Now, all five experimentally known polymorphs are energetically almost identical; all five structures are located with an energy window of only 1.5 kJ/mol.

However, one structure (XXIII-N70) is by about 3 kJ/mol more stable than all experimentally discovered forms and could potentially be another solid form of system XXIII. Systematic benchmarks carried out in Chapter 10 suggest a reliability of our CSP procedure to around 1-2 kJ/mol, suggesting that this so far unobserved structure N70 might be the thermodynamically most stable form at room temperature. Therefore, an experimental confirmation of this structure would be very helpful and we have made several suggestions for crystallization experiments in Chapter 9. Our findings suggest that late-appearing polymorphs [52] are ubiquitous for pharmaceutically relevant systems, which further reinforces other recent predictions of new solid forms [1, 52, 60, 222, 223].

While we have shown that our hierarchical approach enables in principle the calculation of accurate thermodynamic relative stabilities for pharmaceutically relevant systems, we note that further improvements of this procedure are possible and desirable. For example, the geometry optimizations and the harmonic vibrational free energies could be obtained by using PBE+MBD instead of PBE+TS, which has been discussed in Chapter 10 and applied in Chapter 11 for the five polymorphs of coumarin. In this case, our obtained relative stabilities agree within 1 kJ/mol with the experimental measurements. However, further benchmarks on other complex molecular crystals would be helpful for the demonstration of the general applicability of this procedure. Moreover, the accuracy of the free energy calculations could be further improved upon by employing dynamical approaches such as the vibrational self-consistent field approach [203–205] or path-integral molecular dynamics [206].

In addition, more and more studied molecular crystals exhibit very complex polymorphic energy landscapes, with a large number of structures located within a narrow energy window of only a few kJ/mol. Therefore, accurate stabilities are a prerequisite for the better understanding of such landscapes and we hope that our presented computational procedure will contribute to that. However, thermodynamics alone is often insufficient for the understanding of polymorph crystallization and kinetic effects would have to be considered as well [45, 53, 224–226]. Therefore, novel developments would be required for the modeling of kinetic effects during the nucleation phase and the crystal growth, also taking into account the solvent utilized in the crystallization experiment. Moreover, also the effect of disorder should be properly taken into account, which is challenging due to the required increase in crystallographic space complexity and supercell size [57, 60, 227]. In order to make such a CSP approach broadly applicable, it would be necessary to develop more efficient but still accurate methods. For example, density-functional tight binding approaches could be utilized for certain calculation steps [196, 228] or within embedding schemes [229, 230]. Furthermore, PBE0+MBD energies could be utilized as reference data for the development of efficient machine learning energy models or more accurate tailor-made force fields. Moreover, one could also develop a machine-learning model based on PBE+MBD Hessians obtained for a few selected polymorphs of a given system. Finally, combining such efficient machine-learning models with more advanced enhanced sampling techniques [57, 231] could eventually enable path-integral molecular dynamics simulations for practically relevant molecular crystals. [185]

Part IV
Appendices

Appendix A

Benchmark Calculations for Small Molecular Crystals

This appendix presents the calculated relative stabilities for a benchmark set of small molecular crystals discussed in Chapter 9. The corresponding PBE+TS-optimized structures are available in the Supporting Information of Ref. 185.

Table A.1: Relative energies of a benchmark set of structures with small unit cells calculated with different methods. The benchmark set consists of the single $Z = 1$ structure and the top 7 $Z = 2$ structures after the PBE0+MBD ranking for system XXII, and the top 4 $Z = 2$ structures after the PBE0+MBD ranking for system XXIV. The energy of the most stable structure of a system after the PBE0+MBD ranking was set to zero for all methods. The letters after the slash indicate the used basis set and grid settings: *l* corresponds to the *light* species default settings for basis sets and grids in FHI-aims, while *t* refers to the *tight* species default settings. The label *rt* indicates the usage of the *really tight* grid settings together with tier-3 basis functions. All relative energies are in kJ/mol per formula unit. The method PBE0+MBD without a label for basis set and grid settings corresponds to the PBE0+MBD estimate introduced in Chapter 9 and used for the stability rankings in Chapters 9, 10, and 11 [185].

Structure	PBE+MBD/l	PBE+MBD/t	PBE+MBD/rt	PBE0+MBD/l	PBE0+MBD	PBE0+MBD/t
XXII-N17	0.0	0.0	0.0	0.0	0.0	0.0
XXII-N32	-0.1	0.6	0.6	0.4	1.0	0.8
XXII-N36	0.6	1.3	1.3	0.9	1.6	1.6
XXII-N9	-1.3	-0.3	-0.3	0.8	1.9	1.2
XXII-N92	1.8	1.8	1.8	2.1	2.1	2.1
XXII-N24	-2.4	0.6	0.8	-0.3	2.7	2.3
XXII-N39	0.5	1.1	1.0	2.3	2.9	2.4
XXII-N44	0.9	1.5	1.5	2.5	3.1	2.5
XXIV-N53	0.0	0.0	0.0	0.0	0.0	0.0
XXIV-N6	0.2	0.9	0.8	1.1	1.7	2.0
XXIV-N19	3.1	3.4	3.4	2.5	2.8	3.5
XXIV-N12	2.0	2.2	2.2	3.2	3.4	3.4

Appendix B

Accuracy of Optimized Structures for the Blind-Test Systems

This appendix shows the accuracy of the calculated structures for the five blind-test systems w.r.t. the experimentally measured crystal structures (see Chapters 9 and 10). The corresponding optimized structures are available in the Supporting Information of Ref. 195.

Table B.1: Errors of PBE+TS-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD₂₀ is shown in Å [185].

Exp. str.	a	b	c	α	β	γ	Volume	Density	RMSD ₂₀
XXII	0.46	0.94	2.07		0.76		2.99	-2.90	0.103
XXIII-A	-1.25	-2.12	0.06		-1.42		-3.09	3.18	0.198
XXIII-B	1.13	-0.92	-3.95	4.28	0.26	-1.39	-4.53	4.71	0.277
XXIII-C	-1.43	-1.24	-0.97	2.47	2.01	0.11	-3.19	3.30	0.237
XXIII-D	-3.56	0.74	1.24		2.11		-3.51	3.67	0.447
XXIII-E	-1.15	-3.63	-1.69	-3.05	0.93	-0.55	-4.73	4.96	0.479
XXIV	2.17	-0.99	-0.78		1.04		0.11	-0.06	0.119
XXV	-0.06	-1.30	-1.61		0.72		-3.43	3.51	0.129
XXVI	-1.02	-0.67	-1.95	3.30	0.73	0.86	-2.47	2.53	0.268

B Accuracy of Optimized Structures for the Blind-Test Systems

Table B.2: Errors of PBE+MBD-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD₂₀ is shown in Å [195].

Exp. str.	a	b	c	α	β	γ	Volume	Density	RMSD ₂₀
XXII	1.62	0.96	2.01		1.24		3.80	-3.71	0.119
XXIII-A	-1.21	-0.32	-0.03		-1.52		-1.33	1.33	0.161
XXIII-B	2.45	-0.32	-3.28	3.02	1.04	-1.53	-1.83	1.81	0.243
XXIII-C	-1.57	-1.11	-0.21	2.26	1.91	1.11	-2.25	2.30	0.232
XXIII-D	-2.56	1.53	1.19		2.22		-1.92	2.02	0.414
XXIII-E	-0.27	-2.35	-2.02	-2.63	0.99	-0.40	-3.10	3.18	0.453
XXIV	4.69	-1.11	-1.32		0.06		2.15	-2.10	0.158
XXV	0.08	-1.03	-0.75		0.70		-2.16	2.15	0.102
XXVI	-0.56	-0.21	-2.13	3.18	1.04	1.62	-1.27	1.34	0.271

Table B.3: Errors of PBE0+MBD-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD₂₀ is shown in Å [195].

Exp. str.	a	b	c	α	β	γ	Volume	Density	RMSD ₂₀
XXII	0.22	0.18	0.64		0.62		0.63	-0.64	0.059
XXIII-A	-1.98	-2.30	-0.32		-1.41		-4.34	4.52	0.181
XXIII-B	1.58	-1.78	-3.81	3.14	0.65	-1.37	-4.61	4.86	0.251
XXIII-C	-2.64	-1.99	-1.20	2.54	1.88	0.50	-5.24	5.52	0.237
XXIII-D	-2.34	-0.14	-0.21		1.18		-3.71	3.90	0.272
XXIII-E	-1.96	-2.89	-2.57	-4.29	1.16	-0.13	-5.94	6.29	0.471
XXIV	1.68	-1.26	-1.46		0.46		-1.18	1.21	0.109
XXV	-0.70	-1.57	-2.03		0.94		-4.87	5.09	0.156
XXVI	-1.33	-1.45	-3.69	3.83	1.29	2.91	-4.04	4.23	0.357

Table B.4: Errors of PBE+MBD-optimized structures using tight settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD₂₀ is shown in Å [195].

Exp. str.	a	b	c	α	β	γ	Volume	Density	RMSD ₂₀
XXII	2.48	1.21	2.53		1.66		5.15	-4.92	0.152
XXIII-A	-1.16	-0.81	0.05		-1.19		-1.73	1.78	0.159
XXIII-B	1.94	-0.22	-3.21	2.04	-0.29	-1.66	-2.49	2.54	0.219
XXIII-C	-1.59	-1.23	-0.45	2.12	1.69	0.80	-2.72	2.80	0.220
XXIII-D	-1.90	1.37	1.00		2.44		-1.81	1.87	0.397
XXIII-E	-0.66	-1.70	-2.39	-2.70	1.26	-0.36	-3.24	3.33	0.460
XXIV	4.14	-1.02	-0.98		0.26		2.01	-1.91	0.138
XXV	-0.10	-1.02	-1.11		0.74		-2.71	2.72	0.116
XXVI	-0.96	-0.24	-2.03	2.47	0.99	1.30	-1.87	1.93	0.225

Table B.5: Errors of the thermally-expanded structures (corresponding to 300 K within the QHA) calculated with PBE+TS and light settings w.r.t. experimental structures³⁴. This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD₂₀ is shown in Å [185].

Exp. str.	a	b	c	α	β	γ	Volume	Density	RMSD ₂₀
XXIII-A	-0.46	-0.73	0.28		-1.48		-0.70	0.74	0.193
XXIII-B	2.84	0.75	-3.72	3.97	1.56	-1.49	-0.76	0.72	0.293
XXIII-C	-0.69	-0.75	-0.36	2.03	2.06	0.23	-1.37	1.43	0.244
XXIII-D	-2.34	2.13	1.50		2.07		-0.66	0.67	0.408
XXIII-E	0.35	-2.42	-1.49	-3.16	0.56	-1.02	-1.70	1.70	0.479

Appendix C

Stability Rankings for the Blind-Test Systems

This appendix contains all calculated relative stabilities for the five blind-test systems based on the GRACE initial structures (see Chapter 9). The corresponding PBE+TS-optimized structures are available in the Supporting Information of Ref. 185.

Table C.1: Stability ranking for system XXII in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 150 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm³ [185].

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXII-N2 (Exp.)	4	P 21/n	1.677	1.7	0.7	0.0	0.0
XXII-N3	4	P 21/c	1.672	1.7	0.6	0.6	0.5
XXII-N4	4	P 21/c	1.688	2.5	1.2	1.0	1.9
XXII-N1	4	P n a 21	1.715	0.0	0.0	1.8	2.5
XXII-N19	4	P 21/n	1.669	5.2	4.1	3.9	4.1
XXII-N6	4	P 21/n	1.701	3.7	2.4	3.7	4.4
XXII-N7	4	P 21 21 21	1.703	2.8	1.9	4.3	4.8
XXII-N17	2	P 21	1.639	7.3	4.4	4.4	n/d
XXII-N27	4	P 21 21 21	1.668	6.1	5.1	4.5	n/d
XXII-N20	4	P n a 21	1.634	7.6	4.1	4.5	n/d
XXII-N35	4	P 21/n	1.673	6.3	4.6	4.8	n/d
XXII-N5	8	I 2/c	1.728	2.8	3.2	5.1	n/d
XXII-N8	4	P 21/c	1.709	4.0	3.8	5.2	n/d
XXII-N37	4	P 21/c	1.659	6.9	5.3	5.3	n/d
XXII-N32	2	P -1	1.659	7.2	4.9	5.4	n/d
XXII-N46	4	P 21/c	1.657	7.4	5.3	5.4	n/d
XXII-N13	4	P n m a	1.712	4.2	4.0	5.5	n/d

Table C.1: (continued) Stability ranking for system XXII.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXII-N22	4	P 21/c	1.609	7.8	3.9	5.6	n/d
XXII-N67	4	P 21/n	1.626	8.8	5.2	5.7	n/d
XXII-N52	4	P 21/c	1.663	6.8	5.5	5.9	n/d
XXII-N36	2	P -1	1.669	6.7	5.7	6.0	n/d
XXII-N42	4	P 21/n	1.659	6.9	5.3	6.0	n/d
XXII-N11	4	P n m a	1.712	4.0	3.8	6.1	n/d
XXII-N70	8	P b c a	1.607	8.0	6.0	6.2	n/d
XXII-N9	2	P -1	1.705	5.0	4.1	6.3	n/d
XXII-N34	4	P 21 21 21	1.660	7.6	5.3	6.3	n/d
XXII-N21	4	P 21/n	1.721	4.3	5.0	6.4	n/d
XXII-N18	4	P 21 21 21	1.664	7.4	4.4	6.4	n/d
XXII-N41	8	P b c a	1.684	6.4	5.6	6.5	n/d
XXII-N92	2	P 21	1.666	7.8	6.2	6.5	n/d
XXII-N53	4	P 21/c	1.661	8.0	5.9	6.6	n/d
XXII-N76	4	P -1	1.730	4.5	5.1	6.7	n/d
XXII-N16	4	P 21/n	1.708	4.6	4.6	6.7	n/d
XXII-N30	8	P 21/n	1.729	2.9	4.6	6.9	n/d
XXII-N10	4	P 21/n	1.727	3.1	3.7	6.9	n/d
XXII-N85	4	P 21/c	1.649	7.7	6.7	6.9	n/d
XXII-N26	4	P 21/n	1.657	7.2	5.0	7.0	n/d
XXII-N29	4	P 21/c	1.684	6.1	5.4	7.0	n/d
XXII-N47	4	P 21/c	1.664	7.3	6.3	7.0	n/d
XXII-N88	4	P c a 21	1.729	3.9	5.1	7.1	n/d
XXII-N24	2	P c	1.746	2.6	4.9	7.1	n/d
XXII-N25	4	P 21/c	1.724	3.2	5.7	7.2	n/d
XXII-N43	4	P 21/c	1.670	6.4	5.5	7.2	n/d
XXII-N58	4	P 21 21 21	1.696	5.8	6.0	7.3	n/d
XXII-N39	2	P 21/m	1.718	5.2	5.5	7.3	n/d
XXII-N83	2	P -1	1.614	8.6	6.5	7.3	n/d
XXII-N73	4	P 21/c	1.670	8.0	6.5	7.3	n/d
XXII-N12	4	P 21/c	1.706	4.8	4.9	7.4	n/d
XXII-N62	4	P 21/c	1.635	8.4	5.8	7.4	n/d
XXII-N44	1	P 1	1.730	5.8	5.8	7.5	n/d
XXII-N15	4	P 21/n	1.654	6.9	4.7	7.5	n/d
XXII-N40	4	P c a 21	1.744	4.0	5.7	7.6	n/d
XXII-N23	2	P m n 21	1.689	5.6	4.7	7.6	n/d
XXII-N33	4	P 21/n	1.698	5.5	6.0	7.7	n/d
XXII-N64	2	P 21	1.685	7.5	6.4	7.7	n/d
XXII-N31	8	P b c a	1.701	6.1	5.4	7.9	n/d
XXII-N86	4	P 21/n	1.698	6.2	6.7	8.1	n/d

Table C.1: (continued) Stability ranking for system XXII.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXII-N93	4	P c a 21	1.607	10.8	6.7	8.2	n/d
XXII-N48	4	P 21/c	1.701	6.2	5.5	8.2	n/d
XXII-N77	4	P 21/c	1.715	5.6	6.1	8.2	n/d
XXII-N28	4	P 21/n	1.689	6.3	5.1	8.2	n/d
XXII-N65	4	P 21/c	1.693	5.8	5.6	8.3	n/d
XXII-N84	8	C 2/c	1.668	8.4	7.1	8.3	n/d
XXII-N97	2	P -1	1.702	6.2	5.8	8.3	n/d
XXII-N66	4	P n a 21	1.620	10.6	5.7	8.3	n/d
XXII-N63	4	P 21 21 21	1.717	4.3	6.8	8.3	n/d
XXII-N45	8	P b c a	1.719	5.2	6.4	8.4	n/d
XXII-N55	4	P 21/c	1.664	7.6	6.8	8.4	n/d
XXII-N82	4	P 21/n	1.626	10.0	6.3	8.5	n/d
XXII-N71	4	P 21/n	1.708	5.4	6.5	8.6	n/d
XXII-N74	4	P 21/n	1.725	4.6	6.3	8.6	n/d
XXII-N57	4	P n m a	1.696	6.3	5.8	8.7	n/d
XXII-N61	4	P 21/c	1.717	6.8	6.3	8.7	n/d
XXII-N59	4	P 21/n	1.695	7.2	7.2	8.9	n/d
XXII-N69	2	P -1	1.752	4.9	7.7	9.0	n/d
XXII-N78	2	P 21	1.716	4.8	7.3	9.1	n/d
XXII-N72	4	P 21/c	1.723	5.0	7.5	9.2	n/d
XXII-N100	8	I 2/c	1.711	5.9	7.0	9.2	n/d
XXII-N51	4	P 21/c	1.689	7.6	6.6	9.4	n/d
XXII-N79	4	P 21/c	1.710	6.4	8.1	9.5	n/d
XXII-N87	4	P n a 21	1.718	6.1	7.9	9.6	n/d
XXII-N54	4	P 21/c	1.702	6.2	6.7	9.7	n/d
XXII-N89	2	P -1	1.662	7.3	6.7	9.7	n/d
XXII-N90	4	P 21/c	1.671	8.1	6.9	9.7	n/d
XXII-N81	4	P 21/c	1.691	7.7	7.5	10.0	n/d
XXII-N99	4	P -1	1.691	7.4	7.4	10.0	n/d
XXII-N75	8	P b c a	1.703	6.6	7.3	10.1	n/d
XXII-N80	4	P 21/c	1.705	6.9	7.3	10.1	n/d
XXII-N91	2	P -1	1.710	7.6	7.5	10.2	n/d
XXII-N95	4	P 21/c	1.687	7.4	7.4	10.2	n/d
XXII-N68	4	P 21/n	1.719	5.8	7.6	10.3	n/d
XXII-N96	4	P 21 21 21	1.639	9.0	6.7	10.6	n/d
XXII-N94	4	P 21/c	1.671	8.9	7.5	10.7	n/d
XXII-N98	4	P 21/c	1.665	8.6	7.2	10.7	n/d

Table C.2: Stability ranking for system XXIII in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the asymmetric unit (Z'), the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm³ [185].

Name	Z'	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIII-N70	1	4	P 21/n	1.388	11.1	3.4	4.0	0.0
XXIII-N5	2	4	P -1	1.440	6.9	0.7	0.1	1.4
XXIII-E (E)	2	4	P -1	1.418	9.4	2.7	3.2	1.4
XXIII-N6 (C)	2	4	P -1	1.440	6.9	0.7	0.1	1.5
XXIII-N31	1	2	P -1	1.424	6.4	2.7	3.9	2.1
XXIII-N18	1	2	P -1	1.454	6.7	0.5	0.0	2.2
XXIII-N2	2	4	P -1	1.434	4.4	0.0	2.1	2.2
XXIII-N4 (B)	1	2	P -1	1.445	4.6	2.0	1.5	2.5
XXIII-N42	1	2	P -1	1.433	6.0	1.3	1.8	3.0
XXIII-N22	2	4	P -1	1.454	4.6	2.9	3.1	3.5
XXIII-N1	2	8	P 21/n	1.425	11.3	2.6	4.4	3.5
XXIII-N13	2	8	P 21/n	1.415	12.8	3.2	4.7	3.6
XXIII-N100	1	4	P 21/n	1.449	10.5	4.0	3.9	3.7
XXIII-N7	2	4	P -1	1.436	6.4	1.9	2.5	3.8
XXIII-N68	1	4	P 21/n	1.423	7.2	2.0	2.6	3.9
XXIII-N85 (A)	1	4	P 21/c	1.394	9.0	4.6	5.3	4.1
XXIII-N3	1	8	I 2/a	1.440	1.3	1.5	2.7	4.1
XXIII-N39 (D)	1	4	P 21/n	1.384	11.1	4.9	6.7	4.3
XXIII-N53	1	2	P -1	1.415	7.5	4.5	5.1	4.4
XXIII-N28	1	2	P -1	1.430	1.8	2.5	3.5	4.4
XXIII-N33	2	4	P -1	1.433	7.8	4.7	4.5	4.5
XXIII-N10	2	8	P 21/n	1.431	9.6	2.3	3.8	4.6
XXIII-N26	2	8	P 21/c	1.427	8.0	3.7	4.0	4.8
XXIII-N8	2	8	P 21/n	1.425	10.2	2.4	4.2	5.0
XXIII-N46	2	4	P -1	1.440	6.6	3.9	4.7	5.0
XXIII-N38	1	4	P 21/c	1.390	11.6	2.0	4.7	5.1
XXIII-N30	2	4	P -1	1.442	9.2	3.3	3.7	5.2
XXIII-N25	2	4	P -1	1.442	8.9	3.2	3.6	5.3
XXIII-N20	2	8	P 21/c	1.437	1.7	2.8	4.3	5.5
XXIII-N52	2	4	P -1	1.449	7.8	5.0	5.0	5.6
XXIII-N80	1	4	P 21/c	1.411	11.0	5.3	5.2	5.6
XXIII-N23	2	8	P n a 21	1.438	1.8	3.0	4.6	5.8
XXIII-N24	2	8	P 21/c	1.424	6.4	2.4	4.8	6.1
XXIII-N95	1	2	P -1	1.404	10.4	6.3	7.1	7.0
XXIII-N37	1	4	P 21/c	1.459	2.0	3.0	6.6	7.6

Table C.2: (continued) Stability ranking for system XXIII.

Name	Z'	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIII-N41	1	4	P 21/c	1.425	8.8	4.5	7.2	8.0
XXIII-N56	1	2	P -1	1.448	7.7	4.3	5.1	9.0
XXIII-N49	1	2	P -1	1.414	10.0	3.7	5.8	9.2
XXIII-N91	1	2	P -1	1.414	8.7	6.0	7.2	9.2
XXIII-N43	1	4	P 21/n	1.451	8.6	5.8	6.4	9.3
XXIII-N66	1	2	P -1	1.465	0.0	3.8	7.2	9.8
XXIII-N62	1	4	P 21/c	1.406	8.8	5.0	7.2	10.3
XXIII-N88	1	4	P 21/c	1.440	6.4	6.7	8.7	10.9
XXIII-N48	1	4	P 21/c	1.411	8.1	5.8	8.0	11.4
XXIII-N61	1	4	P 21/c	1.459	7.7	5.1	7.5	11.9
XXIII-N81	1	4	P 21/c	1.443	5.9	5.0	8.4	12.5
XXIII-N35	2	16	C 2/c	1.459	6.5	4.5	4.5	n/d
XXIII-N76	2	16	C 2/c	1.421	3.0	3.4	4.8	n/d
XXIII-N47	2	4	P -1	1.445	7.9	5.0	5.0	n/d
XXIII-N21	2	8	P 21/n	1.416	13.0	3.8	5.0	n/d
XXIII-N83	2	4	P 21	1.425	6.1	5.3	5.1	n/d
XXIII-N40	2	4	P -1	1.431	8.6	4.2	5.2	n/d
XXIII-N60	2	4	P -1	1.446	8.0	5.2	5.2	n/d
XXIII-N59	2	4	P -1	1.447	7.9	5.2	5.2	n/d
XXIII-N27	2	16	C 2/c	1.435	3.0	3.3	5.2	n/d
XXIII-N57	2	4	P -1	1.447	7.8	5.1	5.2	n/d
XXIII-N54	2	16	C 2/c	1.450	7.0	5.0	5.3	n/d
XXIII-N63	2	4	P -1	1.433	8.7	5.4	5.4	n/d
XXIII-N36	2	4	P -1	1.428	9.4	5.2	5.4	n/d
XXIII-N32	2	8	P 21/c	1.440	10.0	4.2	5.6	n/d
XXIII-N64	2	8	P 21/c	1.437	4.0	4.6	5.6	n/d
XXIII-N29	2	8	P 21/c	1.429	12.0	4.3	5.9	n/d
XXIII-N71	2	4	P -1	1.447	7.9	4.5	5.9	n/d
XXIII-N73	2	16	I 2/c	1.445	7.4	5.5	6.0	n/d
XXIII-N97	2	4	P -1	1.410	10.3	5.7	6.1	n/d
XXIII-N67	2	4	P -1	1.420	9.0	5.8	6.2	n/d
XXIII-N90	2	4	P -1	1.437	7.5	5.7	6.2	n/d
XXIII-N77	2	4	P -1	1.452	7.4	4.5	6.3	n/d
XXIII-N45	2	16	I 2/c	1.424	5.6	4.3	6.3	n/d
XXIII-N65	2	4	P -1	1.422	8.9	5.8	6.3	n/d
XXIII-N86	2	8	C c	1.432	7.6	5.3	6.3	n/d
XXIII-N89	2	4	P -1	1.450	8.7	4.9	6.4	n/d
XXIII-N87	4	16	P 2/c	1.449	8.3	6.3	6.5	n/d
XXIII-N84	2	8	P 21/c	1.452	2.0	3.9	6.5	n/d
XXIII-N82	2	8	C c	1.424	7.3	5.4	6.6	n/d

Table C.2: (continued) Stability ranking for system XXIII.

Name	Z'	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIII-N44	2	4	P -1	1.431	11.0	4.8	6.6	n/d
XXIII-N55	2	8	P 21/c	1.409	6.7	5.0	6.6	n/d
XXIII-N34	2	16	I 2/c	1.420	4.6	4.3	6.7	n/d
XXIII-N69	2	4	P -1	1.436	8.9	5.2	7.0	n/d
XXIII-N96	2	4	P -1	1.446	9.2	5.0	7.1	n/d
XXIII-N98	2	4	P -1	1.432	9.2	5.5	7.2	n/d
XXIII-N58	2	8	P 21/c	1.403	11.9	5.0	7.3	n/d
XXIII-N93	2	4	P -1	1.445	8.0	5.8	7.3	n/d
XXIII-N51	2	4	P -1	1.422	13.8	5.9	7.5	n/d
XXIII-N99	2	16	I 2/a	1.440	4.4	6.0	7.6	n/d
XXIII-N72	2	4	P -1	1.419	11.8	5.6	7.9	n/d
XXIII-N75	2	4	P -1	1.419	12.0	5.8	7.9	n/d
XXIII-N79	2	8	P 21/n	1.410	14.9	6.5	8.3	n/d
XXIII-N74	2	4	P -1	1.438	12.0	6.7	8.4	n/d
XXIII-N94	2	4	P -1	1.442	10.2	7.0	8.5	n/d

Table C.3: Stability ranking for system XXIV in kJ/mol normalized per formula unit. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 240 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm³ [185].

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIV-N2 (Exp.)	4	P 21/c	1.570	0.3	1.5	0.4	0.0
XXIV-N3	4	P 21/c	1.543	1.2	0.5	0.0	0.3
XXIV-N10	8	C 2/c	1.521	4.1	2.0	2.4	1.1
XXIV-N53	2	P -1	1.566	0.0	0.8	0.8	1.3
XXIV-N7	4	P 21/n	1.549	2.3	1.4	2.7	1.4
XXIV-N4	4	P 21/c	1.551	1.4	2.1	1.3	1.5
XXIV-N6	2	P -1	1.535	2.9	1.7	2.6	1.9
XXIV-N1	4	P 21 21 21	1.534	1.2	0.0	2.6	3.1
XXIV-N8	4	P 21/n	1.561	3.4	3.6	3.3	n/d
XXIV-N9	8	C 2/c	1.571	2.8	3.9	3.5	n/d
XXIV-N19	2	P -1	1.556	4.6	4.2	3.6	n/d
XXIV-N12	2	P -1	1.541	3.9	3.1	4.2	n/d
XXIV-N16	4	P 21/c	1.502	5.0	3.0	5.2	n/d
XXIV-N21	8	I 2/a	1.535	4.6	4.7	5.2	n/d
XXIV-N18	8	P c c n	1.507	4.9	4.1	5.5	n/d
XXIV-N33	4	P 21/c	1.571	5.6	5.7	5.7	n/d
XXIV-N22	8	I 2/c	1.505	5.8	4.8	5.8	n/d
XXIV-N43	4	P 21/c	1.580	7.4	6.8	6.2	n/d
XXIV-N11	4	P 21/c	1.479	6.0	3.3	6.2	n/d
XXIV-N23	2	P -1	1.556	5.4	5.2	6.3	n/d
XXIV-N40	8	P b c n	1.506	7.1	6.0	6.6	n/d
XXIV-N30	4	P 21/n	1.527	7.2	5.5	6.6	n/d
XXIV-N31	8	I b a 2	1.505	6.7	5.6	6.8	n/d
XXIV-N29	2	P -1	1.506	9.0	6.3	6.8	n/d
XXIV-N27	4	P 21/n	1.572	6.6	7.7	6.8	n/d
XXIV-N45	4	P 21/c	1.543	7.8	7.5	7.2	n/d
XXIV-N5	8	P b c a	1.522	4.6	4.3	7.2	n/d
XXIV-N63	4	P 21/c	1.527	9.8	8.1	7.4	n/d
XXIV-N14	4	P 21 21 21	1.523	7.2	4.7	7.4	n/d
XXIV-N13	4	P 21 21 21	1.518	6.8	4.2	7.4	n/d
XXIV-N44	8	C 2/c	1.549	8.0	7.5	7.5	n/d
XXIV-N50	8	P b c n	1.524	7.5	7.2	7.7	n/d
XXIV-N32	8	C 2/c	1.428	8.4	4.9	7.7	n/d
XXIV-N34	4	P 21/c	1.456	8.0	5.3	7.9	n/d
XXIV-N55	4	P 21 21 21	1.548	9.1	8.9	7.9	n/d
XXIV-N17	8	P b c n	1.602	6.5	6.6	8.1	n/d

Table C.3: (continued) Stability ranking for system XXIV.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIV-N24	4	P 21/n	1.536	7.9	6.7	8.2	n/d
XXIV-N58	4	P 21/n	1.538	6.6	7.7	8.2	n/d
XXIV-N78	2	P -1	1.549	8.3	8.9	8.3	n/d
XXIV-N20	4	P n a 21	1.500	7.7	4.6	8.3	n/d
XXIV-N65	4	P 21/n	1.484	11.9	8.0	8.3	n/d
XXIV-N52	4	P 21/c	1.461	10.5	7.5	8.4	n/d
XXIV-N49	8	C 2/c	1.559	7.0	8.3	8.5	n/d
XXIV-N67	4	P -1	1.483	7.7	5.4	8.6	n/d
XXIV-N36	8	P c c n	1.513	8.3	7.2	8.7	n/d
XXIV-N62	8	P c c n	1.419	11.5	6.9	8.7	n/d
XXIV-N39	2	P -1	1.502	7.7	6.0	8.7	n/d
XXIV-N73	4	P 21/c	1.532	9.2	8.2	8.9	n/d
XXIV-N42	2	P -1	1.513	9.5	8.3	9.2	n/d
XXIV-N35	2	P -1	1.473	8.1	6.6	9.2	n/d
XXIV-N54	4	P 21/n	1.463	10.3	6.6	9.2	n/d
XXIV-N37	4	P 21/c	1.479	7.9	5.8	9.3	n/d
XXIV-N47	4	P 21/c	1.466	9.5	6.2	9.5	n/d
XXIV-N25	4	P 21 21 21	1.545	8.5	7.1	10.0	n/d
XXIV-N51	16	F d d 2	1.563	9.5	8.8	10.1	n/d
XXIV-N92	8	C c	1.537	9.3	7.5	10.3	n/d
XXIV-N85	4	P 21 21 21	1.550	9.9	9.9	10.3	n/d
XXIV-N64	4	P 21/c	1.519	11.6	9.3	10.3	n/d
XXIV-N60	2	P -1	1.449	11.7	8.1	10.5	n/d
XXIV-N70	16	F d d 2	1.479	10.2	9.2	10.6	n/d
XXIV-N66	2	P 21	1.555	9.0	9.2	10.8	n/d
XXIV-N90	4	P c a 21	1.472	11.4	8.1	10.8	n/d
XXIV-N94	8	P c c n	1.527	10.6	9.5	11.1	n/d
XXIV-N88	16	F d d 2	1.595	10.0	11.2	11.1	n/d
XXIV-N87	2	P -1	1.473	10.1	8.5	11.1	n/d
XXIV-N48	4	P 21/n	1.417	12.5	7.2	11.2	n/d
XXIV-N84	8	I 2/c	1.514	12.7	9.1	11.2	n/d
XXIV-N91	4	P 21/c	1.427	12.7	8.6	11.4	n/d
XXIV-N59	4	P 21 21 21	1.481	11.5	7.0	11.4	n/d
XXIV-N74	4	P 21/c	1.490	10.6	8.1	11.5	n/d
XXIV-N81	4	P 21/c	1.504	13.0	9.9	11.5	n/d
XXIV-N89	4	P 21/n	1.491	11.2	9.5	11.5	n/d
XXIV-N99	8	P c c n	1.522	10.0	10.2	11.7	n/d
XXIV-N97	4	P 21/c	1.551	11.9	11.0	11.9	n/d
XXIV-N77	4	P 21/c	1.498	11.3	9.7	11.9	n/d
XXIV-N61	2	P -1	1.494	10.1	8.6	12.0	n/d

Table C.3: (continued) Stability ranking for system XXIV.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXIV-N75	4	P 21/c	1.508	11.7	9.6	12.1	n/d
XXIV-N95	4	P 21/c	1.486	10.7	9.6	12.1	n/d
XXIV-N93	4	P 21/n	1.532	12.9	12.1	12.2	n/d
XXIV-N69	4	P n a 21	1.480	13.0	8.7	12.4	n/d
XXIV-N100	4	P 21/n	1.548	13.7	13.0	12.4	n/d
XXIV-N82	4	P 21 21 21	1.508	12.4	9.7	12.7	n/d
XXIV-N96	4	P 21 21 21	1.473	11.8	9.4	12.7	n/d
XXIV-N68	4	P 21/n	1.501	11.6	8.9	12.7	n/d
XXIV-N86	16	F d d 2	1.531	11.9	10.2	12.8	n/d
XXIV-N76	2	P 21	1.415	13.9	8.7	12.8	n/d
XXIV-N71	4	P 21/c	1.344	13.9	7.4	13.0	n/d
XXIV-N80	8	I 2/c	1.446	11.2	8.2	13.1	n/d
XXIV-N83	4	P 21/c	1.469	13.2	9.7	13.7	n/d
XXIV-N79	4	P 21/n	1.461	12.2	9.4	14.4	n/d
XXIV-N38	4	P 21/c	1.419	12.6	7.8	14.9	n/d
XXIV-N46	4	P n a 21	1.417	13.1	8.3	15.2	n/d

Table C.4: Stability ranking for system XXV in kJ/mol normalized per formula unit. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm³ [185].

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXV-N6 (Exp.)	4	P 21/c	1.445	0.5	2.3	0.0	0.0
XXV-N5	8	P 21/c	1.431	0.0	0.0	1.0	3.8
XXV-N11	4	P 2/c	1.411	11.9	3.6	2.5	5.7
XXV-N1	2	P -1	1.453	4.4	3.2	4.2	7.0
XXV-N12	8	P b c n	1.397	13.8	5.1	4.2	n/d
XXV-N15	8	P b c n	1.409	13.3	4.4	4.3	n/d
XXV-N18	4	P c a 21	1.393	14.4	5.3	4.6	n/d
XXV-N22	2	P -1	1.431	7.0	6.7	4.7	n/d
XXV-N3	2	P -1	1.444	5.5	4.3	4.8	n/d
XXV-N16	4	P 21/c	1.422	4.0	3.9	5.1	n/d
XXV-N4	2	P -1	1.417	6.7	5.4	5.7	n/d
XXV-N2	4	P 21/n	1.437	7.6	3.8	6.0	n/d
XXV-N33	8	P b c n	1.407	14.1	7.4	6.1	n/d
XXV-N10	4	P 21/c	1.399	12.5	5.8	6.3	n/d
XXV-N29	4	P -1	1.425	10.9	8.3	6.8	n/d
XXV-N28	2	P -1	1.435	7.6	8.3	7.1	n/d
XXV-N19	4	P 21/c	1.417	13.3	8.5	7.1	n/d
XXV-N7	4	P 21/n	1.430	9.2	5.4	7.1	n/d
XXV-N36	8	P b c a	1.387	17.1	8.6	7.6	n/d
XXV-N25	4	P 21/n	1.404	11.3	6.9	7.8	n/d
XXV-N13	4	P 21/n	1.413	10.8	5.7	7.9	n/d
XXV-N27	8	I 2/c	1.419	14.5	8.8	8.2	n/d
XXV-N48	4	P 21/n	1.409	11.7	7.9	8.6	n/d
XXV-N26	2	P -1	1.418	12.6	7.9	8.7	n/d
XXV-N23	4	P 21/c	1.433	10.7	8.4	9.0	n/d
XXV-N9	4	P 21/c	1.433	10.6	7.8	9.1	n/d
XXV-N42	4	P 21/n	1.424	12.6	9.6	9.2	n/d
XXV-N20	8	P b c n	1.404	12.9	8.6	9.3	n/d
XXV-N49	2	P -1	1.401	11.0	9.7	9.6	n/d
XXV-N39	16	F d d 2	1.392	15.4	9.6	9.9	n/d
XXV-N30	4	P 21/c	1.411	18.0	9.7	9.9	n/d
XXV-N38	4	P 21/c	1.370	17.2	9.0	9.9	n/d
XXV-N32	2	P -1	1.408	15.3	9.2	10.2	n/d
XXV-N17	16	F d d 2	1.391	14.6	7.8	10.6	n/d
XXV-N44	4	P 21/c	1.395	16.2	9.4	10.9	n/d
XXV-N65	4	P 21/n	1.424	17.3	9.9	n/d	n/d

Table C.4: (continued) Stability ranking for system XXV.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXV-N24	4	P 21/c	1.395	14.6	10.0	n/d	n/d
XXV-N66	4	P 21/c	1.396	16.3	10.7	n/d	n/d
XXV-N60	8	A b a 2	1.427	11.6	10.8	n/d	n/d
XXV-N50	4	P 21/n	1.428	14.7	10.8	n/d	n/d
XXV-N45	1	P 1	1.387	18.7	11.1	n/d	n/d
XXV-N47	2	P -1	1.423	13.8	11.6	n/d	n/d
XXV-N59	8	P b c a	1.422	17.4	11.9	n/d	n/d
XXV-N43	4	P 21/c	1.417	19.6	11.9	n/d	n/d
XXV-N100	4	P 21/c	1.397	18.5	11.9	n/d	n/d
XXV-N71	4	P 21/c	1.418	18.4	12.0	n/d	n/d
XXV-N70	4	C c	1.410	14.7	12.0	n/d	n/d
XXV-N77	4	P 21/c	1.383	22.4	12.2	n/d	n/d
XXV-N61	4	P 21/n	1.409	16.1	12.3	n/d	n/d
XXV-N62	4	P 21/c	1.409	18.1	12.4	n/d	n/d
XXV-N40	4	P n a 21	1.437	15.0	12.5	n/d	n/d
XXV-N79	2	P -1	1.424	19.1	12.6	n/d	n/d
XXV-N55	16	I 41/a	1.396	18.2	12.6	n/d	n/d
XXV-N80	16	F d d 2	1.383	17.3	12.7	n/d	n/d
XXV-N57	8	P b c a	1.409	17.8	12.8	n/d	n/d
XXV-N52	8	P b c a	1.419	18.4	12.9	n/d	n/d
XXV-N53	2	P -1	1.437	14.0	13.0	n/d	n/d
XXV-N68	2	P n	1.377	20.6	13.0	n/d	n/d
XXV-N83	4	P 21/c	1.378	21.5	13.0	n/d	n/d
XXV-N81	2	P 21	1.415	15.0	13.0	n/d	n/d
XXV-N41	4	P 21/c	1.381	16.9	13.2	n/d	n/d
XXV-N64	4	P 21/c	1.408	18.7	13.2	n/d	n/d
XXV-N86	4	P 21 21 21	1.401	17.8	13.3	n/d	n/d
XXV-N58	4	P 21/c	1.423	17.3	13.4	n/d	n/d
XXV-N69	2	P -1	1.399	20.6	13.7	n/d	n/d
XXV-N73	2	P 21	1.427	19.7	13.8	n/d	n/d
XXV-N87	2	P -1	1.425	18.2	13.9	n/d	n/d
XXV-N84	8	I 2/a	1.363	27.3	13.9	n/d	n/d
XXV-N72	8	I 2/c	1.407	18.3	13.9	n/d	n/d
XXV-N78	4	P 21/c	1.379	21.3	13.9	n/d	n/d
XXV-N46	4	P 21/n	1.423	15.8	14.1	n/d	n/d
XXV-N89	4	P 21/c	1.420	15.0	14.3	n/d	n/d
XXV-N56	8	I 2/c	1.379	21.1	14.4	n/d	n/d
XXV-N63	2	P -1	1.416	19.2	14.5	n/d	n/d
XXV-N82	2	P -1	1.408	21.2	14.5	n/d	n/d
XXV-N74	4	P 21/c	1.404	18.9	14.7	n/d	n/d

Table C.4: (continued) Stability ranking for system XXV.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXV-N75	4	P 21/n	1.398	19.1	15.1	n/d	n/d
XXV-N85	8	I 2/c	1.429	16.3	15.2	n/d	n/d
XXV-N90	4	P n a 21	1.412	19.1	15.6	n/d	n/d
XXV-N91	4	P 21/c	1.383	20.7	15.7	n/d	n/d
XXV-N54	8	P b c a	1.407	18.9	16.1	n/d	n/d
XXV-N92	8	C 2/c	1.383	22.1	16.1	n/d	n/d
XXV-N88	8	P b c a	1.410	21.3	16.2	n/d	n/d
XXV-N94	8	C 2/c	1.387	23.8	16.6	n/d	n/d
XXV-N97	4	P 21/n	1.360	25.7	16.8	n/d	n/d
XXV-N95	2	P -1	1.395	27.7	17.0	n/d	n/d
XXV-N93	8	I 2/a	1.401	21.4	17.1	n/d	n/d
XXV-N96	4	P 21/n	1.387	19.2	17.2	n/d	n/d
XXV-N99	2	P -1	1.387	21.5	17.4	n/d	n/d
XXV-N76	4	P c a 21	1.401	22.4	17.5	n/d	n/d
XXV-N98	2	P -1	1.426	20.4	20.5	n/d	n/d

Table C.5: Stability ranking for system XXVI in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm³ [185].

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXVI-N1 (Exp.)	2	P -1	1.380	6.1	0.0	0.0	0.0
XXVI-N5	8	P 21/c	1.392	4.7	1.7	1.6	0.2
XXVI-N4	2	P -1	1.377	7.2	1.4	1.1	0.3
XXVI-N9	4	P 21/n	1.410	0.0	0.6	1.2	2.0
XXVI-N6	8	P 21/c	1.407	1.6	2.0	2.5	3.0
XXVI-N3	4	P -1	1.391	6.3	2.6	2.7	3.3
XXVI-N25	2	P -1	1.388	8.0	3.4	2.6	3.4
XXVI-N11	4	P -1	1.403	2.0	2.5	2.7	3.6
XXVI-N17	4	P -1	1.419	4.5	3.7	3.3	n/d
XXVI-N10	4	P -1	1.394	4.6	3.2	3.6	n/d
XXVI-N15	4	P -1	1.417	4.5	3.8	3.7	n/d
XXVI-N99	8	P 21/c	1.390	6.5	3.6	3.8	n/d
XXVI-N23	4	P -1	1.422	4.1	3.9	4.0	n/d
XXVI-N12	4	P 21	1.402	3.1	3.2	4.1	n/d
XXVI-N13	4	P -1	1.381	8.1	3.6	4.4	n/d
XXVI-N8	4	P -1	1.390	7.7	3.7	4.5	n/d
XXVI-N22	4	P -1	1.434	4.1	4.3	4.6	n/d
XXVI-N34	4	P -1	1.381	9.9	4.1	5.0	n/d
XXVI-N19	4	P -1	1.400	7.3	4.1	5.3	n/d
XXVI-N24	8	C c	1.390	8.5	4.0	5.3	n/d
XXVI-N42	4	P -1	1.376	10.4	4.5	5.6	n/d
XXVI-N28	4	P -1	1.374	9.8	4.4	5.9	n/d
XXVI-N40	4	P -1	1.380	9.5	4.6	n/d	n/d
XXVI-N18	4	P -1	1.390	7.9	4.6	n/d	n/d
XXVI-N48	4	P -1	1.414	6.1	4.8	n/d	n/d
XXVI-N63	4	P -1	1.397	8.2	4.9	n/d	n/d
XXVI-N33	4	P -1	1.376	9.4	4.9	n/d	n/d
XXVI-N32	4	P -1	1.379	9.6	4.9	n/d	n/d
XXVI-N38	4	P -1	1.414	5.9	5.0	n/d	n/d
XXVI-N41	4	P -1	1.378	9.4	5.0	n/d	n/d
XXVI-N26	4	P -1	1.426	4.4	5.0	n/d	n/d
XXVI-N72	4	P -1	1.382	9.3	5.0	n/d	n/d
XXVI-N20	16	C 2/c	1.381	9.3	5.1	n/d	n/d
XXVI-N27	4	P -1	1.420	4.6	5.1	n/d	n/d
XXVI-N62	4	P -1	1.412	6.4	5.1	n/d	n/d
XXVI-N30	4	P -1	1.374	9.5	5.1	n/d	n/d

Table C.5: (continued) Stability ranking for system XXVI.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXVI-N16	4	P -1	1.411	5.8	5.2	n/d	n/d
XXVI-N60	4	P -1	1.413	5.9	5.2	n/d	n/d
XXVI-N45	2	P 1	1.429	5.4	5.2	n/d	n/d
XXVI-N47	4	P -1	1.422	5.5	5.2	n/d	n/d
XXVI-N65	4	P -1	1.383	9.6	5.2	n/d	n/d
XXVI-N52	4	P -1	1.419	5.7	5.3	n/d	n/d
XXVI-N67	4	P -1	1.377	10.1	5.3	n/d	n/d
XXVI-N58	4	P -1	1.380	9.6	5.3	n/d	n/d
XXVI-N14	4	P -1	1.385	9.0	5.3	n/d	n/d
XXVI-N54	4	P -1	1.375	9.9	5.4	n/d	n/d
XXVI-N51	4	P -1	1.414	6.3	5.5	n/d	n/d
XXVI-N84	4	P -1	1.412	7.8	5.5	n/d	n/d
XXVI-N61	4	P -1	1.428	5.8	5.5	n/d	n/d
XXVI-N78	4	P -1	1.393	9.0	5.6	n/d	n/d
XXVI-N39	4	P -1	1.410	6.9	5.6	n/d	n/d
XXVI-N31	4	P -1	1.395	8.4	5.6	n/d	n/d
XXVI-N64	4	P -1	1.431	5.4	5.7	n/d	n/d
XXVI-N44	4	P -1	1.378	10.1	5.7	n/d	n/d
XXVI-N85	4	P -1	1.409	7.9	5.7	n/d	n/d
XXVI-N74	2	P 1	1.374	10.5	5.8	n/d	n/d
XXVI-N75	4	P -1	1.375	10.7	5.8	n/d	n/d
XXVI-N35	16	C 2/c	1.375	10.8	5.8	n/d	n/d
XXVI-N50	4	P -1	1.418	6.4	5.8	n/d	n/d
XXVI-N49	4	P -1	1.409	7.6	5.9	n/d	n/d
XXVI-N82	4	P -1	1.376	10.4	5.9	n/d	n/d
XXVI-N83	4	P -1	1.412	7.2	5.9	n/d	n/d
XXVI-N68	4	P -1	1.420	5.9	5.9	n/d	n/d
XXVI-N55	2	P -1	1.416	7.7	5.9	n/d	n/d
XXVI-N71	4	P -1	1.413	7.0	5.9	n/d	n/d
XXVI-N96	4	P -1	1.416	6.9	5.9	n/d	n/d
XXVI-N56	2	P 1	1.414	6.5	6.0	n/d	n/d
XXVI-N100	4	P -1	1.403	8.3	6.0	n/d	n/d
XXVI-N87	4	P n	1.395	8.4	6.0	n/d	n/d
XXVI-N29	16	C 2/c	1.410	8.1	6.1	n/d	n/d
XXVI-N86	4	P -1	1.424	6.2	6.1	n/d	n/d
XXVI-N94	4	P -1	1.391	8.8	6.1	n/d	n/d
XXVI-N70	4	P -1	1.419	6.7	6.1	n/d	n/d
XXVI-N77	4	P -1	1.428	5.1	6.2	n/d	n/d
XXVI-N93	4	P -1	1.386	11.5	6.2	n/d	n/d
XXVI-N73	4	P 21/n	1.401	3.6	6.2	n/d	n/d

Table C.5: (continued) Stability ranking for system XXVI.

Name	Z	Symm.	Density	PBE+TS	PBE+MBD	PBE0+MBD	F
XXVI-N91	4	P 21	1.387	10.8	6.2	n/d	n/d
XXVI-N92	4	P -1	1.372	12.0	6.3	n/d	n/d
XXVI-N80	8	C c	1.391	10.0	6.3	n/d	n/d
XXVI-N57	4	P -1	1.425	5.3	6.3	n/d	n/d
XXVI-N89	4	P -1	1.415	7.4	6.4	n/d	n/d
XXVI-N79	4	P -1	1.382	10.7	6.4	n/d	n/d
XXVI-N37	4	P 21	1.382	10.2	6.4	n/d	n/d
XXVI-N81	4	P -1	1.410	7.1	6.5	n/d	n/d
XXVI-N46	8	C 2/c	1.421	8.3	6.5	n/d	n/d
XXVI-N88	4	P -1	1.421	5.9	6.5	n/d	n/d
XXVI-N43	8	C 2/c	1.387	7.0	6.6	n/d	n/d
XXVI-N59	16	C 2/c	1.396	8.8	6.7	n/d	n/d
XXVI-N90	8	C c	1.392	8.5	6.7	n/d	n/d
XXVI-N98	4	P -1	1.405	7.7	6.8	n/d	n/d
XXVI-N97	4	P -1	1.406	7.9	6.9	n/d	n/d
XXVI-N53	8	C 2/c	1.394	9.5	7.1	n/d	n/d
XXVI-N95	4	P 21/c	1.420	8.7	7.6	n/d	n/d
XXVI-N69	2	P -1	1.446	7.1	8.6	n/d	n/d

Abbreviations

AO	A tomic O rbital
CCDC	C ambridge C ystallographic D ata C entre
CCSD(T)	C oupled C luster with S ingle, D ouble, and perturbative T riple excitations
CI	C onfiguration I nteraction
CSP	C rystal S tructure P rediction
DFA	D ensity F unctional A pproximation
DFT	D ensity F unctional T heory
EOS	E quation O f S tate
GGA	G eneralized G radient A pproximation
HA	H armonic A pproximation
HF	H artree F ock
HMB	H examethyl b enzene
HMBI	H ybrid M any- B ody I nteraction (fragment model)
INS	I nelastic N eutron S cattering
IR	I nfrared
LCAO	L inear C ombination of A tomic O rbitals
LDA	L ocal D ensity A pproximation
MAD	M ean A bsolute D eviation
MAE	M ean A bsolute E rror
MAX	M aximum absolute error
MBD	M any- B ody D ispersion (vdW model)
MO	M olecular O rbital
MP2	S econd order M øller- P lesset perturbation theory

Abbreviations

ND₃	deuterated ammonia
NMR	N uclear M agnetic R esonance
pDOS	p honon D ensity O f S tates
PES	P otential E nergy S urface
PBE	P erdew- B urke- E rnzerhof (density functional)
PBE0	Hybrid version of PBE
QHA	Q uasi- H armonic A pproximation
RMSD₂₀	R oot M ean S quare D eviation calculated using a cluster of 20 molecules
rsSCS	range-separated S elf- C onsistent S creening
SCF	S elf- C onsistent F ield
TS	T katchenko- S cheffler (vdW model)
vdW	v an d er W aals
XDM	E xchange D ipole M oment (vdW model)
2D	T wo- D imensional
3D	T hree- D imensional

Publication List

Publications directly related to this thesis:

- [J. Hoja](#), H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio Jr., A. Tkatchenko, Reliable and Practical Computational Description of Molecular Crystal Polymorphs, *Sci. Adv.* **2019**, *5*, eaau3338.
- [J. Hoja](#), A. Tkatchenko, First-Principles Stability Ranking of Molecular Crystal Polymorphs with the DFT+MBD Approach, *Faraday Discuss.* **2018**, *211*, 253–274.
- G. Folpini, K. Reimann, M. Woerner, T. Elsaesser, [J. Hoja](#), A. Tkatchenko, Strong Local-Field Enhancement of the Nonlinear Soft-Mode Response in a Molecular Crystal, *Phys. Rev. Lett.* **2017**, *119*, 097404.
- [J. Hoja](#), A. M. Reilly, A. Tkatchenko, First-principles Modeling of Molecular Crystals: Structures and Stabilities, Temperature and Pressure, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1294.
- A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, [J. Hoja](#), E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman, B. Kahr, Powder Diffraction and Crystal Structure Prediction Identify Four New Coumarin Polymorphs, *Chem. Sci.* **2017**, *8*, 4926–4940.
- A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr., A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, [J. Hoja](#), R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang,

Publication List

R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu, C. R. Groom, Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods, *Acta Crystallogr. Sect. B: Struct. Sci.* **2016**, *72*, 439–459.

Other publications:

- A. V. Yakutovich, [J. Hoja](#), D. Passerone, A. Tkatchenko, C. A. Pignedoli, Hidden Beneath the Surface: Origin of the Observed Enantioselective Adsorption on PdGa(111), *J. Am. Chem. Soc.* **2018**, *140*, 1401–1408.

List of Figures

1.1	Schematic representation of intermolecular interactions. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	2
1.2	Illustration of a typical molecular CSP procedure.	6
3.1	Convergence of TS and MBD vdW lattice energies w.r.t. a dipole-dipole cut-off radius and a MBD supercell cut-off radius for a ND ₃ and HMB crystal. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	26
4.1	Illustration of the harmonic approximation. The actual potential curve of the total energy is illustrated in black, while the harmonic approximation around the equilibrium position (\mathbf{R}_{eq}) is shown in red.	30
4.2	Illustration of a one-dimensional quantum harmonic oscillator.	35
4.3	Illustration of a one-dimensional Morse oscillator.	36
4.4	Phonon dispersion of a one-dimensional infinite chain of atoms with a single atom in the unit cell.	37
4.5	Illustration of the phonon mode of a homogeneous infinite atomic chain at the $k = 0$ (a) and at $k = \pm\pi/b$ (b). The dashed lines indicate the length of the unit cell.	38
4.6	Phonon dispersion of a one-dimensional infinite chain of atoms with two atoms in the unit cell. The acoustic and optical mode is shown in red and blue, respectively.	39
4.7	Illustration of the acoustic (a) and optical (b) modes in a diatomic infinite chain at $k = 0$	39
4.8	Illustration of the acoustic (a) and optical (b) modes in a diatomic infinite chain at $k = \pm\pi/b$	40

4.9	Vibrational modes of an ammonia molecule.	41
5.1	Unit cell of a cubic ammonia crystal.	45
5.2	The temperature-dependent factor in the second integral of equation 5.5 vs. ω . Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	50
5.3	Full pDOS of ND ₃ (a) and the low-frequency region of the ND ₃ pDOS determined with PBE, PBE+TS and PBE+MBD. The gray lines in (a) show the location of the experimental internal modes as determined by Holt <i>et al.</i> [125] and in (b) they mark experimentally observed IR (61 K) and Raman (18 K) frequencies [126]. Adapted with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	52
5.4	Low-frequency pDOS for HMB. The gray lines mark the peak maxima of the experimental INS spectrum measured at a temperature of 15 K by Ciezak <i>et al.</i> [127]. Adapted with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	52
5.5	Harmonicity of two phonon modes in ND ₃ calculated with PBE+MBD. The displacements are shown at a relative scale in which a displacement of 1 means that the norm of the eigenvectors is equal to one. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	55
5.6	Phonon density of states of solid ND ₃ calculated with PBE at several unit-cell volumes. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	56
5.7	Quasi-harmonic approximation for solid ND ₃ calculated with different methods at three different temperatures. The solid lines represent Mur-naghan EOS fits and the red triangles mark the corresponding minima. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	58
5.8	Heat capacity at constant pressure (a) and linear thermal expansion coef-ficient (b) of solid ND ₃ obtained from the QHA compared to experimental values from Ref. 149. Reproduced with permission from Ref. 13. Copy-right 2016 John Wiley & Sons, Ltd.	60
5.9	Spherical plots of the Young's modulus of ND ₃ (in GPa) obtained at the minima of E_{tot} and at volumes corresponding to 194 K from the QHA for PBE and PBE+MBD, as well as experimental values at 194 K of Ref. 152. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	62

6.1	The unit cell of the studied orthorhombic purine crystal.	65
6.2	Obtained PBE+TS low-frequency THz/IR (left) and Raman (right) spectra for the fully optimized structure within the harmonic approximation (Opt./HA), for the thermally expanded unit cell within the harmonic approximation (QHA/HA), and for the thermally expanded unit cell using Morse oscillators (QHA/Morse) compared with experimental measurements (Exp.) [136]. The individual spectra are offset for clarity.	68
6.3	Obtained low-frequency THz/IR (left) and Raman (right) spectra for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t compared with experimental measurements (Exp.) [136]. All calculations were performed at a unit cell volume of 532.5 \AA^3 and Morse oscillators were used throughout. The individual spectra are offset for clarity.	70
7.1	Unit cell of aspirin form I.	74
7.2	Measured linear absorption of the aspirin sample at $T = 80 \text{ K}$ (a) and the amplitude spectra of applied pulses A and B (b). Data from Ref. 166.	74
7.3	THz spectra for the optimized structure of aspirin form I and for all studied modulated structures. The shown spectra were normalized to the largest observed intensity within the plotted spectral range and offset for clarity. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.	76
7.4	Lowest-frequency IR-active phonon modes of the optimized aspirin structure located at 1.08 THz (a), 1.31 THz (b), 1.66 THz (c), 1.89 THz (d), 2.06 THz (e), and 2.24 THz (f). The mass-weighted displacement vectors (blue) represent the direction and the relative magnitude of the atomic motion. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.	77
7.5	Representative phonon mode within the blue-shifted peak initially located at 1.3 THz for the optimized structure and all modulated structures: a) optimized structure (1.31 THz); b) $A_{\text{max}} = 0.1 \text{ \AA}\sqrt{\text{amu}}$ (1.33 THz); c) $A_{\text{max}} = 0.5 \text{ \AA}\sqrt{\text{amu}}$ (1.38 THz); d) $A_{\text{max}} = 1.0 \text{ \AA}\sqrt{\text{amu}}$ (1.43 THz); e) $A_{\text{max}} = 2.0 \text{ \AA}\sqrt{\text{amu}}$ (1.49 THz); f) $A_{\text{max}} = 4.0 \text{ \AA}\sqrt{\text{amu}}$ (1.67 THz). The mass-weighted displacement vectors (blue) represent the direction and the relative magnitude of the atomic motion. For systems in which more than one mode contribute to this peak, one representative mode is shown. Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.	78

7.6	Overlay of the unit cell of the optimized structure (gray) with the distorted structures of a maximum displacement amplitude of 0.1 (black), 0.5 (orange), 1 (red), 2 (blue), and $4 \text{ \AA}\sqrt{\text{amu}}$ (green). Reprinted with permission from Ref. 166. Copyright 2017 by the American Physical Society.	79
8.1	The five target systems of the sixth CCDC blind test.	84
8.2	Illustration of our complete re-ranking procedure (applied for systems XXII and XXIII).	85
8.3	Impact of the applied geometry and lattice optimizations on PBE+TS energies used for the initial screening of structures.	86
8.4	Results of our blind-test submission. Experimentally observed structures are shown in red while other structures are shown in gray. The first or only result of each system shows the PBE+MBD ranking while the second lists (L2) include in addition also vibrational free energies at room temperature. Note that the results of system XXVI were obtained after the blind test submission deadline.	88
8.5	Relative stability rankings for system XXIII. Experimentally observed structures are shown in color. Reproduced from Ref. 51 (CC-BY).	89
9.1	Illustration of the applied hierarchical CSP procedure.	94
9.2	Relative stabilities for all steps in the present CSP stability ranking procedure for systems XXII, XXIV, XXV, and XXVI. For each ranking, the energy of the most stable crystal structure defines the zero-of-the-energy. Experimentally observed structures are highlighted in color while all other structures are in gray. The final ranking for each system corresponds to the Helmholtz free energies at the PBE0+MBD+ F_{vib} level, computed at 150 K (XXII), 240 K (XXIV), or 300 K (XXV, XXVI). The energies are normalized per molecule for XXII and XXVI and given per formula unit for XXIV and XXV. Adapted from Ref. 185 (CC BY-NC).	97
9.3	Relative stabilities for all steps in the present CSP stability ranking procedure for system XXIII. For each ranking, the energy of the most stable crystal structure defines the zero-of-the-energy. Experimentally observed structures are highlighted in color while all other structures are in gray. The final ranking for each system corresponds to the Helmholtz free energies at the PBE0+MBD+ F_{vib} level, computed at the 300 K. The energies are normalized per molecule. The right part of the figure shows the unit cells for all highlighted structures. Adapted from Ref. 185 (CC BY-NC).	98

9.4	Overlay between the experimentally determined structures (element-specific colors) and the corresponding PBE+TS optimized structures for systems (green): (a) XXII, (b) XXIII-A, (c) XXIII-B, (d) XXIII-C, (e) XXIII-D, (f) XXIII-E, (g) XXIV, (h) XXV, and (i) XXVI. These overlays are shown for the molecules constituting the respective unit cell. Reproduced from Ref. 185 (CC BY-NC).	100
9.5	Structural differences between structure N70 (a) and form A (b) of system XXIII. Both structures share a sheet structure (red boxes) in which the molecules are arranged according to the same pattern, but the sheets are stacked differently.	102
9.6	Stability rankings for thermally-expanded structures. Energetic rankings for all experimentally observed (Form A, B, C, D, E) and theoretically-predicted (N18, N31, N42, N70) structures for system XXIII. All energies were evaluated using thermally-expanded PBE+TS structures optimized at 300 K with the QHA. The last two rankings include harmonic (F_{vib}) and Morse anharmonic (\tilde{F}_{vib}) vibrational free energy contributions. Reproduced from Ref. 185 (CC BY-NC).	104
10.1	Visualization of the unit cell for all explicitly discussed structures. The coordinates of all optimized structures are available in the Supporting Information of Ref. 195. Adapted from Ref. 195 with permission from The Royal Society of Chemistry.	108
10.2	Correlation between PBE+TS and PBE+MBD energies calculated on top of PBE+TS-optimized structures[185]. The most stable PBE+MBD structure is always set to zero and energies are normalized per chemical unit. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	110
10.3	Box plots of all possible absolute deviations in relative energies for system XXIII.	111
10.4	Relative PBE+TS (a) and PBE+MBD (b) energies for several structures of system XXIII. The energy of form C serves as reference and is set to zero. Part (b) shows the relative PBE+MBD energies when MBD interactions are included up to the shown n -atom contribution. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. . . .	112
10.5	Relative PBE0+MBD energies of the five experimental polymorphs of system XXIII for different optimization methods. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	114

-
- 10.6 Low frequency part of the phonon density of states (pDOS) for three structures of system XXIII (left) and the corresponding cumulative relative vibrational free energies w.r.t the wave number (right). The vibrational free energy of structure N70 is always set to zero and the vibrational free energy is evaluated based on the pDOS up to the shown wave number. The right plot shows the low-frequency part up to 300 wave numbers. The final relative vibrational free energies after taking into account all modes are indicated with the three lines outside of the right plot. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. . . . 116
- 10.7 Convergence of vibrational free energies. Part (a) shows the relative stabilities of 9 structures from system XXIII calculated with PBE0+MBD+ F_{vib} (PBE+TS) at 300 K, where supercells with cell lengths of at least 10 Å were used for the finite difference calculation. The vibrational free energy was evaluated using only the Γ point modes of the unit cell (uc), the Γ -point modes of the used supercell (sc), or with a q -point mesh with the shown values used for x (see Methods). In part (b) a converged q -point mesh is used throughout but the finite difference calculations were performed by either using just the unit cell or supercells with minimum cell lengths of 10 and 14 Å, respectively. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. 117
- 10.8 Low-frequency phonon band structure plot along several high-symmetry lines for structure XXIII-N18 calculated using the unit cell (a), a supercell with a minimum length of 10 Å (b), and a supercell with a minimum length of 14 Å (c) in every direction. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. 118
- 10.9 Comparison of PBE+TS and PBE+MBD vibrational free energies for several structures of systems XXII and XXIII. The relative stability consists of the PBE0+MBD energy plus the vibrational free energy evaluated with the respective method. The vibrational free energies were calculated at the temperature of the respective experimental crystal structure measurements (150 K for XXII and 300 K for XXIII). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. 119
- 10.10 Illustration of the QHA thermal expansion for form A of system XXIII as described by PBE+TS (a) and PBE+MBD (b). The dots correspond to the harmonic free energies calculated at various volumes and at three different temperatures. The solid curves correspond to the Murnaghan equation-of-state fit, and the respective minimum is always marked with the red triangle. The light blue line indicates the volume of the fully optimized structure. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry. 122

10.11	Relative stabilities for 9 structures of system XXIII for optimized and thermally expanded structures. The relative stability consists of the PBE0+MBD total energy plus the vibrational free energy evaluated with the respective method. The vibrational free energy was calculated in the harmonic approximation at 300 K. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	123
10.12	Relative stabilities for 9 structures of system XXIII for optimized and thermally expanded structures. The relative stability consists of the PBE0+MBD total energy plus the vibrational free energy evaluated with the respective method. The vibrational free energy was calculated using Morse oscillators at 300 K. Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	124
11.1	Illustration of a coumarin molecule and the unit cells of the five coumarin polymorphs.	127
11.2	PBE+TS and PBE+MBD stability vs. volume for 50 vdW-DF2-optimized structures of coumarin.	129
11.3	Relative Helmholtz free energies vs. temperature for the five coumarin polymorphs. The shown free energies are the sum of the PBE0+MBD static energy and the vibrational free energy calculated with PBE+MBD/light. The energy of form I defines at every temperature the zero of the energy [60].	131
11.4	Low-frequency phonon density of states for coumarin forms I and V according to PBE+MBD.	132

List of Tables

5.1	Unit-cell volumes for ND ₃ phase I for several methods calculated via optimization (opt), extracted from a quasi-harmonic approximation (n K) and optimization under thermal pressure (p_{th}). All volumes are given in Å ³ while p_{th} is in GPa. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	57
5.2	Elastic constants, bulk modulus, and anisotropy ratio for solid ND ₃ calculated at the minima of E_{tot} and at volumes corresponding to 194 K according to respective QHAs, compared to experimental values. Reproduced with permission from Ref. 13. Copyright 2016 John Wiley & Sons, Ltd.	61
6.1	Obtained IR and Raman active PBE+TS phonon frequencies for the fully optimized structure within the harmonic approximation (Opt./HA), for the thermally expanded unit cell within the harmonic approximation (QHA/HA), and for the thermally expanded unit cell using Morse oscillators (QHA/Morse) compared with experimental measurements [136]. Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given.	68
6.2	Obtained IR and Raman active PBE+TS phonon frequencies for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t within the harmonic approximation compared with experimental measurements [136]. All calculations were performed at a thermally expanded unit cell volume of 532.5 Å ³ . Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given.	70
6.3	Obtained IR and Raman active PBE+TS phonon frequencies for PBE+TS/l, PBE+MBD/l, PBE0+MBD/l, and PBE+MBD/t using Morse oscillators compared with experimental measurements [136]. All calculations were performed at a thermally expanded unit cell volume of 532.5 Å ³ . Furthermore, the mean absolute error (MAE) and the maximum absolute error (MAX) w.r.t. experimental measurements are given..	71

8.1	Blind test results of selected groups [51]. The numbers indicate the rank of the experimentally obtained structure within the submitted lists.	87
9.1	Convergence of relative stabilities with basis set and grid settings. This table shows the mean absolute deviation (MAD) and the maximum absolute deviation (MAX) for all possible relative energies (within a system) from Table A.1 w.r.t. PBE+MBD calculations with <i>really tight</i> settings for basis set and grids [185].	93
9.2	This table shows the mean absolute deviation (MAD) and the maximum absolute deviation (MAX) for all possible relative energies (within a system) from Table A.1 w.r.t. PBE0+MBD calculations with <i>tight</i> settings for basis set and grids. The method PBE0+MBD without a label for basis set and grid settings corresponds to the described PBE0+MBD estimate [185].	93
10.1	Statistics for the impact of several interactions on all possible relative energies expressed in terms of the mean absolute deviation (MAD) in kJ/mol and the maximum absolute deviation (MAX) in kJ/mol for all studied systems [185]; PBE+MBD vs. PBE+TS (MBD), PBE0+MBD vs. PBE+MBD (PBE0), PBE0+MBD+ F_{vib} vs. PBE0+MBD (F_{vib}). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	111
10.2	Errors of optimized structures w.r.t. experimental measurements [51]; mean relative deviation (MRD), mean absolute relative deviation (MARD) as well as maximum absolute deviation (MAX) in percent. The labels l and t indicate light and tight species default settings within FHI-aims, respectively. For the cell angles we only take into account values which are not exactly 90 degrees due to symmetry. Adapted from Ref. 195 with permission from The Royal Society of Chemistry.	113
10.3	Relative free energies in kJ/mol. The total energy was always calculated with PBE0+MBD and the structure optimization as well as the vibrational free energy was performed with the labeled method. The vibrational free energies were always calculated at the experimental measuring temperature of the respective crystal structures (150 K for XXII and 300 K for XXIII). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	120
10.4	Relative computation times for one single energy and force calculation of structure XXIII-A (172 atoms in the unit cell). All computation times were normalized to the PBE+TS/ l energy calculation, which amounts to 2.1 CPU hours on Intel Xeon E5-2680 v4 cores (2.4 GHz). Reproduced from Ref. 195 with permission from The Royal Society of Chemistry.	121

11.1	Relative stabilities in kJ/mol calculated using the PBE+MBD/light-optimized structures with different methods compared to the experimental estimates around 340 K (Exp.) [60]. The Helmholtz free energies F at 300 and 340 K are the sum of the PBE0+MBD static energy and the vibrational free energy calculated with PBE+MBD/light.	130
A.1	Relative energies of a benchmark set of structures with small unit cells calculated with different methods. The benchmark set consists of the single $Z = 1$ structure and the top 7 $Z = 2$ structures after the PBE0+MBD ranking for system XXII, and the top 4 $Z = 2$ structures after the PBE0+MBD ranking for system XXIV. The energy of the most stable structure of a system after the PBE0+MBD ranking was set to zero for all methods. The letters after the slash indicate the used basis set and grid settings: l corresponds to the <i>light</i> species default settings for basis sets and grids in FHI-aims, while t refers to the <i>tight</i> species default settings. The label rt indicates the usage of the <i>really tight</i> grid settings together with tier-3 basis functions. All relative energies are in kJ/mol per formula unit. The method PBE0+MBD without a label for basis set and grid settings corresponds to the PBE0+MBD estimate introduced in Chapter 9 and used for the stability rankings in Chapters 9, 10, and 11 [185].	139
B.1	Errors of PBE+TS-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c) , angles (α, β, γ) , cell volume, and density. In addition the RMSD ₂₀ is shown in Å [185].	141
B.2	Errors of PBE+MBD-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c) , angles (α, β, γ) , cell volume, and density. In addition the RMSD ₂₀ is shown in Å [195].	142
B.3	Errors of PBE0+MBD-optimized structures using light settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c) , angles (α, β, γ) , cell volume, and density. In addition the RMSD ₂₀ is shown in Å [195].	142
B.4	Errors of PBE+MBD-optimized structures using tight settings w.r.t. experimental structures. This table shows the relative error in % for cell lengths (a, b, c) , angles (α, β, γ) , cell volume, and density. In addition the RMSD ₂₀ is shown in Å [195].	142

B.5	Errors of the thermally-expanded structures (corresponding to 300 K within the QHA)calculated with PBE+TS and light settings w.r.t. experimental structures ³⁴ . This table shows the relative error in % for cell lengths (a, b, c), angles (α, β, γ), cell volume, and density. In addition the RMSD ₂₀ is shown in Å [185].	143
C.1	Stability ranking for system XXII in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 150 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm ³ [185].	145
C.2	Stability ranking for system XXIII in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the asymmetric unit (Z'), the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm ³ [185].	148
C.2	(continued) Stability ranking for system XXIII.	149
C.2	(continued) Stability ranking for system XXIII.	150
C.3	Stability ranking for system XXIV in kJ/mol normalized per formula unit. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 240 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm ³ [185].	151
C.3	(continued) Stability ranking for system XXIV.	152
C.3	(continued) Stability ranking for system XXIV.	153
C.4	Stability ranking for system XXV in kJ/mol normalized per formula unit. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice energies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm ³ [185].	154

C.4 (continued) Stability ranking for system XXV.	155
C.4 (continued) Stability ranking for system XXV.	156
C.5 Stability ranking for system XXVI in kJ/mol normalized per molecule. The energy of the most stable structure in each ranking was set to zero and the structures are always ordered according to the highest available ranking level. The final ranking (F) includes PBE0+MBD lattice ener- gies and vibrational free energies calculated at 300 K using PBE+TS. In addition, this table contains the number of molecules in the unit cell (Z), the space group (Symm.), and the density in g/cm ³ [185].	157
C.5 (continued) Stability ranking for system XXVI.	158
C.5 (continued) Stability ranking for system XXVI.	159

Bibliography

- [1] J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter, J. Morris, Ritonavir: An Extraordinary Example of Conformational Polymorphism, *Pharm. Res.* **2001**, *18*, 859–866.
- [2] A. G. Davies, A. D. Burnett, W. Fan, E. H. Linfield, J. E. Cunningham, Terahertz Spectroscopy of Explosives and Drugs, *Mater. Today* **2008**, *11*, 18–26.
- [3] C. D. Dimitrakopoulos, P. R. L. Malenfant, Organic Thin Film Transistors for Large Area Electronics, *Adv. Mater.* **2002**, *14*, 99–117.
- [4] M. D. King, W. D. Buchanan, T. M. Korter, Understanding the Terahertz Spectra of Crystalline Pharmaceuticals: Terahertz Spectroscopy and Solid-State Density Functional Theory Study of (S)-(+)-Ibuprofen and (RS)-Ibuprofen, *J. Pharm. Sci.* **2011**, *100*, 1116–1129.
- [5] J. E. Anthony, Functionalized Acenes and Heteroacenes for Organic Electronics, *Chem. Rev.* **2006**, *106*, 5028–5048.
- [6] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, J.-L. Brédas, Charge Transport in Organic Semiconductors, *Chem. Rev.* **2007**, *107*, 926–952.
- [7] J. E. Anthony, The Larger Acenes: Versatile Organic Semiconductors, *Angew. Chem. Int. Ed.* **2008**, *47*, 452–483.
- [8] G. Kaupp, Solid-State Reactions, Dynamics in Molecular Crystals, *Curr. Opin. Solid State Mater. Sci.* **2002**, *6*, 131–138.
- [9] S. Datta, D. J. W. Grant, Crystal Structures of Drugs: Advances in Determination, Prediction and Engineering, *Nat. Rev. Drug Discov.* **2004**, *3*, 42–57.
- [10] G. R. Desiraju, Crystal Engineering: From Molecule to Crystal, *J. Am. Chem. Soc.* **2013**, *135*, 9952–9967.
- [11] A. Stone, *The Theory of Intermolecular Forces*, Second Edition, Oxford Press, Oxford, **2013**.
- [12] B. Jeziorski, R. Moszynski, K. Szalewicz, Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes, *Chem. Rev.* **1994**, *94*, 1887–1930.

- [13] J. Hoja, A. M. Reilly, A. Tkatchenko, First-Principles Modeling of Molecular Crystals: Structures and Stabilities, Temperature and Pressure, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1294.
- [14] A. Szabo, N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, of *Dover Books on Chemistry*, Dover Publications, Mineola, New York, **1989**.
- [15] N. Ferri, R. A. DiStasio Jr., A. Ambrosetti, R. Car, A. Tkatchenko, Electronic Properties of Molecules and Surfaces with a Self-Consistent Interatomic van der Waals Density Functional, *Phys. Rev. Lett.* **2015**, *114*, 176802.
- [16] P. Cieplak, J. Caldwell, P. Kollman, Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases, *J. Comput. Chem.* **2001**, *22*, 1048–1057.
- [17] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, P. Ren, Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins, *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.
- [18] A. Warshel, M. Kato, A. V. Pisliakov, Polarizable Force Fields: History, Test Cases, and Prospects, *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.
- [19] J. F. Dobson, T. Gould, Calculation of Dispersion Energies, *J. Phys. Condens. Matter* **2012**, *24*, 073201.
- [20] G. A. Jeffrey, An Introduction to Hydrogen Bonding, Oxford University Press, New York, **1997**.
- [21] J. Hoja, A. F. Sax, K. Szalewicz, Is Electrostatics Sufficient to Describe Hydrogen-Bonding Interactions?, *Chem. Eur. J.* **2014**, *20*, 2292–2300.
- [22] A. J. Stone, Are Halogen Bonded Structures Electrostatically Driven?, *J. Am. Chem. Soc.* **2013**, *135*, 7005–7009.
- [23] K. E. Riley, P. Hobza, The Relative Roles of Electrostatics and Dispersion in the Stabilization of Halogen Bonds, *Phys. Chem. Chem. Phys.* **2013**, *15*, 17742–17751.
- [24] K. Szalewicz, Symmetry-Adapted Perturbation Theory of Intermolecular Forces, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 254–272.
- [25] G. Jansen, Symmetry-Adapted Perturbation Theory Based on Density Functional Theory for Noncovalent Interactions, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 127–144.
- [26] R. J. Bartlett, Coupled-Cluster Theory and Its Equation-Of-Motion Extensions, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 126–138.

-
- [27] M. Dubecký, P. Jurečka, R. Derian, P. Hobza, M. Otyepka, L. Mitas, Quantum Monte Carlo Methods Describe Noncovalent Interactions with Subchemical Accuracy, *J. Chem. Theory Comput.* **2013**, *9*, 4287–4292.
- [28] P. Jurečka, J. Šponer, J. Černý, P. Hobza, Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- [29] J. Řezáč, K. E. Riley, P. Hobza, S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures, *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.
- [30] A. Ambrosetti, D. Alfè, R. A. DiStasio Jr., A. Tkatchenko, Hard Numbers for Large Molecules: Toward Exact Energetics for Supramolecular Systems, *J. Phys. Chem. Lett.* **2014**, *5*, 849–855.
- [31] S. Grimme, Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction, *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- [32] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A Consistent and Accurate ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu, *J. Chem. Phys.* **2010**, *132*, 154104.
- [33] A. D. Becke, E. R. Johnson, Exchange-Hole Dipole Moment and the Dispersion Interaction Revisited, *J. Chem. Phys.* **2007**, *127*, 154108.
- [34] A. Tkatchenko, M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.* **2009**, *102*, 073005.
- [35] A. Tkatchenko, R. A. DiStasio Jr., R. Car, M. Scheffler, Accurate and Efficient Method for Many-Body van der Waals Interactions, *Phys. Rev. Lett.* **2012**, *108*, 236402.
- [36] S. Grimme, Density Functional Theory with London Dispersion Corrections, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228.
- [37] S. Grimme, A. Hansen, J. G. Brandenburg, C. Bannwarth, Dispersion-Corrected Mean-Field Electronic Structure Methods, *Chem. Rev.* **2016**, *116*, 5105–5154.
- [38] A. D. Boese, Density Functional Theory and Hydrogen Bonds: Are We There Yet?, *Chem. Phys. Chem.* **2015**, *16*, 978–985.
- [39] A. Otero-de-la Roza, E. R. Johnson, A Benchmark for Non-Covalent Interactions in Solids, *J. Chem. Phys.* **2012**, *137*, 054103.
- [40] A. M. Reilly, A. Tkatchenko, Understanding the Role of Vibrations, Exact Exchange, and Many-Body van der Waals Interactions in the Cohesive Properties of Molecular Crystals, *J. Chem. Phys.* **2013**, *139*, 024705.

- [41] A. J. Cruz-Cabeza, S. M. Reutzel-Edens, J. Bernstein, Facts and Fictions about Polymorphism, *Chem. Soc. Rev.* **2015**, *44*, 8619–8635.
- [42] A. M. Reilly, A. Tkatchenko, Role of Dispersion Interactions in the Polymorphism and Entropic Stabilization of the Aspirin Crystal, *Phys. Rev. Lett.* **2014**, *113*, 055701.
- [43] J. Nyman, G. M. Day, Static and Lattice Vibrational Energy Differences Between Polymorphs, *CrystEngComm* **2015**, *17*, 5154–5165.
- [44] R. K. Hylton, G. J. Tizzard, T. L. Threlfall, A. L. Ellis, S. J. Coles, C. C. Seaton, E. Schulze, H. Lorenz, A. Seidel-Morgenstern, M. Stein, S. L. Price, Are the Crystal Structures of Enantiopure and Racemic Mandelic Acids Determined by Kinetics or Thermodynamics?, *J. Am. Chem. Soc.* **2015**, *137*, 11095–11104.
- [45] S. L. Price, Why Don't We Find More Polymorphs?, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2013**, *69*, 313–328.
- [46] J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, D. E. Williams, A Test of Crystal Structure Prediction of Small Organic Molecules, *Acta Crystallogr. Sect. B Struct. Sci.* **2000**, *56*, 697–714.
- [47] W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, D. E. Williams, Crystal Structure Prediction of Small Organic Molecules: A Second Blind Test, *Acta Crystallogr. Sect. B Struct. Sci.* **2002**, *58*, 647–661.
- [48] G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, P. Verwer, A Third Blind Test of Crystal Structure Prediction, *Acta Crystallogr. Sect. B Struct. Sci.* **2005**, *61*, 511–527.
- [49] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf, B. Schweizer, Significant Progress in Predicting the Crystal Structures of Small Organic Molecules – A Report on the Fourth Blind Test, *Acta Crystallogr. Sect. B Struct. Sci.* **2009**, *65*, 107–125.

-
- [50] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. a. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti, I. K. Zhitkov, Towards Crystal Structure Prediction of Complex Organic Compounds – A Report on the Fifth Blind Test, *Acta Crystallogr. Sect. B Struct. Sci.* **2011**, *67*, 535–551.
- [51] A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meeke, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu, C. R. Groom, Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods, *Acta Crystallogr. Sect. B: Struct. Sci.* **2016**, *72*, 439–459.
- [52] D. K. Bučar, R. W. Lancaster, J. Bernstein, Disappearing Polymorphs Revisited, *Angew. Chem. Int. Ed.* **2015**, *54*, 6972–6993.
- [53] S. L. Price, Predicting Crystal Structures of Organic Compounds, *Chem. Soc. Rev.* **2014**, *43*, 2098–2111.
- [54] G. J. O. Beran, Modeling Polymorphic Molecular Crystals with Electronic Structure Theory, *Chem. Rev.* **2016**, *116*, 5567–5613.
- [55] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, G. K.-L. Chan, Ab Initio Determination of the Crystalline Benzene Lattice Energy to Sub-Kilojoule/Mole Accuracy, *Science* **2014**, *345*, 640–643.
- [56] J. Nyman, O. S. Pundyke, G. M. Day, Accurate Force Fields and Methods for Modelling Organic Molecular Crystals at Finite Temperatures, *Phys. Chem. Chem. Phys.* **2016**, *18*, 15828–15837.

- [57] E. Schneider, L. Vogt, M. E. Tuckerman, Exploring Polymorphism of Benzene and Naphthalene with Free Energy Based Enhanced Molecular Dynamics, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 542–550.
- [58] J. G. Brandenburg, S. Grimme, Organic Crystal Polymorphism: A Benchmark for Dispersion-Corrected Mean-Field Electronic Structure Methods, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 502–513.
- [59] S. L. Price, J. G. Brandenburg in *Non-covalent Interactions in Quantum Chemistry and Physics*, A. O. de la Roza, G. A. DiLabio (Eds.), Elsevier, **2017**, pp. 333–363.
- [60] A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, J. Hoja, E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman, B. Kahr, Powder Diffraction and Crystal Structure Prediction Identify Four New Coumarin Polymorphs, *Chem. Sci.* **2017**, *8*, 4926–4940.
- [61] M. U. Schmidt, J. Brüning, J. Glinnemann, M. W. Hützler, P. Mörschel, S. N. Ivashetskaya, J. van de Streek, D. Braga, L. Maini, M. R. Chierotti, R. Gobetto, The Thermodynamically Stable Form of Solid Barbituric Acid: The Enol Tautomer, *Angew. Chem. Int. Ed.* **2011**, *50*, 7924–7926.
- [62] H. P. G. Thompson, G. M. Day, Which Conformations Make Stable Crystal Structures? Mapping Crystalline Molecular Geometries to the Conformational Energy Landscape, *Chem. Sci.* **2014**, *5*, 3173–3182.
- [63] A. J. Cruz-Cabeza, J. Bernstein, Conformational Polymorphism, *Chem. Rev.* **2014**, *114*, 2170–2191.
- [64] A. Ambrosetti, A. M. Reilly, R. A. DiStasio Jr., A. Tkatchenko, Long-Range Correlation Energy Calculated from Coupled Atomic Response Functions, *J. Chem. Phys.* **2014**, *140*, 18A508.
- [65] D. A. McQuarrie, *Quantum Chemistry*, Second Edition, University Science Books, Sausalito, California, **2008**.
- [66] L. Piela, *Ideas of Quantum Chemistry*, Second Edition, Elsevier, San Diego, **2014**.
- [67] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *384*, 361–376.
- [68] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *384*, 489–527.
- [69] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *385*, 437–490.
- [70] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *386*, 109–139.
- [71] P. A. M. Dirac, R. H. Fowler, The Quantum Theory of the Electron, *Proc. R. Soc. Lond. A. Math. Phys. Sci.* **1928**, *117*, 610–624.

-
- [72] M. Born, R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. Phys.* **1927**, *389*, 457–484.
- [73] W. Pauli, Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren, *Zeitschrift für Physik* **1925**, *31*, 765–783.
- [74] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, Ab Initio Molecular Simulations with Numeric Atom-Centered Orbitals, *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.
- [75] D. Cremer, From Configuration Interaction to Coupled Cluster Theory: The Quadratic Configuration Interaction Approach, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 482–503.
- [76] C. Møller, M. S. Plesset, Note on an Approximation Treatment for Many-Electron Systems, *Phys. Rev.* **1934**, *46*, 618–622.
- [77] J. Čížek, On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods, *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- [78] R. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules, of *International Series of Monographs on Chemistry*, Oxford University Press, New York, **1994**.
- [79] A. D. Becke, Perspective: Fifty Years of Density-Functional Theory in Chemical Physics, *J. Chem. Phys.* **2014**, *140*, 18A301.
- [80] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.* **1964**, *136*, B864–B871.
- [81] W. Kohn, L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [82] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [83] S. F. Sousa, P. A. Fernandes, M. J. Ramos, General Performance of Density Functionals, *J. Phys. Chem. A* **2007**, *111*, 10439–10452.
- [84] C. Adamo, V. Barone, Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- [85] S. Grimme, S. Ehrlich, L. Goerigk, Effect of the Damping Function in Dispersion Corrected Density Functional Theory, *J. Comp. Chem.* **2011**, *32*, 1456–1465.
- [86] A. D. Becke, E. R. Johnson, Exchange-Hole Dipole Moment and the Dispersion Interaction: High-Order Dispersion Coefficients, *J. Chem. Phys.* **2006**, *124*, 014104.
- [87] E. R. Johnson, Dependence of Dispersion Coefficients on Atomic Environment, *J. Chem. Phys.* **2011**, *135*, 234109.

- [88] A. M. Reilly, A. Tkatchenko, van der Waals Dispersion Interactions in Molecular Materials: Beyond Pairwise Additivity, *Chem. Sci.* **2015**, *6*, 3289–3301.
- [89] M. A. Blood-Forsythe, T. Markovich, R. A. DiStasio Jr., R. Car, A. Aspuru-Guzik, Analytical Nuclear Gradients for the Range-Separated Many-Body Dispersion Model of Noncovalent Interactions, *Chem. Sci.* **2016**, *7*, 1712–1728.
- [90] F. L. Hirshfeld, Bonded-Atom Fragments for Describing Molecular Charge Densities, *Theor. Chim. Acta* **1977**, *44*, 129–138.
- [91] Q. Wu, W. Yang, Empirical Correction to Density Functional Theory for van der Waals Interactions, *J. Chem. Phys.* **2002**, *116*, 515–524.
- [92] P. Jurečka, J. Černý, P. Hobza, D. R. Salahub, Density Functional Theory Augmented with an Empirical Dispersion Term. Interaction Energies and Geometries of 80 Noncovalent Complexes Compared with ab Initio Quantum Mechanics Calculations, *J. Comput. Chem.* **2007**, *28*, 555–569.
- [93] D. C. Langreth, J. P. Perdew, Exchange-Correlation Energy of a Metallic Surface: Wave-Vector Analysis, *Phys. Rev. B* **1977**, *15*, 2884–2901.
- [94] O. Gunnarsson, B. I. Lundqvist, Exchange and Correlation in Atoms, Molecules, and Solids by the Spin-Density-Functional Formalism, *Phys. Rev. B* **1976**, *13*, 4274–4298.
- [95] A. Tkatchenko, A. Ambrosetti, R. A. DiStasio Jr., Interatomic Methods for the Dispersion Energy Derived from the Adiabatic Connection Fluctuation-Dissipation Theorem, *J. Chem. Phys.* **2013**, *138*, 074106.
- [96] J. Hermann, R. A. DiStasio Jr., A. Tkatchenko, First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications, *Chem. Rev.* **2017**, *117*, 4714–4758.
- [97] A. Tkatchenko, Current Understanding of Van der Waals Effects in Realistic Materials, *Adv. Func. Mat.* **2015**, *25*, 2054–2061.
- [98] A. Ambrosetti, N. Ferri, R. A. DiStasio Jr., A. Tkatchenko, Wavelike Charge Density Fluctuations and van der Waals Interactions at the Nanoscale, *Science* **2016**, *351*, 1171–1176.
- [99] V. V. Gobre, A. Tkatchenko, Scaling Laws for van der Waals Interactions in Nanostructured Materials, *Nat. Commun.* **2013**, *4*, 2341.
- [100] N. W. Ashcroft, N. D. Mermin, Solid State Physics, Saunders College, Philadelphia, **1976**.
- [101] P. Hofmann, Solid State Physics: An Introduction, of *Physics Textbook*, Wiley, **2008**.
- [102] C. Kittel, Introduction to Solid State Physics, Wiley, **1996**.

-
- [103] S. Baroni, S. de Gironcoli, A. Dal Corso, P. Giannozzi, Phonons and Related Crystal Properties from Density-Functional Perturbation Theory, *Rev. Mod. Phys.* **2001**, *73*, 515–562.
- [104] P. M. Morse, Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels, *Phys. Rev.* **1929**, *34*, 57–64.
- [105] J. P. Dahl, M. Springborg, The Morse Oscillator in Position Space, Momentum Space, and Phase Space, *J. Chem. Phys.* **1988**, *88*, 4535–4547.
- [106] A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, A. Michaelides, Fast and Accurate Quantum Monte Carlo for Molecular Crystals, *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 1724–1729.
- [107] K. D. Nanda, G. J. O. Beran, Prediction of Organic Molecular Crystal Geometries from MP2-Level Fragment Quantum Mechanical/Molecular Mechanical Calculations, *J. Chem. Phys.* **2012**, *137*, 174106.
- [108] S. Wen, G. J. O. Beran, Accurate Molecular Crystal Lattice Energies from a Fragment QM/MM Approach with On-the-Fly Ab Initio Force Field Parametrization, *J. Chem. Theory Comput.* **2011**, *7*, 3733–3742.
- [109] L. Maschio, B. Civalleri, P. Ugliengo, A. Gavezzotti, Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller-Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations, *J. Phys. Chem. A* **2011**, *115*, 11179–11186.
- [110] A. D. Becke, Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior, *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [111] C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density, *Phys. Rev. B* **1988**, *37*, 785–789.
- [112] A. D. Becke, Density-Functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [113] B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano, P. Ugliengo, B3LYP Augmented with an Empirical Dispersion Term (B3LYP-D*) as Applied to Molecular Crystals, *Cryst. Eng. Comm.* **2008**, *10*, 405–410.
- [114] J. Moellmann, S. Grimme, DFT-D3 Study of Some Molecular Crystals, *J. Phys. Chem. C* **2014**, *118*, 7615–7621.
- [115] R. S. Bradley, S. Cotson, 347. The Vapour Pressure and Lattice Energy of Hydrogen-Bonded Crystals. Part II. α - and β -Anhydrous Oxalic Acid and Tetragonal Pentaerythritol, *J. Chem. Soc.* **1953**, 1684–1688.
- [116] H. G. M. de Wit, J. A. Bouwstra, J. G. Blok, C. G. de Kruif, Vapor Pressures and Lattice Energies of Oxalic Acid, Mesotartaric Acid, Phloroglucinol, Myoinositol, and their Hydrates, *J. Chem. Phys.* **1983**, *78*, 1470–1475.

- [117] N. Marom, R. A. DiStasio Jr., V. Atalla, S. Levchenko, A. M. Reilly, J. R. CheLIKowsky, L. Leiserowitz, A. Tkatchenko, Many-Body Dispersion Interactions in Molecular Crystal Polymorphism, *Angew. Chem. Int. Ed.* **2013**, *52*, 6629–6632.
- [118] B. Schatschneider, S. Monaco, J.-J. Liang, A. Tkatchenko, High-Throughput Investigation of the Geometry and Electronic Structures of Gas-Phase and Crystalline Polycyclic Aromatic Hydrocarbons, *J. Phys. Chem. C* **2014**, *118*, 19964–19974.
- [119] N. D. Mermin, Thermal Properties of the Inhomogeneous Electron Gas, *Phys. Rev.* **1965**, *137*, A1441–A1443.
- [120] A. Marek, V. Blum, R. Johanni, V. Havu, B. Lang, T. Auckenthaler, A. Heinicke, H.-J. Bungartz, H. Lederer, The ELPA Library: Scalable Parallel Eigenvalue Solutions for Electronic Structure Theory and Computational Science, *J. Phys. Condens. Matter* **2014**, *26*, 213201.
- [121] T. Auckenthaler, V. Blum, H.-J. Bungartz, T. Huckle, R. Johanni, L. Krämer, B. Lang, H. Lederer, P. Willems, Parallel Solution of Partial Symmetric Eigenvalue Problems from Electronic Structure Calculations, *Parallel Comput.* **2011**, *37*, 783–794.
- [122] V. Havu, V. Blum, P. Havu, M. Scheffler, Efficient $O(N)$ Integration for All-Electron Electronic Structure Calculation Using Numeric Basis Functions, *J. Comput. Phys.* **2009**, *228*, 8367–8379.
- [123] A. Togo, I. Tanaka, First Principles Phonon Calculations in Materials Science, *Scr. Mater.* **2015**, *108*, 1–5.
- [124] S. A. Rivera, D. G. Allis, B. S. Hudson, Importance of Vibrational Zero-Point Energy Contribution to the Relative Polymorph Energies of Hydrogen-Bonded Species, *Cryst. Growth Des.* **2008**, *8*, 3905–3907.
- [125] J. S. Holt, D. Sadoskas, C. J. Pursell, Infrared Spectroscopy of the Solid Phases of Ammonia, *J. Chem. Phys.* **2004**, *120*, 7153.
- [126] O. Binbrek, A. Anderson, Raman Spectra of Molecular Crystals. Ammonia and 3-Deutero-Ammonia, *Chem. Phys. Lett.* **1972**, *15*, 421–427.
- [127] J. A. Ciezak, B. S. Hudson, Vibrational Analysis of the Inelastic Neutron Scattering Spectrum of Tetracyanoethylene-Hexamethylbenzene by Electronic Structure Calculations, *J. Mol. Struct. THEOCHEM* **2005**, *755*, 195–202.
- [128] C. Motta, S. Sanvito, Charge Transport Properties of Durene Crystals from First-Principles, *J. Chem. Theory Comput.* **2014**, *10*, 4624–4632.
- [129] M. Takahashi, Terahertz Vibrations and Hydrogen-Bonded Networks in Crystals, *Crystals* **2014**, *4*, 74–103.
- [130] A. I. McIntosh, B. Yang, S. M. Goldup, M. Watkinson, R. S. Donnan, Terahertz Spectroscopy: A Powerful New Tool for the Chemical Sciences?, *Chem. Soc. Rev.* **2012**, *41*, 2072–2082.

-
- [131] A. M. Reilly, D. S. Middlemiss, M. M. Siddick, D. A. Wann, G. J. Ackland, C. C. Wilson, D. W. H. Rankin, C. A. Morrison, The Phonon Spectrum of Phase-I Ammonia: Reassignment of Lattice Mode Symmetries from Combined Molecular and Lattice Dynamics Calculations, *J. Phys. Chem. A* **2008**, *112*, 1322–1329.
- [132] N. A. Spaldin, A Beginner’s Guide to the Modern Theory of Polarization, *J. Solid State Chem.* **2012**, *195*, 2 – 10.
- [133] D. G. Allis, D. A. Prokhorova, T. M. Korter, Solid-State Modeling of the Terahertz Spectrum of the High Explosive HMX, *J. Phys. Chem. A* **2006**, *110*, 1951–1959.
- [134] M. D. King, W. D. Buchanan, T. M. Korter, Application of London-Type Dispersion Corrections to the Solid-State Density Functional Theory Simulation of the Terahertz Spectra of Crystalline Pharmaceuticals, *Phys. Chem. Chem. Phys.* **2011**, *13*, 4250–4259.
- [135] M. D. King, T. M. Korter, Modified Corrections for London Forces in Solid-State Density Functional Theory Calculations of Structure and Lattice Dynamics of Molecular Crystals, *J. Phys. Chem. A* **2012**, *116*, 6927–6934.
- [136] M. T. Ruggiero, J. A. Zeitler, A. Erba, Intermolecular Anharmonicity in Molecular Crystals: Interplay Between Experimental Low-Frequency Dynamics and Quantum Quasi-Harmonic Simulations of Solid Purine, *Chem. Commun.* **2017**, *53*, 3781–3784.
- [137] M. Takahashi, Y. Ishikawa, Terahertz Vibrations of Crystalline α -D-Glucose and the Spectral Change in Mutual Transitions Between the Anhydride and Monohydrate, *Chem. Phys. Lett.* **2015**, *642*, 29–34.
- [138] Z.-P. Zheng, W.-H. Fan, H. Yan, Terahertz Absorption Spectra of Benzene-1,2-Diol, Benzene-1,3-Diol and Benzene-1,4-Diol, *Chem. Phys. Lett.* **2012**, *525-526*, 140–143.
- [139] M. D. King, W. D. Buchanan, T. M. Korter, Identification and Quantification of Polymorphism in the Pharmaceutical Compound Diclofenac Acid by Terahertz Spectroscopy and Solid-State Density Functional Theory, *Anal. Chem.* **2011**, *83*, 3786–3792.
- [140] I. Azuri, A. Hirsch, A. M. Reilly, A. Tkatchenko, S. Kendler, O. Hod, L. Kronik, Terahertz Spectroscopy of 2,4,6-Trinitrotoluene Molecular Solids from First Principles, *Beilstein J. Org. Chem.* **2018**, *14*, 381–388.
- [141] N. Bedoya-Martínez, B. Schrode, A. O. F. Jones, T. Salzillo, C. Ruzié, N. Demitri, Y. H. Geerts, E. Venuti, R. G. D. Valle, E. Zojer, R. Resel, DFT-Assisted Polymorph Identification from Lattice Raman Fingerprinting, *J. Phys. Chem. Lett.* **2017**, *8*, 3690–3695.
- [142] N. Bedoya-Martínez, A. Giunchi, T. Salzillo, E. Venuti, R. G. D. Valle, E. Zojer, Toward a Reliable Description of the Lattice Vibrations in Organic Molecular Crystals: The Impact of van der Waals Interactions, *J. Chem. Theory Comput.* **2018**, *14*, 4380–4390.

- [143] R. E. Allen, F. W. De Wette, Calculation of Dynamical Surface Properties of Noble-Gas Crystals. I. The Quasiharmonic Approximation, *Phys. Rev.* **1969**, *179*, 873–886.
- [144] F. D. Murnaghan, The Compressibility of Media under Extreme Pressures, *Proc. Natl. Acad. Sci.* **1944**, *30*, 244–247.
- [145] A. W. Hewat, C. Riekel, The Crystal Structure of Deuteroammonia Between 2 and 180 K by Neutron Powder Profile Refinement, *Acta Crystallogr. Sect. A* **1979**, *35*, 569–571.
- [146] B. Schatschneider, S. Monaco, A. Tkatchenko, J.-J. Liang, Understanding the Structure and Electronic Properties of Molecular Crystals Under Pressure: Application of Dispersion Corrected DFT to Oligoacenes, *J. Phys. Chem. A* **2013**, *117*, 8323–8331.
- [147] Y. N. Heit, K. D. Nanda, G. J. O. Beran, Predicting Finite-Temperature Properties of Crystalline Carbon Dioxide from First Principles with Quantitative Accuracy, *Chem. Sci.* **2016**, *7*, 246–255.
- [148] A. Erba, J. Maul, B. Civalleri, Thermal Properties of Molecular Crystals Through Dispersion-Corrected Quasi-Harmonic ab Initio Calculations: The Case of Urea, *Chem. Commun.* **2016**, *52*, 1820–1823.
- [149] V. G. Manzhelii, A. M. Tolkachev, I. N. Krupskii, E. I. Voitovich, V. A. Popov, L. A. Koloskova, Thermal Properties of Solid ND₃, *J. Low Temp. Phys.* **1972**, *7*, 169–182.
- [150] M. Jamal, S. Jalali Asadabadi, I. Ahmad, H. Rahnamaye Aliabad, Elastic Constants of Cubic Crystals, *Comput. Mater. Sci.* **2014**, *95*, 592–599.
- [151] B. M. Powell, G. Dolling, G. S. Pawley, J. W. Leech, Lattice Dynamics of Ammonia, *Can. J. Phys.* **1980**, *58*, 1703–1711.
- [152] H. Kiefte, S. W. Breckon, R. Penney, M. J. Clouter, Elastic Constants of Ammonia by Brillouin Spectroscopy, *J. Chem. Phys.* **1985**, *83*, 4738.
- [153] J. J. Wortman, R. A. Evans, Young’s Modulus, Shear Modulus, and Poisson’s Ratio in Silicon and Germanium, *J. Appl. Phys.* **1965**, *36*, 153–156.
- [154] J. Turley, G. Sines, The Anisotropy of Young’s Modulus, Shear Modulus and Poisson’s Ratio in Cubic Materials, *J. Phys. D: Appl. Phys.* **1971**, *4*, 264.
- [155] T. Böhlke, C. Brüggemann, Graphical Representation of the Generalized Hooke’s Law, *Technische Mechanik* **2001**, *21*, 145–158.
- [156] P. Vannucci, General Anisotropic Elasticity, Springer Singapore, Singapore, **2018**, pp. 19–73.
- [157] K. Szalewicz, Determination of Structure and Properties of Molecular Crystals from First Principles, *Acc. Chem. Res.* **2014**, *47*, 3266–3274.

-
- [158] M. A. Palafox, M. Gill, N. J. Nunez, V. K. Rastogi, L. Mittal, R. Sharma, Scaling Factors for the Prediction of Vibrational Spectra. II. The Aniline Molecule and Several Derivatives, *Int. J. Quant. Chem.* **2005**, *103*, 394–421.
- [159] D. G. Watson, R. M. Sweet, R. E. Marsh, The Crystal and Molecular Structure of Purine, *Acta Cryst.* **1965**, *19*, 573–580.
- [160] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, M. Scheffler, Resolution-Of-Identity Approach to Hartree-Fock, Hybrid Density Functionals, RPA, MP2 and GW with Numeric Atom-Centered Orbital Basis Functions, *New J. Phys.* **2012**, *14*, 053020.
- [161] A. C. Ihrig, J. Wieferink, I. Y. Zhang, M. Ropo, X. Ren, P. Rinke, M. Scheffler, V. Blum, Accurate Localized Resolution of Identity Approach for Linear-Scaling Hybrid Density Functionals and for Many-Body Perturbation Theory, *New J. Phys.* **2015**, *17*, 093020.
- [162] F. Knuth, C. Carbogno, V. Atalla, V. Blum, M. Scheffler, All-Electron Formalism for Total Energy Strain Derivatives and Stress Tensor Components for Numeric Atom-Centered Orbitals, *Comput. Phys. Commun.* **2015**, *190*, 33 – 50.
- [163] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, R. Dovesi, Ab Initio Analytical Raman Intensities for Periodic Systems Through a Coupled Perturbed Hartree-Fock/Kohn-Sham Method in an Atomic Orbital Basis. I. Theory, *J. Chem. Phys.* **2013**, *139*, 164101.
- [164] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, R. Dovesi, Ab Initio Analytical Raman Intensities for Periodic Systems Through a Coupled Perturbed Hartree-Fock/Kohn-Sham Method in an Atomic Orbital Basis. II. Validation and Comparison with Experiments, *J. Chem. Phys.* **2013**, *139*, 164102.
- [165] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, R. Dovesi, Comment on “Ab Initio Analytical Infrared Intensities for Periodic Systems Through a Coupled Perturbed Hartree-Fock/Kohn-Sham Method” [J. Chem. Phys. 137, 204113 (2012)], *J. Chem. Phys.* **2013**, *139*, 167101.
- [166] G. Folpini, K. Reimann, M. Woerner, T. Elsaesser, J. Hoja, A. Tkatchenko, Strong Local-Field Enhancement of the Nonlinear Soft-Mode Response in a Molecular Crystal, *Phys. Rev. Lett.* **2017**, *119*, 097404.
- [167] M. Woerner, W. Kuehn, P. Bowlan, K. Reimann, T. Elsaesser, Ultrafast Two-Dimensional Terahertz Spectroscopy of Elementary Excitations in Solids, *New J. Phys.* **2013**, *15*, 025039.
- [168] P. Hamm, M. Meuwly, S. L. Johnson, P. Beaud, U. Staub, Perspective: THz-Driven Nuclear Dynamics from Solids to Molecules, *Struct. Dyn.* **2017**, *4*, 061601.
- [169] C. Ouvrard, S. L. Price, Toward Crystal Structure Prediction for Conformationally Flexible Molecules: The Headaches Illustrated by Aspirin, *Cryst. Growth Des.* **2004**, *4*, 1119–1127.

- [170] A. Bond, R. Boese, G. Desiraju, On the Polymorphism of Aspirin: Crystalline Aspirin as Intergrowths of Two “Polymorphic” Domains, *Angew. Chem. Int. Ed.* **2007**, *46*, 618–622.
- [171] N. Laman, S. S. Harsha, D. Grischkowsky, Narrow-Line Waveguide Terahertz Time-Domain Spectroscopy of Aspirin and Aspirin Precursors, *Appl. Spectrosc.* **2008**, *62*, 319–326.
- [172] C. C. Wilson, Interesting Proton Behaviour in Molecular Structures. Variable Temperature Neutron Diffraction and ab Initio Study of Acetylsalicylic Acid: Characterising Librational Motions and Comparing Protons in Different Hydrogen Bonding Potentials, *New J. Chem.* **2002**, *26*, 1733–1739.
- [173] F. Wang, K. D. Jordan, A Drude-Model Approach to Dispersion Interactions in Dipole-Bound Anions, *J. Chem. Phys.* **2001**, *114*, 10717–10724.
- [174] T. Sommerfeld, K. D. Jordan, Quantum Drude Oscillator Model for Describing the Interaction of Excess Electrons with Water Clusters: An Application to $(\text{H}_2\text{O})_{13}^-$, *J. Phys. Chem. A* **2005**, *109*, 11531–11538.
- [175] A. P. Jones, J. Crain, V. P. Sokhan, T. W. Whitfield, G. J. Martyna, Quantum Drude Oscillator Model of Atoms and Molecules: Many-body Polarization and Dispersion Interactions for Atomistic Simulation, *Phys. Rev. B* **2013**, *87*, 144103.
- [176] P. S. Venkataram, J. Hermann, A. Tkatchenko, A. W. Rodriguez, Phonon-Polariton Mediated Thermal Radiation and Heat Transfer among Molecules and Macroscopic Bodies: Nonlocal Electromagnetic Response at Mesoscopic Scales, *Phys. Rev. Lett.* **2018**, *121*, 045901.
- [177] P. G. Karamertzanis, C. C. Pantelides, Ab Initio Crystal Structure Prediction—I. Rigid Molecules, *J. Comput. Chem.* **2005**, *26*, 304–324.
- [178] P. G. Karamertzanis, C. C. Pantelides, Ab Initio Crystal Structure Prediction. II. Flexible Molecules, *Mol. Phys.* **2007**, *105*, 273–291.
- [179] A. J. Stone, Distributed Multipole Analysis: Stability for Large Basis Sets, *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.
- [180] S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis, G. M. Day, Modelling Organic Crystal Structures Using Distributed Multipole and Polarizability-Based Model Intermolecular Potentials, *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.
- [181] A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman, C. C. Pantelides, Efficient Handling of Molecular Flexibility in Lattice Energy Minimization of Organic Crystals, *J. Chem. Theory Comput.* **2011**, *7*, 1998–2016.
- [182] M. A. Neumann, F. J. J. Leusen, J. Kendrick, A Major Advance in Crystal Structure Prediction, *Angew. Chem. Int. Ed.* **2008**, *47*, 2427–2430.

-
- [183] M. A. Neumann, Tailor-Made Force Fields for Crystal-Structure Prediction, *J. Phys. Chem. B* **2008**, *112*, 9810–9829.
- [184] M. A. Neumann, M.-A. Perrin, Energy Ranking of Molecular Crystals Using Density Functional Theory Calculations and an Empirical van der Waals Correction, *J. Phys. Chem. B* **2005**, *109*, 15531–15541.
- [185] J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio Jr., A. Tkatchenko, Reliable and Practical Computational Description of Molecular Crystal Polymorphs, *Sci. Adv.* **2019**, *5*, eaau3338, (arXiv preprint arXiv:1803.07503).
- [186] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, P. A. Wood, *Mercury CSD 2.0 – New Features for the Visualization and Investigation of Crystal Structures*, *J. Appl. Cryst.* **2008**, *41*, 466–470.
- [187] A. L. Spek, Structure Validation in Chemical Crystallography, *Acta Cryst. Sect. D* **2009**, *65*, 148–155.
- [188] R. A. DiStasio Jr., O. A. von Lilienfeld, A. Tkatchenko, Collective Many-Body van der Waals Interactions in Molecular Systems, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 14791–14795.
- [189] R. A. DiStasio Jr., V. V. Gobre, A. Tkatchenko, Many-Body van der Waals Interactions in Molecules and Condensed Matter, *J. Phys. Condens. Matter* **2014**, *26*, 213202.
- [190] F. Curtis, X. Wang, N. Marom, Effect of Packing Motifs on the Energy Ranking and Electronic Properties of Putative Crystal Structures of Tricyano-1,4-Dithiino[c]-Isothiazole, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 562–570.
- [191] L. J. Simons, B. W. Caprathe, M. Callahan, J. M. Graham, T. Kimura, Y. Lai, H. LeVine, W. Lipinski, A. T. Sakkab, Y. Tasaki, L. C. Walker, T. Yasunaga, Y. Ye, N. Zhuang, C. E. Augelli-Szafran, The Synthesis and Structure-Activity Relationship of Substituted N-Phenyl Anthranilic Acid Analogs as Amyloid Aggregation Inhibitors, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 654–657.
- [192] J. L. McKinley, G. J. O. Beran, Identifying Pragmatic Quasi-Harmonic Electronic Structure Approaches for Modeling Molecular Crystal Thermal Expansion, *Faraday Discuss.* **2018**, *211*, 181–207.
- [193] M. J. Schuler, T. S. Hofer, C. W. Huck, Assessing the Predictability of Anharmonic Vibrational Modes at the Example of Hydroxyl Groups – Ad Hoc Construction of Localised Modes and the Influence of Structural Solute–Solvent Motifs, *Phys. Chem. Chem. Phys.* **2017**, *19*, 11990–12001.
- [194] D. P. Schofield, J. R. Lane, H. G. Kjaergaard, Hydrogen Bonded OH-Stretching Vibration in the Water Dimer, *J. Phys. Chem. A* **2007**, *111*, 567–572.

- [195] J. Hoja, A. Tkatchenko, First-Principles Stability Ranking of Molecular Crystal Polymorphs with the DFT+MBD Approach, *Faraday Discuss.* **2018**, *211*, 253–274.
- [196] M. Mortazavi, J. G. Brandenburg, R. J. Maurer, A. Tkatchenko, Structure and Stability of Molecular Crystals with Many-Body Dispersion-Inclusive Density Functional Tight Binding, *J. Phys. Chem. Lett.* **2018**, *9*, 399–405.
- [197] J. Nyman, G. M. Day, Modelling Temperature-Dependent Properties of Polymorphic Organic Molecular Crystals, *Phys. Chem. Chem. Phys.* **2016**, *18*, 31132–31143.
- [198] G. J. O. Beran, J. D. Hartman, Y. N. Heit, Predicting Molecular Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography, *Acc. Chem. Res.* **2016**, *49*, 2501–2508.
- [199] J. G. Brandenburg, J. Potticary, H. A. Sparkes, S. L. Price, S. R. Hall, Thermal Expansion of Carbamazepine: Systematic Crystallographic Measurements Challenge Quantum Chemical Calculations, *J. Phys. Chem. Lett.* **2017**, *8*, 4319–4324.
- [200] R. Sure, S. Grimme, Corrected Small Basis Set Hartree-Fock Method for Large Systems, *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- [201] J. G. Brandenburg, S. Grimme, Accurate Modeling of Organic Molecular Crystals by Dispersion-Corrected Density Functional Tight Binding (DFTB), *J. Phys. Chem. Lett.* **2014**, *5*, 1785–1789.
- [202] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids, *Phys. Rev. Lett.* **2003**, *91*, 146401.
- [203] B. Monserrat, N. D. Drummond, R. J. Needs, Anharmonic Vibrational Properties in Periodic Systems: Energy, Electron-Phonon Coupling, and Stress, *Phys. Rev. B* **2013**, *87*, 144302.
- [204] N. D. Drummond, B. Monserrat, J. H. Lloyd-Williams, P. L. Ríos, C. J. Pickard, R. J. Needs, Quantum Monte Carlo Study of the Phase Diagram of Solid Molecular Hydrogen at Extreme Pressures, *Nat. Commun.* **2015**, *6*, 7794.
- [205] E. A. Engel, B. Monserrat, R. J. Needs, Anharmonic Nuclear Motion and the Relative Stability of Hexagonal and Cubic Ice, *Phys. Rev. X* **2015**, *5*, 021033.
- [206] M. Rossi, P. Gasparotto, M. Ceriotti, Anharmonic and Quantum Fluctuations in Molecular Crystals: A First-Principles Study of the Stability of Paracetamol, *Phys. Rev. Lett.* **2016**, *117*, 115702.
- [207] T. Bučko, S. Lebègue, T. Gould, J. G. Ángyán, Many-Body Dispersion Corrections for Periodic Systems: An Efficient Reciprocal Space Implementation, *J. Phys. Condens. Matter* **2016**, *28*, 45201.

-
- [208] F. Bernauer, "Gedrilte" Kristalle: Verbreitung, Entstehungsweise und Beziehungen zu optischer Aktivität und Molekülasymmetrie, of *Forschungen zur Kristallkunde*, Gebrüder Borntraeger, **1929**.
- [209] A. Kofler, J. Geyr, Zum mikrochemischen Nachweis des Cumarins, *Mikrochemie* **1934**, *15*, 67–73.
- [210] E. Lindpaintner, Mikroskopische Untersuchungen an polymorphen Substanzen, *Mikrochem. Ver. Mikrochim. Acta* **1939**, *27*, 21–41.
- [211] E. Gavuzzo, F. Mazza, E. Giglio, Determination of the Molecular Packing in the Crystal of Coumarin by Means of Potential-Energy Calculations, *Acta Crystallogr. Sect. B* **1974**, *30*, 1351–1357.
- [212] T.-J. Hsieh, C.-C. Su, C.-Y. Chen, C.-H. Liou, L.-H. Lu, Using Experimental Studies and Theoretical Calculations to Analyze the Molecular Mechanism of Coumarin, p-Hydroxybenzoic Acid, and Cinnamic acid, *J. Mol. Struct.* **2005**, *741*, 193 – 199.
- [213] P. Munshi, T. N. G. Row, Exploring the Lower Limit in Hydrogen Bonds: Analysis of Weak C–H···O and C–H··· π Interactions in Substituted Coumarins from Charge Density Analysis, *J. Phys. Chem. A* **2005**, *109*, 659–672.
- [214] A. R. Oganov, A. O. Lyakhov, M. Valle, How Evolutionary Crystal Structure Prediction Works—and Why, *Acc. Chem. Res.* **2011**, *44*, 227–237.
- [215] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, Q. Zhu, New Developments in Evolutionary Structure Prediction Algorithm USPEX, *Comput. Phys. Commun.* **2013**, *184*, 1172–1182.
- [216] Q. Zhu, A. R. Oganov, C. W. Glass, H. T. Stokes, Constrained Evolutionary Algorithm for Structure Prediction of Molecular Crystals: Methodology and Applications, *Acta Crystallogr. Sect. B: Struct. Sci.* **2012**, *68*, 215–226.
- [217] A. R. Oganov, C. W. Glass, Crystal Structure Prediction Using ab Initio Evolutionary Techniques: Principles and Applications, *J. Chem. Phys.* **2006**, *124*, 244704.
- [218] D. S. Coombes, S. L. Price, D. J. Willock, M. Leslie, Role of Electrostatic Interactions in Determining the Crystal Structures of Polar Organic Molecules. A Distributed Multipole Study, *J. Phys. Chem.* **1996**, *100*, 7352–7360.
- [219] B. P. van Eijck, J. Kroon, Upack Program Package for Crystal Structure Prediction: Force Fields and Crystal Structure Generation for Small Carbohydrate Molecules, *J. Comput. Chem.* **1999**, *20*, 799–812.
- [220] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

- [221] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, D. C. Langreth, Higher-Accuracy van der Waals Density Functional, *Phys. Rev. B* **2010**, *82*, 081101.
- [222] H.-M. Wolff, L. Quere, J. Riedner, Polymorphic Form of Rotigotine, **2010**, EP 2215072A2.
- [223] M. A. Neumann, J. van de Streek, F. P. A. Fabbiani, P. Hidber, O. Grassmann, Combined Crystal Structure Prediction and High-Pressure Crystallization in Rational Pharmaceutical Polymorph Screening, *Nat. Commun.* **2015**, *6*, 7793.
- [224] M. Salvalaglio, T. Vetter, F. Giberti, M. Mazzotti, M. Parrinello, Uncovering Molecular Details of Urea Crystal Growth in the Presence of Additives, *J. Am. Chem. Soc.* **2012**, *134*, 17221–17233.
- [225] J. Anwar, D. Zahn, Uncovering Molecular Processes in Crystal Nucleation and Growth by Using Molecular Simulation, *Angew. Chem. Int. Ed.* **2011**, *50*, 1996–2013.
- [226] S. L. Price, Is Zeroth Order Crystal Structure Prediction (CSP₀) Coming to Maturity? What Should We Aim for in an Ideal Crystal Structure Prediction Code?, *Faraday Discuss.* **2018**, *211*, 9–30.
- [227] G. R. Woollam, M. A. Neumann, T. Wagner, R. J. Davey, The Importance of Configurational Disorder in Crystal Structure Prediction: The Case of loratadine, *Faraday Discuss.* **2018**, *211*, 209–234.
- [228] L. Iuzzolino, P. McCabe, S. Price, J. G. Brandenburg, Crystal Structure Prediction of Flexible Pharmaceutical-Like Molecules: Density Functional Tight-Binding as an Intermediate Optimisation Method and for Free Energy Estimation, *Faraday Discuss.* **2018**, *211*, 275–296.
- [229] G. A. Dolgonos, O. A. Loboda, A. D. Boese, Development of Embedded and Performance of Density Functional Methods for Molecular Crystals, *J. Phys. Chem. A* **2018**, *122*, 708–713.
- [230] O. A. Loboda, G. A. Dolgonos, A. D. Boese, Towards Hybrid Density Functional Calculations of Molecular Crystals via Fragment-Based Methods, *J. Chem. Phys.* **2018**, *149*, 124104.
- [231] T.-Q. Yu, M. E. Tuckerman, Temperature-Accelerated Method for Exploring Polymorphism in Molecular Crystals Based on Free Energy, *Phys. Rev. Lett.* **2011**, *107*, 015701.