

THERMAL CONTRIBUTIONS TO RELATIVE FREE ENERGIES
OF ALLOTROPES AND POLYMORPHS FROM
DENSITY-FUNCTIONAL THEORY

by

Joseph A. Weatherby

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
April 2022

© Copyright by Joseph A. Weatherby, 2022

Table of Contents

List of Tables	iv
List of Figures	v
List of Abbreviations and Symbols Used	vi
Abstract	xi
Acknowledgements	xii
Chapter 1 Introduction	1
Chapter 2 Theory	7
2.1 Density-Functional Theory	7
2.2 Modelling Periodic Solid-State Systems	10
2.3 The Pseudopotential Approximation	12
2.4 The Exchange-Hole Dipole Moment Model	14
2.5 Lattice Dynamics in Crystals	18
2.6 Calculating Vibrational Contributions to the Gibbs Free Energy	21
2.7 The Quasi-Harmonic Approximation	23
2.8 Low-Cost Computational Methods for Modelling Solid-State Systems	24
Chapter 3 The Relative Stability of Graphite and Diamond	26
3.1 Introduction	26
3.2 Computational Methodology	27
3.3 Results and Discussion	28

3.4	Summary	34
Chapter 4	Helicenes	36
4.1	Introduction	36
4.2	Computational Methodology	39
4.3	Results and Discussion	40
4.3.1	4,13-difluoro[6]helicene	43
4.3.2	13-fluoro[6]helicene	43
4.3.3	6-triacetylene[6]helicene	44
4.3.4	7,10-dialdehyde[6]helicene	44
4.4	Conclusion	45
Chapter 5	A Density-Functional Benchmark of Thermal Free-Energy Corrections for Polymorphic Molecular Crystals . . .	46
5.1	Introduction	46
5.2	Background Theory	49
5.2.1	Phonons and Free-Energy Corrections	49
5.2.2	Mode Matching	50
5.2.3	Treatment of The Acoustic Modes	53
5.3	Computational Methodology	54
5.4	Results	55
5.4.1	The PV17 Benchmark	55
5.4.2	Assessment of Low-Level Methods	59
5.5	Discussion and Conclusions	64
Chapter 6	Conclusions and Future Work	66
6.1	Summary of Chapters	66
6.2	Future Work	68
Bibliography	70

List of Tables

3.1	Optimized DFT cell parameters and volumes for graphite and diamond, compared with experimental data.	29
3.2	Experimentally determined values for the ΔH , ΔS , and ΔG between graphite and diamond.	30
3.3	Electronic and Gibbs energies of graphite, relative to diamond, obtained with selected functionals and XDM dispersion corrections.	31
3.4	Ground state electronic energy of graphite, relative to diamond, based on different DFT models.	32
5.1	Molecular structures, CCDC codes, and F_{vib} results for the PV17 dataset.	56
5.2	Relative B86bPBE-XDM electronic energies, converged vibrational free-energy corrections, and relative free energies for the PV17 benchmark.	60
5.3	Absolute F_{vib} values for all crystals in our PV17 benchmark set with selected electronic structure methods.	61
5.4	Relative F_{vib} values for all polymorph pairs in the PV17 benchmark set with selected computational methods.	62

List of Figures

1.1	Simplified 1D illustration of the difference between optical and acoustical phonon modes for a linear diatomic chain.	3
2.1	Phonon density of states for CO ₂ computed with DFPT and finite difference for various sampling grids.	21
3.1	Exfoliation curves of graphite computed with selected density functionals and dispersion corrections.	30
3.2	Thermodynamic energy differences for graphite and diamond, as a function of temperature at ambient pressure.	33
4.1	The two enantiomers of the chiral [6]helicene molecule.	37
4.2	Image of the translational dimer motif commonly observed in [6]helicene crystals.	38
4.3	Four promising OSC molecular materials identified by a computational screening approach.	40
4.4	Crystal-energy landscapes for four promising [6]helicene materials.	41
4.5	Low-energy crystal structures for four promising [6]helicene materials. The translational dimer motifs are highlighted.	42
4.6	Electron mobilities for functionalized [6]helicene structures.	43
5.1	Overlay of the DFT phDOS for the $I\beta$ phase of ethylenediamine computed at Γ and using a converged \mathbf{q} -point grid.	51
5.2	Overlays of the phDOS of the $I\beta$ phase of ethylenediamine computed with QE, DFTB, and mode matching.	52

List of Abbreviations and Symbols Used

Abbreviation	Description
a.u.	Atomic unit
AE	All Electron
AM	Acoustical phonon Modes
AO	Atomic Orbital
APW	Augmented Plane-Wave method
ATM	Axilrod-Teller-Muto three-body dispersion-energy term
B86bPBE	Becke's B86b exchange and PBE correlation GGA density functional
B86bPBE-25X	B86bPBE-based hybrid functional with 25% exact exchange
BJ	Becke-Johnson
BR	Becke-Roussel
BSSE	Basis-Set Superposition Error
CCDC	Cambridge Crystallographic Data Center
CP	Counterpoise
CSD	Cambridge Structure Database
CSP	Crystal Structure Prediction
DFA	Density-Functional Approximation
DFPT	Density-Functional Perturbation Theory
DFTB	Density-Functional Tight Binding
DFT	Density-Functional Theory
DMA	Distributed Multipole Analysis
DMACRYS	Price's distributed multipole force-field code
D_n	Grimme's series of dispersion corrections (e.g., $n= 2, 3, 4$)
DZP	Double-Zeta basis set with Polarization functions
Expt	Experimental
FHIaims	Fritz Haber Institute <i>ab initio</i> molecular simulations
gCP	Geometrical Counterpoise Method
GGA	Generalized-Gradient Approximation
HA	Harmonic Approximation

Abbreviation	Description
HF	Hartree–Fock theory
(s)HF-3c	Grimme’s (scaled) minimal basis-set Hartree–Fock method with semi-empirical corrections
HSE06	Heyd–Scuseria–Ernzerhof screened exchange hybrid functional
KB49	Kannemann–Becke benchmark set of molecular dimers
KS	Kohn and Sham
LAPW	Linearised Augmented Plane-Wave methods
LSDA	Local Spin Density Approximation
MO	Molecular Orbital
MP	Monkhorst–Pack \mathbf{k} -point grid sampling scheme
NAO	Numerical Atomic Orbitals
NC	Norm Conserving
NPL	Nyman Polymorph Library
OSC	Organic Semi-Conductor
PAW	Projector Augmented-Wave
PBE	Perdew–Burke–Ernzerhof GGA density functional
PBEsol	Perdew–Burke–Ernzerhof functional revised for solids
PBE0	Perdew–Burke–Ernzerhof hybrid density functional
phDOS	Phonon Density of States
PV17	Benchmark set of 17 polymorphic organic molecular crystal pairs for vibrational free-energy corrections
QHA	Quasi-Harmonic Approximation
QE	QUANTUM ESPRESSO
QM	Quantum Mechanics
SCF	Self-Consistent Field
SIESTA	Spanish Initiative for Electronic Simulations with Thousands of Atoms
SRB	Short-Ranged Basis
TM	Troullier–Martins
VASP	Vienna <i>ab initio</i> Simulation Package
vdW	van der Waals
W99	Williams ’99 Force Field
X23	Benchmark set of 23 molecular organic solids
XDM	Exchange-hole Dipole Moment dispersion model

Symbol	Description
a_1, a_2	XDM Becke-Johnson Damping Parameters
a_x	Fraction of Exact Exchange
\mathbf{a}	Unit Cell Vector
\mathbf{b}	Reciprocal Cell Vector
$c_{\mathbf{k}+\mathbf{G}}$	Fourier Coefficient
$C_{n,ij}$	Two-Body Dispersion Coefficient (e.g., $n = 6, 8, 10$)
C_v	Constant-Volume Heat Capacity
C	Force-Constant Matrix
c_x	Constant Defined in LSDA and GGA Exchange Energy Expressions
$d_{X\sigma}$	Exchange-Hole Dipole Moment
E_{base}	Total DFT Energy
E_c	Correlation Energy
E^{cut}	Planewave Kinetic-Energy Cutoff
E_{disp}	Dispersion Energy
E_{DFT}	Total DFT Energy
E_x^{GGA}	GGA Exchange Energy
E_x^{LSDA}	LSDA Exchange Energy
E_{static}	Static Electronic Energy
E_x	Exchange Energy
E_x^{HF}	Hartree-Fock Exact Exchange Energy
E_x^{hybrid}	Hybrid Exchange Energy
E_{XC}	Exchange-Correlation Energy
F	Enhancement Factor in GGA Functionals
\mathbf{F}_{disp}	Dispersion Contribution to Atomic Forces
F_I	Forces Acting on Atom I
$f(R_{ij,\mathbf{L}})$	Damping Function
F	Helmholtz Free Energy
G	Gibbs Free Energy
\mathbf{G}	Reciprocal Lattice Vector
\hat{H}	Electronic Hamiltonian
H	Enthalpy
\hbar	Reduced Planck's Constant
$h_{X\sigma}$	Spin-Dependant Exchange Hole

Symbol	Description
i, j	Atomic Index, Electronic Index
J	Classical Coulomb Self Energy
k_B	Boltzmann's Constant
\mathbf{k}	Reciprocal Space Vector
\mathbf{L}	Lattice Vector
M	Atomic Mass
M_l	l -th Order Multipole Moment
n	Occupation Number
\tilde{p}_i	Projector Functions
p_{th}	Thermal Pressure
p	Pressure
q	Electronic Charge
\mathbf{q}	Reciprocal Space Vector
r	Radius
r_{12}	Interelectronic Separation
r_c	Radial Cutoff
$r_{c,l}$	Radial Cutoff with Angular Momentum l
\mathbf{R}	Atomic Coordinates
\mathbf{R}	Bravais Lattice Vector
$R_{c,ij}$	Critical Radius in XDM Model
R_{ij}	Interatomic Distance
$R_{ij,\mathbf{L}}$	Interatomic Distance in a Periodic System
$R_{\text{vdW},ij}$	van-der-Waals Radius
R_k	Length Parameter For Computing \mathbf{k} Grids
S	Entropy
T_0	Kinetic-Energy of Non-Interacting Electrons
T	Temperature
V	Volume, Periodic Potential
V_{ext}	External Potential
α_i	Atom-in-Solid Polarizability
α_i^{free}	Free-Atom Polarizability
χ_σ	Reduced Density Gradient
β	Constant Defined in B86bPBE Enhancement Factor Expression

Symbol	Description
γ	Constant Defined in B86bPBE Enhancement Factor Expression
γ_{ab}	DFTB Function Describing On-Site Self Repulsion
Γ	Brillouin-Zone Center
Γ_{ij}	Christoffel Matrices
δ_{ij}	Kronecker Delta
ε_i	Orbital Energy
η, ξ	Cartesian Directions
∇^2	Laplacian Operator
$\rho(\mathbf{r})$	Electron Density Function
$\mu \nu$	Atom Orbital Indices
μ	Electron Mobility
σ	Electron Spin Index
σ_{disp}	van-der-Waals Contribution to Cell Stress
ϕ	Atomic Orbital
$\tilde{\phi}$	Pseudo Partial Wave
ψ	Electronic Wavefunction
Ψ	Total Electronic Wavefunction
$\tilde{\Psi}$	Pseudised Wavefunction
$\psi_{\mathbf{k}}$	Bloch Wavefunction
ω_{ac}	Acoustic Phonon Frequency
ω	Hirshfeld Weight
ω	Phonon Frequency
ω_D	Debye Frequency

Abstract

Relative stabilities of two or more crystalline phases, such as allotropes or polymorphs, can be predicted theoretically using density-functional theory (DFT). Understanding the stability landscape of a given system has far-reaching applications in the pharmaceutical industry and materials modelling. For example, the focus could be to screen for compounds with specific properties, or to complement experiment in determining the isolable crystal structure.

Crystal structure prediction (CSP) is a rapidly evolving field of computational chemistry. The over-arching goal of CSP is to predict the crystal structure of a given organic molecule beginning from its 2D chemical diagram. Being able to routinely conduct CSP studies is highly desirable, but is complicated by the complexity of the potential-energy surface that must be explored due to the many possible ways molecules can arrange themselves in the solid-state. DFT is routinely used to compute the relative energies of polymorphs in CSP studies, but temperature effects are frequently neglected. While DFT phonon calculations provide the zero-point and thermal contributions to the relative free energies of polymorphic systems, they are often intractable for the size of systems commonly encountered in CSP studies.

The work contained in this thesis aims to study several problems concerning allotropes, polymorphism, and free-energy corrections. We examine two allotropes of carbon, diamond and graphite, and apply DFT to compute the relative free-energy difference. By undertaking this study, we can use high-accuracy theoretical data to determine which allotrope of carbon is more thermodynamically stable. With regards to polymorphism, we examine functionalized [6]helicene systems for organic electronic applications and use DFT to propose several low-energy crystal structures that may be crystallized experimentally. Finally, we conduct a benchmark study of thermal corrections of polymorphic molecular crystals and assess the accuracy of selected low-cost methods in hopes of finding a cheaper alternative to computationally expensive DFT phonon calculations.

Acknowledgements

First, a general thank you to *everyone* who has supported and believed in me throughout these last three years, it has been an absolutely amazing experience. I'd like to start by acknowledging my incredible supervisor, Dr. Erin Johnson, for everything that she has contributed to not only this work but also to my growth as a person. I appreciate everything that I've learned from you about doing research in science as well as being a mentor to others. Thank you, Erin.

I extend my sincere thanks to my committee members, Drs. Josef Zwanziger, Aaron Kelly, and Saurabh Chitnis. Thank you for your amazing support and insight which contributed to this work.

I'd like to acknowledge Dr. Alberto Otero de la Roza from the Universidad de Oviedo in Spain. Thank you for mentoring me throughout my training as a computational scientist.

I acknowledge several amazing scientists for their contribution to this work via collaborations or providing software codes: Cameron Cook, Drs. Greg Beran, Jonas Nyman, Graeme Day, Kim Jelfs, Julia Schmidt, and Mary Anne White.

Thank you to the past and present Johnson group members whom I've had the absolute pleasure of working alongside: Stephen, Mat, Luc, Fred, Alastair, Tilas, Kyle, Alex, Sarah, Adrian, Grace, and Adebayo. Thank you guys for being amazing friends and colleagues.

Thank you to the greater Dalhousie community and chemistry department for the amazing support, comraderie, and encouragement. Thank you to the Dasog group members for their incredible support and friendship during the COVID-19 pandemic.

I am grateful to the Government of Nova Scotia for graduate scholarship opportunities (NSGS). I am grateful for the computational resources and friendly technical support of Compute Canada and ACENET. Lastly, thank you, the reader, for taking the time to read this work.

CHAPTER 1

Introduction

The field of computational chemistry exploits models based on physics and mathematics to make predictions about systems that are of interest to chemists. There are a large number of computational models that exist in the literature that are designed specifically for a given purpose. What sets the different models apart is the underlying theory (quantum mechanics or classical mechanics, for example) and the approximations or assumptions invoked, which leads to varying degrees of accuracy and computational time. Researching and understanding the pitfalls and successes of a given model is paramount to make sure one is using an appropriate model for the application at hand.

There are many applications of computational chemistry, such as modelling reaction pathways,¹ bond-dissociation energies,² and binding energies for molecules on surfaces.³ Computational chemistry can be used to calculate thermodynamic quantities of molecules and materials, such as the Gibbs or Helmholtz free energy, enthalpy, entropy, etc.⁴ One particular focus of this work is computing relative (free) energies of different polymorphs (phases) of a solid-state material to determine which may be the likely structure(s) to be observed experimentally.⁵ The work contained in this thesis aims to apply a specific computational model, known as density-functional theory (DFT), to allotropes and polymorphs. We will show results for three projects in total, consisting of two short, collaborative projects (Chs. 3 and 4) and one more in-depth project (Ch. 5). In Ch. 3, we study the relative thermodynamic stability of two allotropes of carbon: diamond and graphite. We use very accurate DFT models to compute the relative energy and subsequently perform phonon calculations to ascertain the temperature contribution to the relative Gibbs free energy. Next, Ch. 4 focuses on applying a crystal structure prediction protocol to determine accurate energetic rankings of polymorphs of functionalized [6]helicene systems. These

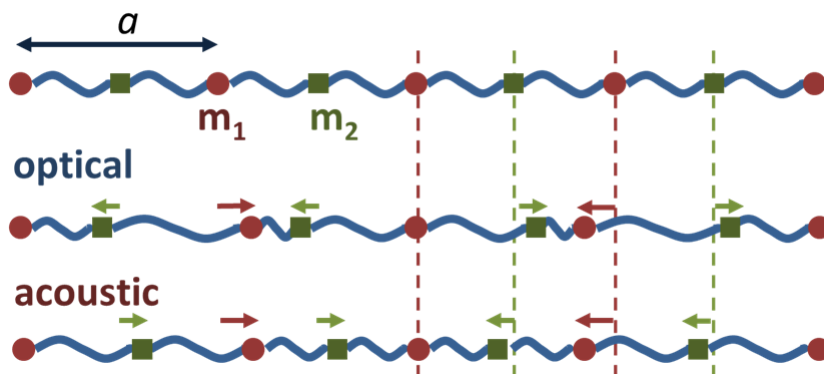
[6]helicenes are interesting as they are relevant materials in electronic applications,⁶⁻⁸ such as organic light-emitting diodes.⁹ Lastly, we undertake an in-depth benchmark study of thermal corrections in polymorphic organic molecular crystals. We use accurate (and expensive) DFT models to compute free-energy corrections for a small benchmark set and subsequently explore the viability of various low-cost models to reproduce the high-level data.

As stated above, DFT¹⁰⁻¹² is the main computational framework that we use in this work. DFT is a quantum-mechanical model where the central quantity is the electron density function, $\rho(\mathbf{r})$, where \mathbf{r} contains the three spatial coordinates. The electron density is used to model chemical interactions, such as bonding. DFT is widely used in chemistry and physics research as a tool to supplement experimental results, as well as to make predictions that inform experiment.^{13,14} In our case, we are interested in using DFT to calculate minimum-energy crystal structures and phonon frequencies, which are then used to calculate finite-temperature thermodynamic quantities.

Phonons can be thought of as the vibrations of atoms that occur in a crystalline lattice. Solids exhibit two types of phonon modes: acoustical and optical phonons. To illustrate the difference between the two phonon modes, consider a one-dimensional chain of atoms as shown in Fig. 1.1. Acoustic phonons are vibrations that involve coherent movements of atoms, while optical phonons are those that involve incoherent movements of atoms. Intramolecular vibrations can be influenced by the surrounding molecules in the crystal lattice due to intermolecular interactions, such as hydrogen bonding and π stacking. Finally, we can treat the individual vibrations as harmonic oscillators, allowing for straightforward formulae to evaluate the entropy, heat capacity, enthalpy corrections, and free-energy corrections, based on only the vibrational frequencies.

Most DFT calculations are performed assuming no energetic or structural effects arising from finite temperature. Under these circumstances, we are only calculating the energy of the static electrons, while ignoring thermal motion. While neglecting thermal effects can be a valid approximation in many cases, in order to obtain a more accurate description of the system of interest, we can account for temperature effects. Moreover, many properties in crystals are affected by temperature, such as lattice expansion/molar volumes,¹⁶ thermochemical properties,^{16,17} mechanical properties,^{17,18} vibrational spectra,¹⁹ and nuclear

Figure 1.1: Simplified 1D illustration of the difference between optical and acoustical phonon modes for a linear diatomic chain, with atoms of mass m_1 and m_2 and lattice parameter a . Optical modes are those where the atoms are moving incoherently, whereas acoustical modes are those where the atoms are moving coherently. Image taken from Ref. 15.



magnetic resonance (NMR) chemical shifts.²⁰

Crystal structure prediction (CSP) is one field of computational chemistry that can benefit from the computation of thermal effects. CSP aims to predict the most stable solid-state polymorph of a given molecular crystal, beginning from the 2D chemical diagram. The CSP problem is extremely challenging from a computational perspective due to the many possible ways that molecules can arrange themselves in the solid-state, known as polymorphism.^{21,22} Large-scale statistical studies on hundreds of polymorphic organic crystals showed that thermal effects have the ability to rearrange relative ordering of about 10-20% of structures.^{23,24} These results suggest that thermal effects have a small, but sizeable, impact on the relative stability ordering of polymorphs. Therefore, it is important that we consider the energetic contributions arising from temperature so we are correctly ranking structures, especially at the final stages of a CSP protocol.

It is commonplace for CSP to be broken down into a series of steps (we refer to the collection of these steps as a protocol) involving several different theoretical models and computational software packages. The reason for this is that molecular crystals, especially those of practical relevance, tend to adopt very large unit-cell geometries and very inexpensive (albeit inaccurate) models are frequently used for initial structure generation. Due to their high cost and relatively low contribution (compared to relative electronic energies), calculating phonon frequencies and thermal effects is usually the very last step of a CSP protocol.

Another application of computing phonon frequencies and thermodynamic functions is in modeling phase diagrams. A phase diagram is a plot of temperature versus pressure and displays the most stable phase of a compound at the given conditions. Phase diagrams are particularly relevant for both allotropes and polymorphs because different phases can be stable under different temperature or pressure conditions. For example, one polymorph may be stable at room temperature and ambient pressure; however, upon applying pressure, another polymorph may become more stable. DFT electronic energies are insufficient to give us enough information to be able to conclude the presence of different phases under these different conditions. As a result, adding free-energy and pressure-volume corrections allows us to study the temperature and pressure dependence of allotropes and polymorphs.

The most common way to supplement the electronic energy with thermal effects within the DFT framework is to calculate phonon frequencies and apply formulas derived from statistical mechanics within either the harmonic or quasi-harmonic approximations. The harmonic approximation assumes that the phonon frequencies are independent of volume, which leads to behaviour such as infinite thermal conductivity and zero thermal expansion.²⁵ By extension of zero thermal expansion being predicted, mechanical properties like elastic constants and bulk moduli are calculated independent of temperature (which is strictly an approximation). The quasi-harmonic approximation is a more sophisticated approach to introduce phonon dependence on volume, yielding improved results when temperature and pressure effects are important.^{17,26} Both the harmonic and quasi-harmonic approximations are prescriptions that can be applied to any system and depend on the computed phonon frequencies. We now discuss the practical aspects of such phonon calculations in more detail.

In general, a phonon frequency calculation is significantly more expensive compared to one single-point energy calculation (evaluating the energy of the electrons at a given geometry) or even a complete geometry optimization. Broadly speaking, in the solid state, there are two main contributors that dictate the cost of a phonon calculation: the number of atoms in the unit cell and the sampling of the Brillouin zone. The number of phonon modes in a system is $3N$, where N is the number of atoms, and each individual mode must be solved for iteratively. Moreover, we must calculate $3N$ phonon modes for each \mathbf{q} -point that is sampled in the Brillouin zone. We must converge a quantity of interest (for example, the Gibbs free energy) with respect to \mathbf{q} -point sampling, which quickly becomes

very expensive computationally. Symmetry can be exploited to reduce the computational time; however, it remains quite impractical to calculate phonons for crystal structures with large unit cells. As a result, it is desirable to find ways to compute phonon frequencies at a reduced cost.

In CSP, one way to address the problem of phonon frequencies being expensive to compute is to limit the number of structures for which we calculate thermal contributions. That is, we can ensure that we are using very accurate energy-ranking methods to filter out as many structures as possible, minimizing the number of phonon calculations we need to perform. Additionally, because full plane-wave DFT can be very expensive, we can use cheaper computational models, such as semi-empirical approaches like HF-3c,²⁷ which employs a minimal basis set. The orbitals and electron density are represented as linear combinations of functions, termed a basis set, which typically consists of either plane waves or Gaussian functions. HF-3c uses a minimal Gaussian basis set to construct the orbitals, greatly decreasing its computational cost, as well as its accuracy. To compensate for using a minimal basis set, the HF-3c method uses three distinct energy corrections to try and recover higher accuracy. Another low-cost computational model is density-functional based tight binding (DFTB).²⁸ DFTB models are derived by expansion of the total-energy functional from DFT via a Taylor series up to third order. The terms in this Taylor expansion are then approximated in several different ways, such as by using DFT or empirical data, or superposition of atomic densities. Finally, DFTB models are parametrized for a given reference system, leading to many different parameterizations. Using a cheaper method, such as HF-3c or DFTB, ultimately brings up the cost/accuracy dilemma. Semi-empirical methods may be too inaccurate to be viable options, which is why it is crucial to extensively test and validate these models.

There are several projects contained in this thesis two of which are short, collaborative projects and one is an in-depth computational benchmark study. The first of the two collaborative projects involves addressing the question of which allotrope of carbon is the most thermodynamically stable, graphite or diamond. We achieve this by applying high-level periodic DFT calculations within the quasi-harmonic approximation to compute the relative entropies, enthalpies, and Gibbs free energies. Next, we apply DFT within the SIESTA²⁹ framework to functionalized [6]helicene molecules and examine their potential as organic-semiconducting (OSC) materials. Helicene systems are an excellent case to

apply CSP as the conformational landscape of the crystals can be large and, therefore, polymorphism is an inherent problem that must be studied. Lastly, we generate a benchmark set of relative thermal corrections for a list of polymorph pairs obtained from the Nyman-Day polymorph library.^{23,24} This benchmark is then used to assess the performance of lower-cost models, with the hope that one will be sufficiently accurate that we can forego the expensive full plane-wave DFT phonon calculations and use a more expedient method to obtain reliable results.

CHAPTER 2

Theory

2.1 Density-Functional Theory

Density-functional theory (DFT) is a quantum-mechanical theory that revolves around a system's electron density function, $\rho(\mathbf{r})$. Formally, we can express the total density as a summation of the probability densities of the occupied (spin) orbitals,

$$\rho(\mathbf{r}) = \sum_{i,\sigma} |\phi_{i,\sigma}(\mathbf{r})|^2. \quad (2.1)$$

The theoretical foundations for DFT arise from early works by Hohenberg, Kohn, and Sham.^{10,11} There are several theorems that set the stage for DFT, of which we provide an overview here. The first theorem, proved by Hohenberg and Kohn, is that the ground-state energy from Schrödinger's equation ($\hat{H}\Psi = E\Psi$) is a unique functional of the electron density, $E[\rho]$. By functional, we mean a function of a function, that is, a function of the electron-density function. The first Hohenberg-Kohn theorem is useful because it tells us that the ground-state electron density uniquely determines all the properties of our system, analogous to the N -electron wavefunction, Ψ .¹² Despite the fact that a unique energy functional of the electron density *exists*, we are not given any indication of its mathematical form — a problem which we shall discuss shortly. The second Hohenberg-Kohn theorem tells us that the electron density that minimizes the energy functional is the electron density corresponding to the full solution of the Schrödinger equation, allowing us to invoke the variational principle. Finally, we consider Kohn-Sham (KS) theory, in which a system of noninteracting electrons is used to make several approximations to address the problem of the unknown energy functional. Assuming a noninteracting electron reference system, the DFT energy can be formulated as a functional of the electron density and is expressed

as

$$E_{\text{DFT}} \equiv E[\rho] = T_0[\rho] + \int V_{\text{ext}}\rho(\mathbf{r})\mathbf{d}\mathbf{r} + J[\rho] + E_{\text{XC}}[\rho]. \quad (2.2)$$

Here, $T_0[\rho]$ is the approximation to the kinetic energy considering a noninteracting reference system, given by

$$T_0 = - \sum_i \int \phi_i^*(\mathbf{r})\nabla^2\phi_i(\mathbf{r})\mathbf{d}\mathbf{r}, \quad (2.3)$$

where ∇^2 is the Laplacian operator. The second term is the electron-nuclear interaction energy, where V_{ext} is determined from the nuclear positions. $J[\rho]$ is the classical Coulomb electron repulsion energy

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}}\mathbf{d}\mathbf{r}_1\mathbf{d}\mathbf{r}_2, \quad (2.4)$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. Lastly, E_{XC} , the exchange-correlation energy term, can be thought of as the error made by considering classical electron-electron interactions. Unfortunately, there is no exact expression for the exchange-correlation energy and, in practice, we resort to approximations. We can variationally minimize Eq. 2.2 with respect to the orbitals, ϕ_i , to obtain the Kohn-Sham equations, solved self-consistently, and expressed as:

$$-\frac{1}{2}\nabla^2\phi_i + v_{\text{KS}}\phi_i = \epsilon_i\phi_i, \quad (2.5)$$

where ϵ_i is the orbital energy and v_{KS} is the Kohn-Sham potential, given by

$$v_{\text{KS}} = v_{\text{ext}} + v_{\text{el}} + \frac{\delta E_{\text{XC}}}{\delta\rho}. \quad (2.6)$$

The KS potential includes three terms: the external potential (electron-nuclear interactions), the classical Coulomb potential, v_{el} , given by

$$v_{\text{el}} = \int \frac{\rho(\mathbf{r}_2)}{r_{12}}\mathbf{d}\mathbf{r}_2, \quad (2.7)$$

and the functional derivative, $\delta E_{\text{XC}}/\delta\rho$.¹²

All practical calculations under the DFT framework use density-functional approximations (DFAs), simply called functionals. For each functional, the exchange-correlation energy can be split into exchange and correlation terms as

$$E_{\text{XC}} = E_{\text{X}} + E_{\text{C}}. \quad (2.8)$$

Throughout this section, we are mainly interested in the choice of exchange functional, as E_X contributes more to E_{XC} than E_C . We now address various ways to approximate E_X , beginning with the local spin-density approximation (LSDA). The LSDA approximates the exchange-energy contribution from each point in space as being equal to the corresponding result for a uniform electron gas with the same density. Here, E_X is

$$E_X^{\text{LSDA}} = -\frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \sum_{\sigma} \int \rho_{\sigma}^{4/3}(\mathbf{r}) d\mathbf{r}, \quad (2.9)$$

where ρ_{σ} is the spin-dependent density. Note that E_X^{LSDA} involves only the electron density, $\rho(\mathbf{r})$, which is the reason for the LSDA being termed local in the density. The LSDA is reasonable for systems with slowly varying density, such as metals; however, for almost all chemical systems of interest the density is not slowly varying. Therefore, a more sophisticated approach is required to capture more information about how the density is varying in space. The generalized gradient approximation (GGA) was formulated to expand upon the LSDA by including dependence on the gradient of the density, $\nabla\rho$. We can write the form of a GGA functional as follows:

$$E_X^{\text{GGA}} = c_X \sum_{\sigma} \int F(\chi_{\sigma}) \rho_{\sigma}^{4/3}(\mathbf{r}) d\mathbf{r}, \quad (2.10)$$

where c_X is a constant and $F(\chi_{\sigma})$ is referred to as the enhancement factor and is dependent on the reduced density gradient, χ_{σ} , given by

$$\chi_{\sigma} = \frac{|\nabla\rho_{\sigma}(\mathbf{r})|}{\rho_{\sigma}^{4/3}(\mathbf{r})}. \quad (2.11)$$

The enhancement factor varies depending on the choice of functional used and is normally formulated to obey known limits. The most commonly used exchange functional is that of Perdew, Burke, and Ernzerhof (PBE),³⁰ while the exchange functional we employ most often in this work is B86b.³¹ Both PBE and B86b are often paired with the PBE correlation functional, with the full names of these exchange-correlation functionals written as PBE/PBE (simply PBE) and B86b/PBE. B86b/PBE is our functional of choice as it yields excellent results in molecular crystal applications.³² Both exchange functionals share the form for a GGA given in Eq. 2.10 with

$$F^{\text{PBE}}(\chi_{\sigma}) = 1 + \frac{\beta}{c_X} \frac{\chi_{\sigma}^2}{1 + \gamma\chi_{\sigma}^2} \quad (2.12)$$

and

$$F^{\text{B86b}}(\chi_\sigma) = 1 + \frac{\beta}{c_x} \frac{\chi_\sigma^2}{(1 + \gamma\chi_\sigma^2)^{4/5}}, \quad (2.13)$$

where β and γ are constants, which vary depending on the functional. PBE recovers the correct behaviour of the enhancement factor, in the zero-gradient limit, which is ideal for metals, but not in the large-gradient limit. B86bPBE, on the other hand, recovers correct behaviour in the large-gradient limit, which is ideal for intermolecular interactions. GGAs represent a massive improvement over the LDA, particularly with regards to thermochemistry.³³ However, GGAs tend to overbind systems, in contrast with Hartree-Fock (HF), which tends to underbind.¹² With regards to binding, we are referring to the bond energy, not necessarily the bond length.

We now consider another class of DFAs, known as hybrid functionals, which are based on a GGA but also include some percentage of exact (HF) exchange. The rationale for the good performance of hybrids is straightforward; GGAs tend to overbind and HF tends to underbind, so a combination of the two should therefore offer an improvement. In general, a hybrid functional can be expressed as

$$E_X^{\text{hybrid}} = a_X E_X^{\text{HF}} + (1 - a_X) E_X^{\text{GGA}}, \quad (2.14)$$

where a_X is between 0 and 1 (depending on the particular hybrid functional) and E_X^{HF} is the exact-exchange energy given by

$$E_X^{\text{HF}} = -\frac{1}{2} \sum_\sigma \iint \phi_i^*(\mathbf{r}_1) \phi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \phi_j(\mathbf{r}_1) \phi_i(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.15)$$

We mainly use two hybrid functionals in this work, PBE0³⁴ and B86bPBE-25X,³⁵ which contain 25% exact exchange ($a_X = 0.25$) and are based on PBE and B86bPBE, respectively. Due to the inherent nonlocal nature of HF exchange, production calculations involving hybrid functionals quickly become intractable with plane-wave basis sets. Fortunately, algorithmic advances have been made to allow hybrid functionals to be used in small solid-state calculations.³⁶

2.2 Modelling Periodic Solid-State Systems

Given the task of modelling a solid-state system, such as a material or molecular crystal, we are presented with the challenge of representing an infinite number of electrons. Thanks

largely to the advances of solid-state physics, we are able to minimize the problem down to repeating sub units, or unit cells. Unit cells are segments of the overall crystal structure that can be translated in the three spatial dimensions to construct the entire system. Usually, we want to work with the primitive unit cell as that is the smallest unit cell that we can have for a given system with a particular atomic arrangement. Note that when we are considering unit cells, there is an assumption of high structural order, that is, the regularly repeating units are all exactly the same, which is not necessarily true for a bulk material in reality. We can specify any unit cell by three lattice vectors given by

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (2.16)$$

where a_i are the primitive lattice vectors which are scaled by a coefficient n_i that can be any integer value. Eq. 2.16 allows us to build the Bravais lattice, of which there are 14 possible symmetries. Bravais lattices are an indispensable tool in solid-state computations because of the ability to exploit their symmetry to reduce computational cost.

Consider the potential, $V(\mathbf{r})$, within the full crystal, which will remain unchanged when considering displacements by a Bravais vector,

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R}). \quad (2.17)$$

Due to the periodicity of the potential, Bloch's theorem³⁷ gives us a prescription for the one-electron wavefunctions, ϕ , in terms of plane waves, $e^{i\mathbf{k}\cdot\mathbf{r}}$, expressed as

$$\phi_{j,\mathbf{k}}(\mathbf{r}) = u_j(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (2.18)$$

where \mathbf{k} is a wave vector. Here, $u(\mathbf{r}) = u(\mathbf{r} + \mathbf{R})$ is a periodic function that can be expanded in terms of a Fourier series, given by

$$u_j(\mathbf{r}) = \sum_{\mathbf{G}} c_{j,\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.19)$$

where \mathbf{G} are the reciprocal lattice vectors and $c_{j,\mathbf{G}}$ are the Fourier coefficients. Combining Eqs. 2.18 and 2.19, we can then write the one-electron wavefunctions as

$$\phi_{j,\mathbf{k}} = \sum_{\mathbf{G}} c_{j,\mathbf{k}+\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}. \quad (2.20)$$

This provides a method through which we can model solid-state systems with periodic potentials in terms of plane waves in reciprocal space. Instead of treating infinitely many electrons, we can treat a finite number of electrons in the unit cell only, with the caveat that we must consider an infinite number of \mathbf{k} -points. It may appear that we have simply exchanged one infinity for another, but in practice, we only need to sample a finite number of \mathbf{k} -points.

We are left with the question of how to sample the infinite \mathbf{k} -space. Convergence tests can be undertaken whereby one increases the size of \mathbf{k} -point sampling until there is no change (below a chosen threshold) in a quantity of interest, such as the total energy or the band gap. A simple and straightforward scheme for sampling \mathbf{k} -space is that proposed by Monkhorst and Pack, in which a uniform grid is used.³⁸ Throughout this work, we use the notation $i \times j \times k$, where i, j and, k are integers, to represent the number of equally spaced \mathbf{k} -points for a given real-space direction, x, y , and z .

2.3 The Pseudopotential Approximation

Calculating the Kohn-Sham orbitals as a linear combination of plane waves (Eq. 2.21) is commonplace in performing solid-state calculations, because of the periodic nature of crystal lattices. This approach is used in several solid-state electronic-structure codes, such as QUANTUM ESPRESSO^{39,40} and ABINIT,^{41,42} for example. For a given KS orbital wavefunction, ϕ_a , using plane waves as our basis, we have

$$\phi_a = \sum_j c_j e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (2.21)$$

where c_j are variational coefficients, which are optimized to minimize the energy. However, one of the downsides of using plane waves as the basis set is that we require many plane waves to accurately describe both the oscillatory nature of the wavefunctions near the nuclei and the flat, near-constant behaviour between atoms. The number of plane waves that is used to express the KS orbitals and electron density is modulated by a simulation parameter called the kinetic-energy cutoff, E^{cut} , where a higher value indicates using plane waves of higher energy (more oscillations). We want the total energy (or some other property of interest) to be converged with respect to the kinetic-energy cutoff; therefore, we must perform convergence tests with respect to varying E^{cut} . Furthermore, calculations can become quickly intractable as the number of plane waves employed increases; therefore,

it is necessary that one uses an E^{cut} value that is as low as possible for a particular purpose and avoid using a value that is unnecessarily large.

The most popular computational methodology in solid-state electronic-structure theory is that of the pseudopotential approximation,⁴³ in which core electrons in an atom are represented by an effective potential while valence electrons are treated with plane waves. To justify the pseudopotential approximation from a chemist’s perspective, the valence electrons are considered to be the descriptors of chemical phenomena, such as bonding. Moreover, the chemically inert core electrons act as spectators of sorts and are not (usually) involved in chemical processes. As a result, we see an increase in computational efficiency, because we no longer have to model all of the core orbitals individually. Another benefit of the pseudopotential approximation is that we no longer require a large number of plane waves for an accurate representation of the rapidly-oscillating behaviour of the wavefunction near the core region of an atom.

We now examine several different ways to construct pseudopotentials,^{43–48} beginning with norm-conserving (NC) potentials. Pseudopotentials are commonly referred to as norm-conserving if they satisfy four general conditions.^{44,49} The first condition is that valence pseudo-wavefunctions should not contain any nodes, to ensure a smooth potential. It is important for the potential to be smooth, so that we can use a fewer number of plane waves to represent it. Secondly, the radial pseudo-wavefunctions must be equal to the radial all-electron wavefunction beyond a chosen cutoff radius, $r_{c,l}$,

$$\phi_l^{\text{PS}}(\mathbf{r}) = \phi_l^{\text{AE}}(\mathbf{r}) \text{ for } r > r_{c,l}, \quad (2.22)$$

where ϕ_l is the wavefunction with angular momentum, l , for the pseudized case (PS) and the all-electron case (AE). The second condition for norm-conserving pseudopotentials suggests that the PS wavefunction must be identical to the AE wavefunction beyond a certain distance from the nucleus. Thirdly, the norm (charge density) must be conserved within $r_{c,l}$ for both the pseudo- and the all-electron wavefunctions,

$$\int_0^{r_{c,l}} |\phi_l^{\text{PS}}(\mathbf{r})|^2 \text{d}\mathbf{r} = \int_0^{r_{c,l}} |\phi_l^{\text{AE}}(\mathbf{r})|^2 \text{d}\mathbf{r}. \quad (2.23)$$

Lastly, the pseudo- and all-electron eigenvalues of the Kohn-Sham equations must be equal,

that is

$$\epsilon_l^{\text{PS}} = \epsilon_l^{\text{AE}}, \quad (2.24)$$

which ensures that the pseudopotential is able to accurately represent the eigenvalues of orbitals in the core region.

Alternatively, the projector-augmented wave (PAW) method, developed by Blöchl,⁴⁵ was designed to combine the versatility of linear-augmented plane wave methods⁵⁰ with the simplicity of the pseudopotential approximation. PAW is fundamentally an all-electron method (operating under the frozen-core approximation),⁴⁵ which can provide the full wavefunction, in addition to the full potential, which is determined from the full charge densities. Note that the normalization condition of Eq. 2.23 is relaxed in the PAW method. In PAW schemes, a transform operator is defined that ultimately leads to a prescription for calculating the all-electron wavefunction, $|\Psi_n\rangle$, from the pseudo-wavefunction, $|\tilde{\Psi}_n\rangle$, given by⁴⁷

$$|\Psi_n\rangle = |\tilde{\Psi}_n\rangle + \sum_i (|\phi_i\rangle - |\tilde{\phi}_i\rangle) \langle \tilde{p}_i | \tilde{\Psi}_n \rangle, \quad (2.25)$$

where the sum runs over all atomic sites and the \tilde{p}_i are projector functions. The so-called AE partial waves, ϕ_i , are obtained for a reference atom and are equivalent to the PS partial waves, $\tilde{\phi}_i$, outside a chosen $r_{c,l}$, just as with NC pseudopotentials. We employ the PAW approach within this work as it has been shown to be more accurate than NC pseudopotentials in modelling molecular crystals and van der Waals dimers in the solid-state.⁵¹ However, we do use NC pseudopotentials for hybrid DFT calculations, for which the PAW method is not yet implemented in QUANTUM ESPRESSO.

2.4 The Exchange-Hole Dipole Moment Model

London dispersion interactions are the long-range, weak, attractive forces that arise from induced dipoles in fluctuating electron densities and are not properly accounted for by routine DFT methods. In particular, widely used GGA functionals depend only on the electron density and its gradient. Hence, GGA functionals are termed semi-local, and they are intrinsically unable to account for non-local dispersion physics.

Of the many ways to account for noncovalent interactions in DFT,^{52,53} the exchange-hole dipole moment (XDM) model, developed by Becke and Johnson,^{54,55} will be the focus of this section and is used extensively herein. In particular, we use XDM in this work to

accurately treat chemical systems that are heavily influenced by dispersion interactions, such as molecular crystals³² and graphite.⁵¹ XDM belongs to a class of dispersion models known as post-SCF dispersion corrections, which supplement the base DFT energy (E_{base}) with the dispersion energy, (E_{disp}):

$$E = E_{\text{base}} + E_{\text{disp}}. \quad (2.26)$$

XDM has been shown to accurately describe noncovalent interactions for a wide array of systems, such as gas-phase molecules,⁵⁶ surface adsorption,⁵⁷ layered materials,⁵⁸ and molecular crystals.³² XDM has been implemented in several electronic-structure codes such as QUANTUM ESPRESSO,^{32,39,40,51} SIESTA,^{29,59–61} and Gaussian⁶² (with the postg program).⁶³

We now take a closer look at the inner workings of the XDM model within a solid-state DFT framework. Considering a periodic lattice, the XDM dispersion energy takes the following form:^{32,51,61}

$$E_{\text{disp}} = -\frac{1}{2} \sum_{n=6,8,10} \sum_{\mathbf{L}} \sum_{ij} \frac{C_{n,ij} f_n(R_{ij,\mathbf{L}})}{R_{ij,\mathbf{L}}^n}, \quad (2.27)$$

where the $C_{n,ij}$ are the n -th order dispersion coefficients for a given atom pair, ij . Each atom pair is separated by a distance $R_{ij,\mathbf{L}}$, given by

$$R_{ij,\mathbf{L}} = |\mathbf{R}_i - \mathbf{R}_j + \mathbf{L}|, \quad (2.28)$$

which is the distance between atoms i and j separated by lattice vector \mathbf{L} . For the $\mathbf{L} = 0$ term, i cannot equal j . Moreover, in practice, the sum over the lattice vectors is truncated at a point such that the remaining interatomic contributions fall below a given energy threshold. $f_n(R_{ij,\mathbf{L}})$ is known as the Becke-Johnson (BJ) damping function that, by damping to a constant value, prevents the dispersion energy from diverging at small internuclear separations. We will examine the BJ damping function later in this section. Returning to Eq. 2.27, the first sum extends infinitely, in theory; however, in the canonical implementation of XDM it is truncated at the C_{10} term ($n = 10$). Additionally, XDM only considers pairwise interactions, as opposed to, for example, higher-order many-body terms such as the three-body Axilrod-Teller-Muto (ATM) term.^{64,65} The combination of considering dispersion coefficients up to C_{10} and two-body terms has been shown to be

sufficient for describing noncovalent interactions in intermolecular complexes,^{56,66} molecular crystals,^{32,66} and layered materials.⁵⁸

We now discuss the dispersion coefficients, $C_{n,ij}$, derived from second-order perturbation theory.^{55,67} For $n = 6, 8, 10$, they are:

$$C_{6,ij} = \alpha_i \alpha_j \frac{\langle M_1^2 \rangle_i \langle M_1^2 \rangle_j}{\alpha_j \langle M_1^2 \rangle_i + \alpha_i \langle M_1^2 \rangle_j}, \quad (2.29)$$

$$C_{8,ij} = \frac{3}{2} \alpha_i \alpha_j \frac{\langle M_1^2 \rangle_i \langle M_2^2 \rangle_j + \langle M_2^2 \rangle_i \langle M_1^2 \rangle_j}{\alpha_j \langle M_1^2 \rangle_i + \alpha_i \langle M_1^2 \rangle_j}, \quad (2.30)$$

and,

$$C_{10,ij} = \alpha_i \alpha_j \frac{2 \langle M_1^2 \rangle_i \langle M_3^2 \rangle_j + 2 \langle M_3^2 \rangle_i \langle M_1^2 \rangle_j + \frac{21}{5} \langle M_2^2 \rangle_i \langle M_2^2 \rangle_j}{\alpha_j \langle M_1^2 \rangle_i + \alpha_i \langle M_1^2 \rangle_j}. \quad (2.31)$$

Here, $\langle M_l^2 \rangle$ ($l = 1, 2, 3$) are the expectation values of the l -th order exchange-hole multipole moments and α_i are the atom-in-solid polarizabilities given by

$$\langle M_l^2 \rangle_i = \sum_{\sigma} \int \omega_i(\mathbf{r}) \rho_{\sigma}(\mathbf{r}) [r^l - (r - d_{X\sigma})^l]^2 d\mathbf{r}, \quad (2.32)$$

and

$$\alpha_i = \frac{V_i}{V_{i,\text{free}}} \alpha_{i,\text{free}}, \quad (2.33)$$

respectively. The exchange-hole dipole moment, $d_{X\sigma}$, is the distance between the reference electron and the center of its exchange hole. The value of $d_{X\sigma}$ at each reference point is obtained from the Becke-Roussel (BR)⁶⁸ exchange-hole model, due to its implicit inclusion of correlation effects, in addition to its computational efficiency.

With regards to Eq. 2.33, the free-atomic polarizability, $\alpha_{i,\text{free}}$ (the subscript free denotes *in vacuo*), is scaled by a ratio of the atom-in-solid volume to the free-atom volume. The atom-in-solid volume is given by

$$V_i = \int r^3 \omega_i(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (2.34)$$

and the free-atom volume is

$$V_{i,\text{free}} = \int r^3 \rho_{i,\text{free}}(\mathbf{r}) d\mathbf{r}. \quad (2.35)$$

The Hirshfeld weights,^{69,70} $\omega_i(\mathbf{r})$, are given by the ratio of the free-atomic electron density of atom i and the promolecular density (the sum of all free-atomic densities) and are expressed as

$$\omega_i(\mathbf{r}) = \frac{\rho_i^{at}(\mathbf{r})}{\sum_j \rho_j^{at}(\mathbf{r})}, \quad (2.36)$$

which are 1 near atom i and close to 0 everywhere else. It is necessary to use a partitioning scheme, such as the Hirshfeld weights, because we require the atom-in-solid polarizabilities and multipole moments.

We continue our discussion on the XDM dispersion model by examining the damping function first introduced in Eq. 2.27. The main purpose of introducing a damping function is to prevent the dispersion energy from diverging at small internuclear separations. The BJ damping function takes the form

$$f_n(R_{ij,\mathbf{L}}) = \frac{R_{ij,\mathbf{L}}^n}{R_{ij,\mathbf{L}}^n + R_{\text{vdW},ij}^n}. \quad (2.37)$$

A measure of the range of interaction of atomic pair ij can be given by the sum of the van der Waals radii, $R_{\text{vdW},ij}$, which is approximated as

$$R_{\text{vdW},ij} = a_1 R_{c,ij} + a_2, \quad (2.38)$$

where $R_{c,ij}$ describes the distance range at which the multipolar expansion breaks down, causing successive terms in the series expansion to be equal, as opposed to decreasing in magnitude. $R_{c,ij}$ is written in terms of the dispersion coefficients and is given by

$$R_{c,ij} = \frac{1}{3} \left[\left(\frac{C_{8,ij}}{C_{6,ij}} \right)^{1/2} + \left(\frac{C_{10,ij}}{C_{8,ij}} \right)^{1/2} + \left(\frac{C_{10,ij}}{C_{6,ij}} \right)^{1/4} \right]. \quad (2.39)$$

There are two empirical parameters present in the BJ damping function, a_1 and a_2 . These parameters are fit for use with a given density functional (and basis set), most often by minimizing binding-energy errors for the KB49 set of molecular dimers.^{55,56,71}

We conclude by looking at the XDM dispersion contributions to both the atomic forces and stress tensor. The force on atom i , is given by

$$\mathbf{F}_{\text{disp},i} = \sum_{n=6,8,10} \sum_{\mathbf{L}} \sum_j \frac{n C_{n,ij} R_{ij,\mathbf{L}}^{n-2}}{(R_{ij,\mathbf{L}}^n + R_{\text{vdW},ij}^n)^2} \mathbf{R}_{ij,\mathbf{L}}, \quad (2.40)$$

where the sum runs over all other atoms in the system, j . For the stress tensor, we consider two cartesian components, η and ξ from x, y , and z , to build the matrix. The XDM contribution to the stress tensor is

$$\sigma_{\text{disp},\eta\xi} = -\frac{1}{2V} \sum_{n=6,8,10} \sum_{\mathbf{L}} \sum_j \frac{n C_{n,ij} R_{ij,\mathbf{L}}^{n-2} (R_{ij,\mathbf{L}})_\eta (R_{ij,\mathbf{L}})_\xi}{(R_{ij,\mathbf{L}}^n + R_{\text{vdW},ij}^n)^2}, \quad (2.41)$$

where V is the unit-cell volume. The XDM contributions to the atomic forces and stress tensor assume that the $C_{n,ij}$ dispersion coefficients remain constant with respect to changes in the crystal geometry.⁵¹

2.5 Lattice Dynamics in Crystals

Lattice vibrations are responsible for a wide range of physical properties, and give rise to infrared and Raman spectra, for example. In the quantum description of solids, lattice dynamics are understood as phonons. Phonons represent the normal modes of vibration in a solid system. In the context of chemical systems studied in this work, there are two types of vibrations of interest: intramolecular and intermolecular modes. Intramolecular modes are those vibrations that occur within a molecule in the lattice and intermolecular modes occur between more than one molecule. Intramolecular vibrations are high-frequency modes (high energy) and consist of short-wavelength phonons; conversely, intermolecular modes are with low-frequency modes (low energy) and consist of long-wavelength phonons. There are two main types of phonons: acoustic and optical. Acoustic phonons are the coherent movement of atoms in the lattice out of their equilibrium positions and, due to translational invariance, are 0 at the gamma point. Secondly, there are optical phonons, which are the out-of-phase movements of atoms in the lattice. An example of an optical phonon is one atom moving to the left and the neighbouring atom moving to the right.

Phonons have defined states and energies; a proper computational treatment of these states is paramount to achieving reliable results with regards to properties such as thermal transport or, in our case, vibrational energies. We can determine the eigenvalues, ε , and eigenfunctions, Ψ , of the phonons from the Schrödinger equation:⁷²

$$\left(\sum_I \frac{\hbar^2}{2M_I} \frac{\partial^2}{\partial \mathbf{R}_I^2} + E_{GS} \right) \Psi = \varepsilon \Psi, \quad (2.42)$$

where \mathbf{R}_I and M_I is the coordinate and mass, respectively, of the I -th nucleus. E_{GS}

is the ground-state energy of the system, which, in practice, is calculated from routine electronic-structure calculations.

In order to calculate phonons, our system must be in an equilibrium configuration. The condition for the equilibrium geometry of a system requires that the forces acting on individual atoms are 0, formally expressed as

$$F_I \equiv -\frac{\partial E(\mathbf{R})}{\partial \mathbf{R}_I} = 0, \quad (2.43)$$

where $E(\mathbf{R})$ is the potential-energy surface as a function of all the nuclear coordinates, \mathbf{R} . The vibrational frequencies, ω , are determined from the Hessian of the potential-energy surface, given by

$$\det \left| \frac{1}{\sqrt{M_I M_J}} \frac{\partial^2 E(\mathbf{R})}{\partial \mathbf{R}_I \partial \mathbf{R}_J} - \omega^2 \right| = 0. \quad (2.44)$$

Calculating the equilibrium geometry and the vibrational properties of a system requires us to compute the first and second derivatives of the potential-energy surface.

We can apply the Hellmann-Feynman theorem,^{73,74} which relates the derivative of the energy with respect to some parameter, λ , to the expectation value of the derivative of the Hamiltonian (with respect to that same parameter). It is formally expressed as

$$\frac{\partial E_\lambda}{\partial \lambda} = \left\langle \Psi_\lambda \left| \frac{\partial H_\lambda}{\partial \lambda} \right| \Psi_\lambda \right\rangle, \quad (2.45)$$

where Ψ_λ is the eigenfunction of H_λ with corresponding eigenvalue E_λ . Applying the Hellmann-Feynman theorem to our definition of a system at equilibrium (Eq. 2.43), where nuclear coordinates are taken to be parameters, we have

$$F_I = - \left\langle \Psi(\mathbf{R}) \left| \frac{\partial H}{\partial \mathbf{R}_I} \right| \Psi(\mathbf{R}) \right\rangle. \quad (2.46)$$

Next, we differentiate the Hellmann-Feynman forces with respect to nuclear positions to obtain the Hessian, usually called the matrix of inter-atomic force constants (IFC), first introduced in Eq. 2.44. It can be shown that⁷²

$$\frac{\partial^2 E(\mathbf{R})}{\partial \mathbf{R}_I \partial \mathbf{R}_J} \equiv -\frac{\partial F_I}{\partial \mathbf{R}_J} = \int \frac{\partial \rho(\mathbf{r})}{\partial \mathbf{R}_J} \frac{\partial V(\mathbf{r})}{\partial \mathbf{R}_I} d\mathbf{r} + \int \rho(\mathbf{r}) \frac{\partial^2 V(\mathbf{r})}{\partial \mathbf{R}_I \partial \mathbf{R}_J} d\mathbf{r} + \frac{\partial^2 E_N(\mathbf{R})}{\partial \mathbf{R}_I \partial \mathbf{R}_J}, \quad (2.47)$$

where $V(\mathbf{r})$ and $E_N(\mathbf{R})$ are the electron-nuclear and nuclear-repulsion energies, respectively.

There are two main computational methods that are used to calculate the IFC: finite difference and density-functional perturbation theory (DFPT).⁷² In finite-difference calculations, the condition for equilibrium must first be satisfied (Eq. 2.43), then a series of small atomic displacements are made and the energy is recalculated for each. The IFC are then approximated as

$$\frac{\partial^2 E(\mathbf{R})}{\partial \mathbf{R}_I^\alpha \partial \mathbf{R}_J^\beta} \approx \frac{E_i^\alpha + E_j^\beta - 2E_0}{\Delta \mathbf{R}_i^\alpha \Delta \mathbf{R}_j^\beta}, \quad (2.48)$$

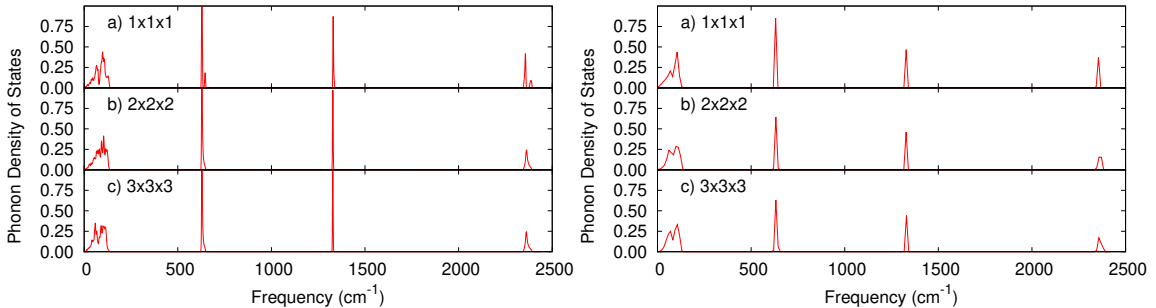
where greek superscripts indicate cartesian directions, E_i are the energies calculated for small atomic displacements, $\Delta \mathbf{R}_i$, and E_0 is the equilibrium energy. In finite-difference calculations, a supercell must be constructed from the primitive unit cell to avoid spurious self-interaction errors with atoms that are too close. Generally, large supercells are required to yield converged properties and can be upwards of hundreds of atoms, greatly increasing computational time.

DFPT, starts with the dynamical matrix, and then is Fourier transformed to yield the IFC. The eigenvalue equation for the dynamical matrix, \mathbf{D} , takes the form

$$\mathbf{D}(\mathbf{q})\epsilon(p, \mathbf{q}) = \omega^2(p, \mathbf{q})\epsilon(p, \mathbf{q}), \quad (2.49)$$

with eigenvectors $\epsilon(p, \mathbf{q})$ and eigenvalues $\omega^2(p, \mathbf{q})$, where p represents the $3N$ solutions to this equation for N atoms in the primitive unit cell. The dynamical matrix is calculated for each \mathbf{q} -vector in reciprocal space by treating the displacements of each phonon mode as a perturbation. We are now left with the task of sampling the reciprocal space, which uses the same MP uniform grid as with \mathbf{k} -points discussed in Section 2.2. The size of the MP grid is varied until convergence in some parameter, in our case, thermodynamic quantities, is reached with respect to \mathbf{q} -point sampling. Note that in both finite-difference and DFPT calculations, the IFC is Fourier interpolated (with effectively no additional computational cost) onto a denser \mathbf{q} -grid; we employ a $12 \times 12 \times 12$ dense sampling. Phonon calculations become intractable significantly more quickly than total energy calculations; therefore, we must use the smallest number of \mathbf{q} -points as possible. Figure 2.1 displays phonon density of states (phDOS) for CO_2 computed with DFPT and finite-difference calculations for increasing \mathbf{q} -point grids (supercells). At higher-frequency regimes, the phDOS peaks broaden out with denser sampling. Additionally, the change is more drastic at the lower-frequency region and it should be noted that the phDOS computed from both

Figure 2.1: Phonon density of states for CO₂ computed with DFPT (implemented in QE, left) and finite difference (implemented in Phonopy,⁷⁵ right) for various sampling grids. The level of theory used was B86bPBE-XDM, kinetic-energy and charge-density cutoffs of 80 and 800 Ry, respectively, and a $4 \times 4 \times 4$ \mathbf{k} -point grid. Symmetric sampling was used and is indicated in each plot, either via \mathbf{q} -points or supercells.



DFPT and finite difference converge to similar results at larger sampling.

In this work, we are interested in computationally modelling phonon frequencies using a variety of different electronic-structure methods, as well as comparing their performance to high-level reference calculations.

2.6 Calculating Vibrational Contributions to the Gibbs Free Energy

We are ultimately interested in predicting polymorph stability under ambient conditions (300 K, 1 atm), which necessitates computing the Gibbs free energy, G , by applying finite-temperature and pressure corrections to the total electronic energy from DFT. From standard thermodynamics, the change in Gibbs free energy for a given process can be calculated as

$$\Delta G = \Delta H - T\Delta S, \quad (2.50)$$

where T is the temperature, H is the enthalpy, and S is the entropy. Vibrational contributions to G and S can be computed by either integrating the PDOS or directly from the frequencies; the vibrational correction to H is then calculated via rearrangement of Eq. 2.50.

We begin by demonstrating the calculation of the vibrational entropy, S_{vib} , from the

phDOS, which is a straightforward integration over the first Brillouin zone, given by

$$S_{\text{vib}} = \int_0^\infty \left[-k_B \ln \left(1 - e^{-\omega/k_B T} \right) + \frac{\omega}{T} \frac{1}{e^{\omega/k_B T} - 1} \right] g(\omega) d\omega, \quad (2.51)$$

where $g(\omega)$ is the PDOS and k_B is Boltzmann's constant. Note that the integral can be exchanged for a sum over frequencies, which is equally valid. For example, we can write

$$S_{\text{vib}} = \sum_j \left[-k_B \ln \left(1 - e^{-\omega_j/k_B T} \right) + \frac{\omega_j}{T} \frac{1}{e^{\omega_j/k_B T} - 1} \right]. \quad (2.52)$$

Here, ω_j are the individual frequencies and the sum runs over all \mathbf{q} -points where frequencies have been calculated.

Next, we can calculate the Helmholtz free energy, F , given by

$$F = E_{\text{sta}} + F_{\text{vib}}, \quad (2.53)$$

where E_{sta} is the static-electronic energy (static refers to the absence of thermal contributions) and is obtained from routine ground-state DFT calculations. F_{vib} is the vibrational contribution to the Helmholtz free energy. Like the vibrational entropy calculation, it involves an integration of the PDOS, expressed as

$$F_{\text{vib}} = \int_0^\infty \left[\frac{\omega}{2} + k_B T \ln \left(1 - e^{-\omega/k_B T} \right) \right] g(\omega) d\omega. \quad (2.54)$$

The first term is known as the zero-point energy and is independent of temperature, while the second term is the thermal contribution to the F_{vib} . G can be expressed as

$$G = F + pV, \quad (2.55)$$

where p is the applied pressure (1 atm at ambient conditions) and V is the unit-cell volume. In this work, we assume the pV term to be negligible compared to F and, therefore, $G = F$. The last thermodynamic function we will discuss is the constant-volume heat capacity, given by

$$C_v = \int_0^\infty \left[k_B \left(\frac{\omega}{k_B T} \right)^2 \frac{e^{-\omega/k_B T}}{(e^{-\omega/k_B T} - 1)^2} \right] g(\omega) d\omega. \quad (2.56)$$

In Ch. 3, we use this equation to compare directly to experimental data for the heat capacity.

2.7 The Quasi-Harmonic Approximation

Routine DFT geometry optimization calculations are performed at *effectively* 0 K. Here, we mean that there is no consideration of finite-temperature or zero-point energy contributions, which both arise due to phonons. Moreover, the unit-cell volumes of crystals change with respect to temperature. Most materials undergo thermal expansion, that is, increasing the temperature expands the crystal; however, there are materials which can undergo negative thermal expansion (compression). Changes in volume due to thermal expansion are neglected in standard DFT optimizations and must be accounted for to compare directly to experimental cell geometries. The simplest way to account for thermal effects is the harmonic approximation (HA), which can be improved by a more accurate and sophisticated approach, the quasi-harmonic approximation (QHA).⁴

We first examine the harmonic approximation, which assumes that the phonon density of states does not change with respect to volume. Within the HA, phonon frequencies are calculated only for the equilibrium volume and the equations of Section 2.6 are used to calculate the free-energy contributions. However, it is strictly false that phonon frequencies do not change with volume and, for temperatures approaching the melting point, the HA breaks down significantly.²³ While there are successes of the HA, there are also failures arising from this model, such as predictions of infinite thermal conductivity and zero thermal expansion.²⁵ The QHA builds upon the HA by assuming that the latter holds for any given crystal geometry, even if it is not the equilibrium geometry. The QHA provides a prescription to account for changes in volume due to phonons, which we consider next.

In a typical QHA run, we first must generate an energy versus volume curve, $E(V)$. This is usually achieved by an unconstrained optimization of the cell volume and atomic positions, followed by modifying the unit-cell volume only and performing a series of fixed-volume relaxations. For example, one could choose to increase and decrease the cell volume by +/- 10% in increments of 1%. In QUANTUM ESPRESSO, the `cell.dofree='shape'` keyword is used to allow all atomic coordinates and lattice parameters to relax under the constraint that the cell volume must remain constant. Phonon calculations are run at all the geometries and the F_{vib} (Eq. 2.54) is computed and plotted versus the grid of volumes. From here, we fit this data using linear regression and the thermal pressure,

$$p_{\text{th}} = -\frac{\partial F_{\text{vib}}}{\partial V}, \quad (2.57)$$

is extracted as the slope, which in most cases will be a negative value (expansion of the crystal). We then apply the thermal pressure to the equilibrium structure and perform a final geometry optimization. Applying p_{th} is how we account for volume changes arising from temperature, which is necessary to compare DFT results directly to experimental geometries.

2.8 Low-Cost Computational Methods for Modelling Solid-State Systems

It quickly becomes prohibitive to calculate phonon frequencies for very large systems, especially those that may be found in a CSP study. Therefore, it is desirable to have computationally expedient models to carry out these large calculations. One way to approach this problem is by using semi-empirical methods. Semi-empirical methods make approximations and use fitted parameters to replace many of the mathematical terms with predetermined values that would normally be evaluated by explicit integration over the orbitals or electron density. As a result, there can be different parameterizations of semi-empirical methods that are suited for a variety of purposes, such as parameter sets being optimized for organic molecular crystals or inorganic network solids. Two semi-empirical methods that we are interested in assessing are density-functional tight binding (DFTB)²⁸ and the HF-3c method, corresponding to minimal-basis Hartree-Fock with three empirical corrections.²⁷

DFTB is based on a Taylor expansion of the Kohn-Sham DFT energy functional, assuming a given reference density, ρ_0 . The ground-state density is represented as the reference, perturbed by density fluctuations: $\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \delta\rho(\mathbf{r})$. The Taylor expansion of the functional (up to third order) is given by

$$E^{\text{DFTB3}}[\rho_0 + \delta\rho] = E^0[\rho_0] + E^1[\rho, \delta\rho] + E^2[\rho, (\delta\rho)^2] + E^3[\rho_0, (\delta\rho)^3]. \quad (2.58)$$

Because we use the DFTB3 model in this work, we will center our discussion around this framework. The DFTB3 total energy is given by

$$E^{\text{DFTB3}} = \sum_i \sum_{AB} \sum_{\mu \in A} \sum_{\nu \in B} n_i c_{\mu i} c_{\nu i} H_{\mu\nu}^0 + \frac{1}{2} \sum_{AB} V_{AB}^{\text{rep}} + \frac{1}{2} \sum_{AB} \Delta q_A \Delta q_B \gamma_{AB}^h + \frac{1}{3} \sum_{AB} \Delta q_A^2 \Delta q_B \Gamma_{AB}. \quad (2.59)$$

In the first term we have the contributions arising from the Hamiltonian matrix elements,

$H_{\mu\nu}^0$. The Hamiltonian matrix elements only consider two-center terms, which are pre-computed once as a function of inter-atomic distance for all pairs of atoms, A and B. The sum runs over all molecular orbitals (MOs) with occupation numbers, n_i . Here, $c_{\mu i}$ and $c_{\nu i}$ are the expansion coefficients of the i th MO in terms of a minimal, valence-only basis set of atomic orbitals (AOs). The index μ is used for the sum over AOs on atom A, and the index ν is used for the sum over AOs on atom B. The second term gives the repulsion between each pair of atoms, A and B, which is parameterized by fitting to either DFT calculations⁷⁶ or empirical data.⁷⁷ The remaining two terms correspond to the second- and third-order terms in the Taylor expansion that involve density (or charge) fluctuations. The atomic charge fluctuations, relative to the free, neutral atoms are $\Delta q_A = q_A - Z_A$, where q_A is the electron population and Z_A is the nuclear charge. In practice, Δq_A is evaluated as the Mulliken charge of atom A. The analytical function γ_{AB} approaches $1/R_{AB}$ at large separations to recover the Coloumb interaction between the two charges, while at short distances it represents the electron-electron interactions within a single atom. Finally, the Γ_{AB} function in the third-order term is the first derivative of γ_{AB} with respect to charge.

In terms of computational efficiency, DFTB methods are several orders of magnitude faster than plane-wave DFT; however, there are several limitations of the core DFTB model.²⁸ For example, since DFTB is fundamentally based on DFT, noncovalent interactions are not described and it is therefore necessary to use a dispersion correction, such as Grimme’s D3 model.⁷⁸

We now examine the HF-3c method. The general form for HF-3c is given by

$$E^{\text{HF-3c}} = E^{\text{HF/MINIX}} + E_{\text{disp}}^{\text{D3}} + E_{\text{BSSE}}^{\text{gCP}} + E_{\text{SRB}}. \quad (2.60)$$

The first term is the HF energy computed with the MINIX⁷⁹ basis set, the second term is the geometrical counterpoise correction (gCP)⁸⁰ used to correct for basis-set superposition error, and the last term is the short-ranged basis-set incompleteness correction (SRB).²⁷ In total, there are nine empirically determined parameters in the HF-3c method, three for the D3 dispersion, four in the gCP scheme, and two for the SRB correction. Since HF-3c uses a minimal basis set, it will be considerably faster computationally compared to DFT, although slower in general than DFTB due to the requirement of performing a full HF calculation. The HF-3c method was specifically tailored and benchmarked for molecular crystal applications; therefore, it is a desirable option for our study on molecular crystals.

CHAPTER 3

The Relative Stability of Graphite and Diamond

This chapter was adapted from: White, M. A.; Kahwaji, S.; Freitas, V. L.; Siewert, R.; Weatherby, J. A.; Ribeiro da Silva, M. D.; Verevkin, S. P.; Johnson, E. R.; Zwanziger, J. W. The Relative Thermodynamic Stability of Diamond and Graphite. *Angew. Chem. Int. Ed.* **2021**, *60*, 1546–1549.

3.1 Introduction

Allotropes are different physical forms in which an element can exist. For example, graphite and diamond are two common allotropes of carbon. Other notable allotropes of carbon are graphene, carbon nanotubes, and buckminsterfullerenes (C_{60}). What separates these allotropes apart from one another is their chemical bonding arrangement, and thus their physical properties. For example, diamond consists of a σ -bonding arrangement of carbon atoms, whereas graphite is a layered material of sp^2 -hybridized carbons. As a result, diamond is an exceptionally hard material that is not a good conductor of electricity, whereas graphite is softer and an excellent conductor of electricity.

At standard temperature and pressure, graphite is more stable than diamond and this is conventionally thought to be true for 0 K as well. However, a recent density-functional theory (DFT) study by Grochala on the relative stability of graphite and diamond found that, as temperatures approached 0 K, the two allotropes became degenerate.⁸¹ In an effort to clarify this discrepancy, we undertook a collaborative project with several experimental groups, led by Prof. M. A. White,⁸² to determine whether graphite or diamond is the more stable allotrope of the two.

Experimental determination of the relative stability of graphite and diamond comes from thermodynamics measurements. Here, the relative enthalpy values (ΔH) are derived from the heats of combustion of graphite and diamond, while the relative entropy values are obtained from the integrated heat capacities. However, there are several complications that arise in the determination of the relative thermodynamics of graphite and diamond, specifically the absence of heat capacity data for diamond at very low temperatures, as well as the existence of relatively few combustion measurements for diamond leading to experimental uncertainty. Moreover, thermodynamic differences consist of small differences between very large numbers, making exceptional accuracy a must.

From our computational perspective employed in this work, we aim to supplement the experimental results by performing high-level DFT calculations with various functionals and dispersion methods, as well as several solid-state software packages. In order to assess thermodynamic stability, we must compute relative free energies, which are obtained within the quasi-harmonic approximation (QHA)⁴ for a more accurate determination of contributions arising from thermal expansion. The combination of experimental and theoretical work presented here aims to provide conclusive evidence of the relative thermodynamic stability of graphite and diamond.

3.2 Computational Methodology

Density-functional theory geometry optimizations were performed starting from experimental structures of graphite ($c = 6.696 \text{ \AA}$)^{83,84} and diamond⁸³ (COD ID: 9008564), using either the PBE³⁰ or the B86bPBE^{30,31} generalized-gradient approximation (GGA) functional, paired with the exchange-hole dipole moment (XDM)³² dispersion correction. Single-point energies at the equilibrium geometries were evaluated using several hybrid functionals: a B86bPBE-based hybrid with 25% exact exchange (B86bPBE-25X),³⁵ PBE0,³⁴ and HSE06,⁸⁵ all paired with XDM dispersion. The Becke-Johnson damping parameters used with each functional were the same as in Ref. 35.

All calculations were performed with the QUANTUM ESPRESSO (QE)^{39,40} software package (v 6.4), using periodic-boundary conditions and the planewaves/pseudopotentials approach. Projector augmented-wave (PAW)⁴⁵ datasets were used for geometry optimizations and norm-conserving (NC)⁴⁴ pseudopotentials were used for single-point energy

calculations with the hybrid functionals. Convergence thresholds of 10^{-5} Ry and 10^{-4} Ry/Bohr in the energy and forces, respectively, were used. The kinetic-energy and charge-density cutoffs were set to 80 and 800 Ry, respectively, for PAW and to 80 and 320 Ry, respectively, for NC calculations. Uniform \mathbf{k} -point meshes (Monkhorst-Pack,³⁸ offset from the origin) with $8 \times 8 \times 4$ and $5 \times 5 \times 5$ points were used for graphite and diamond, respectively. These values were chosen by imposing a convergence criterion of less than 1 mRy in the energy with respect to the kinetic-energy cutoff and \mathbf{k} -point grid.

Phonon frequencies were calculated from density-functional perturbation theory (DFPT)^{40,72} using the same methodology as for the geometry optimizations. Uniform \mathbf{q} -point grids of $3 \times 3 \times 1$ and $2 \times 2 \times 2$ were used for graphite and diamond, respectively, to satisfy a convergence criterion of 0.01 kcal/mol (0.04 kJ/mol) in the vibrational contribution to the Helmholtz free energy (F_{vib}). Relaxation and phonon calculations were carried out for a range of fixed cell volumes, centered about the equilibrium volume (spanning from -5% to +10% for graphite, and -10% to +10% for diamond, in 5% increments) to obtain the thermodynamic quantities of interest within the quasi-harmonic approximation.⁴ At each volume, the vibrational contribution to the Helmholtz energy was determined from integrating the computed phonon density of states. The thermodynamic properties (enthalpies, H ; Gibbs energies, G ; entropies, S ; constant-volume heat capacities, C_v) were then determined as functions of temperature, over the range of 0-400 K. At each considered temperature, T , the Gibbs energy was computed as a sum of the electronic energy, E_{el} obtained from the hybrid functional, and the thermal correction obtained from the GGA,

$$G(V, T) = E_{\text{el}}(V) + F_{\text{vib}}(V, T), \quad (3.1)$$

and was minimized with respect to volume. The other thermodynamic properties were then evaluated at that optimum volume, through interpolation of the corresponding quantities from the volume grid.

3.3 Results and Discussion

First, we compare the computed unit-cell geometries for both graphite and diamond at 0 K (static conditions) and 298 K (quasi-harmonic approximation) to the experimental reference, shown in Table 3.1. From these results, we see that both B86bPBE-XDM and PBE-XDM are able to accurately reproduce the cell parameters for graphite as well as

diamond at 0 K. With regards to the c lattice parameter for graphite, B86bPBE-XDM gives a value very close to the experimental reference. However, upon applying the QHA, we see an overestimation in the lattice parameters, due to the negative thermal pressure. The largest change due to the QHA is seen in the c parameter for graphite (as opposed to a or b) since this direction of the unit cell is dominated by weak dispersion interactions and, therefore, any applied pressure would easily affect it.

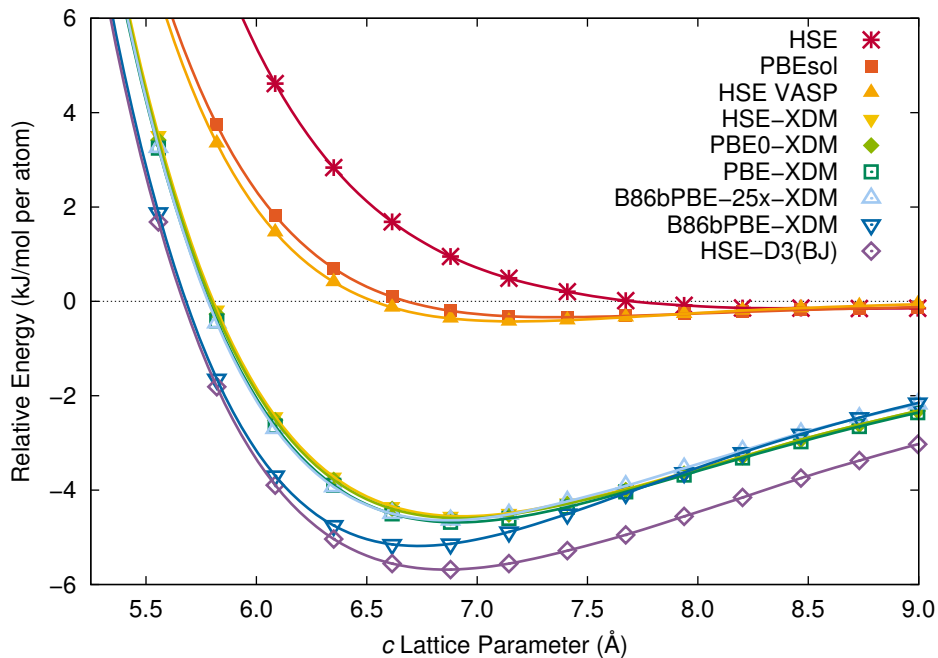
Table 3.1: B86bPBE-XDM/PAW and PBE-XDM/PAW optimized cell parameters (\AA) and volumes (\AA^3) for graphite and diamond. DFT relaxations were performed in the absence (0 K) and presence (298 K) of applied negative thermal pressure. Experimental data, at a temperature of 298 K, are given for comparison.^{83,84}

Quantity	B86bPBE-XDM		PBE-XDM		Expt. 298 K
	0 K	298 K	0 K	298 K	
Graphite a, b	2.459	2.460	2.460	2.461	2.456
Graphite c	6.695	6.781	6.838	6.929	6.696
Graphite V	35.07	35.55	35.85	36.35	34.98
Diamond a	2.515	2.525	2.516	2.526	2.522
Diamond V	11.25	11.39	11.27	11.40	11.34

Figure 3.1 shows the potential energy curves for the graphite exfoliation obtained with several dispersion-corrected DFT methods, as well as those used in a previous study by Grochala.⁸¹ In that work, HSE06 single points were computed on PBEsol geometries (HSE06//PBEsol), omitting the use of a dispersion correction. We see that the absence of a dispersion correction yields extremely weak binding between the graphite layers and therefore overly large c lattice parameters. Adding a dispersion correction reduces the c lattice parameter, imparts an approximately additional 5 kJ/mol per carbon stabilization to graphite, and is needed for a reasonable treatment of the relative thermodynamic stability of diamond and graphite. Since graphite is more stable than diamond with $\Delta G = -3.17 \pm 0.15$ kJ/mol experimentally,⁸² diamond would be calculated to be the thermodynamic ground state, without the dispersion stabilization from interlayer interactions, provided that an accurate base functional is used. This result is in agreement with a recent analytical model, which predicted diamond to be more thermodynamically stable than graphene when accounting for the graphite exfoliation energy.⁸⁶

We are comparing our computationally determined values to experimental results

Figure 3.1: Exfoliation curves of graphite computed with selected density functionals and dispersion corrections comparing methods used in this work with those from Ref. 81. All calculations used the QUANTUM ESPRESSO program and the same parameters as described in Section 3.2, except one of the HSE curves, where the VASP program was used, as noted in the key.



calculated recently by White et al.⁸² using state-of-the-art thermodynamic measurement techniques. Table 3.2 shows experimental values for ΔH , ΔS , and ΔG at both 0 K and 298.15 K.

Table 3.2: Experimentally determined values for ΔH , ΔS , and ΔG for graphite relative to diamond, at 0 K and 298.15 K. All values are in J/mol, except S , which is in J/mol/K.

Temperature (K)	ΔH	ΔS	ΔG
0	-2690±150	0	-2690±150
298.15	-2150±150	3.436±0.030	-3170±150

Table 3.3 contains energy differences for graphite and diamond obtained with various DFT models, relative to diamond (a negative value indicates graphite is lower in energy, i.e. more stable). From these results, B86bPBE and PBE functionals paired with the XDM dispersion correction yield a large energy difference between graphite and diamond of around -10 kJ/mol. This overstabilization can be attributed to density-functional delocalization

error^{12,87} which is known to preferentially stabilize delocalized electron densities; this is the case for graphite, which contains a conjugated π -system of sp^2 -hybridized carbons. While delocalization error is more commonly seen for charge-transfer complexes, its effects have also been observed in molecules with extended conjugation, notably manifesting in overstabilization of cumulenes relative to poly-enes,⁸⁸ overstabilization of the ring isomer of C_{20} relative to the cage and bowl forms,⁸⁹ and overstabilization of delocalized carbon-centered radicals.^{90,91} Thus, it is expected that delocalization error will preferentially stabilize the delocalized electrons in the conjugated π -system of graphite relative to the more localized electrons in the σ -bonding network of diamond. As is characteristic of delocalization error, improved performance can be obtained with hybrid functionals (B86bPBE-25X, PBE0, and HSE06). Indeed, using the B86bPBE-25X-XDM and PBE0-XDM dispersion-corrected hybrid functionals afford a 1 kJ/mol accuracy of the experimental values for ΔG of -2.69 kJ/mol at 0 K and -3.17 kJ/mol at 298 K.⁸² To our knowledge, this is the first demonstration of density-functional delocalization error in graphite.

Table 3.3: Electronic (ΔE_{el}) and Gibbs (ΔG) energies of graphite, relative to diamond, obtained with selected functionals and XDM dispersion corrections. Energies are in kJ/mol per carbon atom. The geometries were obtained from relaxations in the absence (0 K) or presence (298 K) of applied negative thermal pressure. The zero-point energy contributions are -0.96 and -0.95 kJ/mol per atom for B86bPBE-XDM and PBE-XDM, respectively, which correspond to the differences between $\Delta G(0)$ and $\Delta E_{el}(0)$.

Functional	Pseudo	$\Delta E_{el}(0)$	$\Delta G(0)$	$\Delta E_{el}(298)$	$\Delta G(298)$
B86bPBE-XDM Geometries					
B86bPBE-XDM	PAW	-8.72	-9.69	-8.63	-10.07
B86bPBE-XDM	NC	-8.63	-9.59	-8.43	-9.87
B86bPBE-25X-XDM	NC	-2.07	-3.03	-2.21	-3.65
PBE0-XDM	NC	-1.33	-2.29	-1.52	-2.96
HSE-XDM	NC	-2.47	-3.43	-2.68	-4.12
PBE-XDM Geometries					
PBE-XDM	PAW	-8.84	-9.79	-8.77	-10.23
PBE-XDM	NC	-8.62	-9.56	-8.47	-9.92
B86bPBE-25X-XDM	NC	-2.06	-3.00	-2.16	-3.61
PBE0-XDM	NC	-1.35	-2.29	-1.50	-2.96
HSE-XDM	NC	-2.52	-3.46	-2.69	-4.14

It has been shown previously by Grochala that HSE06 single-point energies on PBEsol

geometries for graphite and diamond predicted diamond to have a lower electronic energy than graphite in contrast to experiment.⁸¹ To put this result in context, we performed additional calculations of the diamond-graphite electronic-energy difference with a range of common DFT methods using three different solid-state codes: QUANTUM ESPRESSO, ABINIT,^{41,42} and VASP.⁹² The results are shown in Table 3.4. The three codes give comparable results using the PBE functional paired with the D3 dispersion correction (with either zero or Becke-Johnson damping). While zero-damping imparts slightly greater stability to graphite, the D3(BJ) results are in close agreement with the PBE-XDM results in Table 3.3. Finally, using the PBE0 hybrid functional improves agreement with experiment by reducing delocalization error, in agreement with the XDM-corrected hybrids. Interestingly, it has been found that, using the HSE06//PBEsol approach, diamond is only lower in energy with the VASP code, as used in Ref. 81. Analogous calculations using both the ABINIT* and QUANTUM ESPRESSO codes have shown graphite to be lower in energy, which indicates there may be differences in the HSE06 implementation. The latter two results are in agreement with previous HSE06 results in the literature,⁹³ in addition to all other functionals considered here.

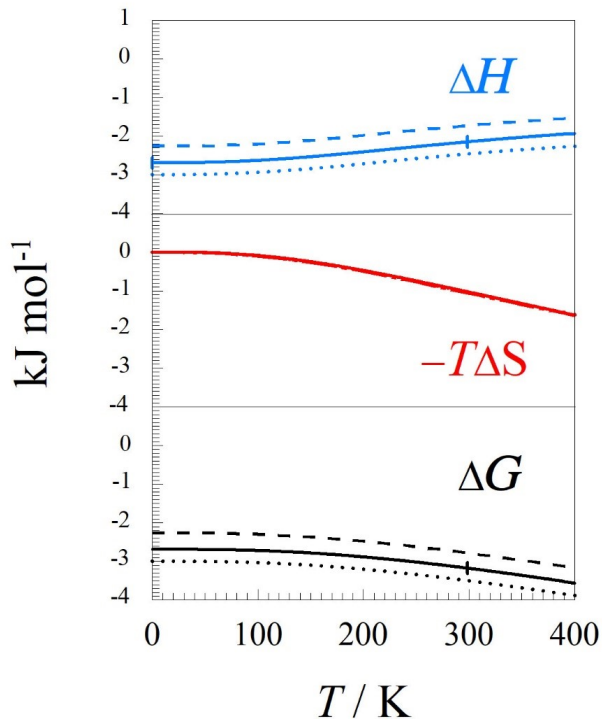
Table 3.4: Ground state electronic energy of graphite, relative to diamond, based on different DFT models. For each entry, the model used for structural relaxation is given first, then the model used for the single-point energy, and finally the code.

Model	$\Delta E(0 \text{ K})$ (kJ/mol)
PBE-D3/PBE-D3/ABINIT	-11.7
PBE-D3/PBE-D3/VASP	-10.8
PBE-D3/PBE-D3/QE	-10.8
PBE-D3(BJ)/PBE-D3(BJ)/ABINIT	-8.6
PBE-D3(BJ)/PBE-D3(BJ)/VASP	-8.5
PBE-D3(BJ)/PBE-D3(BJ)/QE	-8.4
PBEsol/HSE06/ABINIT	-7.8
PBEsol/HSE06/VASP	0.6
PBEsol/HSE06/QE	-6.4
PBE-D3/PBE0-D3/QE	-3.0
PBE-D3(BJ)/PBE0-D3(BJ)/QE	-1.1

Figure 3.2 displays the relative thermodynamic functions, ΔH , $-T\Delta S$, and ΔG for

*ABINIT and VASP calculations were performed by Prof. Josef W. Zwanziger

Figure 3.2: Thermodynamic energy differences for graphite and diamond, as a function of temperature at ambient pressure. Solid lines are experimental results, shown with associated uncertainties (too small to be visible for $-T\Delta S$). Dotted lines are B86bPBE-25X-XDM//B86bPBE-XDM results and dashed lines are PBE0-XDM//PBE-XDM calculated data.



graphite and diamond computed with the two XDM-corrected hybrid functionals, B86bPBE-25X-XDM//B86bPBE-XDM and PBE0-XDM//PBE-XDM, which are expected to be the most accurate due to minimization of delocalization error. For ΔH and ΔG , we see that the calculated thermodynamic functions bracket the experimental results and are displaced by about $2\times$ the experimental uncertainty. Moreover, $-T\Delta S$ compares very well to experiment with both methods due to the heat capacities being computed accurately. Ultimately, these results show excellent agreement with experiment, providing supporting and conclusive evidence that graphite is in fact more thermodynamically stable than diamond, contrary to previously reported results.⁸¹

3.4 Summary

In this chapter, we have provided state-of-the-art high-level computational evidence that, together with conclusive experimental data, support the claim that graphite is more thermodynamically stable than diamond for $T < 400\text{K}$. These findings corroborate the conventional knowledge and also provide up-to-date values for the relative free energies and enthalpies.

The importance of modelling dispersion interactions in graphite is highlighted and it is found that omitting the use of a dispersion correction will lead to diamond being calculated as the lower-energy allotrope with hybrid functionals. We found that inclusion of a dispersion model afforded an approximately 5 kJ/mol per C atom stabilization resulting from dispersion, and neglect of this is enough to yield diamond as more stable. Therefore, it is necessary to account for dispersion interactions when predicting the relative stability of graphite and diamond.

Interestingly, the previous electronic energy ordering at 0 K reported by Grochala⁸¹ was reproduced in the VASP program (diamond being more stable than graphite); however, analogous calculations using QUANTUM ESPRESSO and ABINIT were not able to reproduce those results using the same HSE06//PBEsol methodology. Additionally, none of the other dispersion-corrected models we used were able to reproduce the ordering from Ref. 81. As a result, we suggest that there may be a difference in the HSE06 implementation in the VASP program compared to QE or ABINIT.

Delocalization error, the tendency for GGA models to overstabilize delocalized electron densities, is believed to manifest in graphite, which is rationalized due to the in-plane sp^2 -hybridization of the carbon atoms, allowing for a conjugated π -system to exist. Furthermore, calculations performed with hybrid functionals vastly reduce the energy difference between graphite and diamond at 0 K from -8.72 to -2.07 kJ/mol per carbon atom (using B86bPBE-XDM and B86bPBE-25X-XDM, respectively), supporting our hypothesis.

Using the B86bPBE-25X-XDM//B86bPBE-XDM and PBE0-XDM//PBE-XDM methods, the ΔG (298 K, relative to diamond) was computed to be -3.65 and -2.96 kJ/mol per C atom, respectively, compared to the experimental value of -3.17 kJ/mol per C atom. By calculating contributions arising from thermal effects, the thermodynamic functions for

graphite and diamond were computed using the aforementioned methods and we found that the ΔS function is nearly identical to the experimental findings, indicating that our methods are able to accurately represent the heat capacity. Similarly, for ΔG and ΔH , the computational results bracket the experimental data within only $2\times$ the experimental uncertainty.

CHAPTER 4

Crystal-Energy Landscapes and Polymorphism of Functionalized [6]Helicene Structures

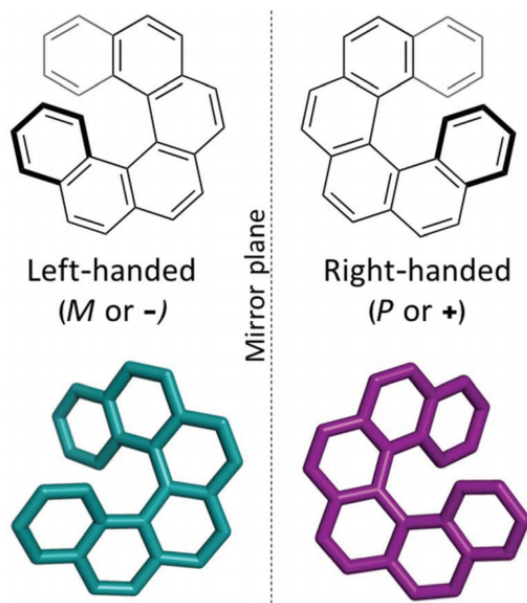
This chapter was adapted from: Schmidt, J. A.; **Weatherby, J. A.**; Sugden, I. J.; Santana-Bonilla, A; Salerno, F; Fuchter, M. J.; Johnson, E. R.; Nelson, J; Jelfs, K. E. Computational Screening of Chiral Organic Semiconductors: Exploring Side-Group Functionalization and Assembly to Optimize Charge Transport *Cryst. Growth Des.* **2021**, *21*, 5036–5049.

In this chapter, we focus on JW’s contribution to the above manuscript, in addition to the conclusions that can be drawn in the broader scope of this thesis. All DFT calculations under the SIESTA framework were performed by JW. The computational screening protocol and generation of CSP crystal structures, and electron mobility calculations were performed by JS.

4.1 Introduction

Helicenes are chiral structures which generally consist of fused benzene rings with varying substitution patterns decorating a given backbone. One minimal example is the chiral [6]helicene structure which contains 6 benzene rings, adopting a helical structure. The helical nature of [6]helicene gives rise to two enantiomers: *M*- and *P*-[6]helicene, as shown in Fig. 4.1. Moreover, the crystallized version of [6]helicene can either exist as an enantiopure structure, containing only one of the two enantiomers, or as a racemic structure, containing both enantiomers. Helicenes show promise as low-cost organic semiconductors (OSCs) and have been used in applications such as chiroptical switches,⁶ circularly polarized (CP)

Figure 4.1: The two enantiomers of the chiral [6]helicene molecule, shown in a top-down view (bold lines indicate bonds projecting towards the viewer). Image obtained from Ref. 94.

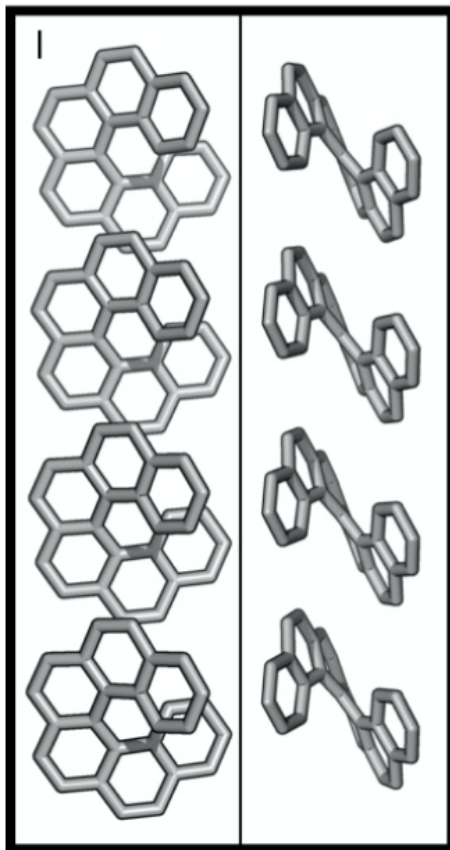


organic light-emitting diodes,⁹ CP light photodetectors,⁷ and transistor materials.⁸

The 3-dimensional solid-state crystal packing of helicenes is highly dependant on the molecular geometry. Noncovalent interactions, such as $\pi - \pi$ stacking, are responsible for the unique arrangement that the molecules can adopt in the solid state. The crystal energy landscapes of helicenes contain many local minima corresponding to subtle distortions in the packing motifs that give rise to polymorphism (as is the case with many molecular crystals). Additionally, properties such as charge-carrier mobility for electronic applications are heavily dependent on the crystal packing. The relationship between molecular structure, crystal packing, and charge-carrier mobility is the underlying principle of this work.

We aim to perform an exhaustive survey of the substitution landscape to qualitatively assess the relative transport properties of substituted *M*-[6]helicene crystals. The *M*-[6]helicene backbone was populated via mono- and di-substitution patterns on all possible hydrogen positions (di-substitution only occurred on symmetry equivalent positions). Functional groups such as halogens, alcohols, esters, ethers, carboxylic acids, phenyl groups, and aryl groups were selected—resulting in 1344 structures. In order to screen these hypothetical structures for their suitability as OSC materials, we first assume that a simple translational

Figure 4.2: Image displaying the translational dimer motif that is commonly observed in [6]helicene crystal structures. Image obtained from Ref. 95.



dimer model is sufficient to describe the transport properties of the given material.* The translational dimer has been found to be the predominant packing motif in helicene crystals and is displayed in Fig. 4.2. The four compounds with translational dimers that possessed the most favourable predictors of high charge mobility were selected for more extensive study, which allows us to assess the validity of the translational-dimer model in this work.

Upon screening for desirable transport properties, first-principles crystal structure prediction (CSP) was conducted on the four best substituted-helicene candidates. Crystal structures were generated via the CrystalPredictorII program⁹⁶ followed by ranking of the energetics using the DMACRYS force-field code.⁹⁷ A subset of the list of structures is then passed to density-functional theory (DFT) calculations for accurate energy ranking. Here, we are operating under the zeroth-order CSP assumption,⁹⁸ which only considers electronic (or lattice) energies and omits thermal effects. Problems for zeroth-order CSP

*For a detailed discussion on the metrics used to assess OSC suitability, please see Ref. 95.

arise when polymorphs are ranked very close in energy (in the range of 0-3 kJ mol⁻¹) and thermal effects can reorder these low-lying structures (see Ch. 5).^{24,99} Additionally, the DFT energies were evaluated for the optimized force-field geometries, which are assumed to be a good approximation to the DFT optimized structures. While this assumption could lead to errors in correctly ordering the structures energetically, it has been shown to yield reasonable crystal-energy landscapes for pharmaceutical compounds.⁵ We are limited by computational resources and it would be intractable to run full plane-wave DFT optimizations on all the candidate helicene structures.

For our DFT study, we employed the SIESTA method,²⁹ which has been shown to yield higher accuracies in generating crystal-energy landscapes compared to distributed-multipole force fields.⁵ We assess the validity of the translational dimer model as a common packing motif for helicene crystals by noting its overwhelming presence in the lowest-energy structures generated from CSP. Furthermore, for 3 out of the 4 compounds studied, the crystal-energy landscapes show one or more candidates possessing energies within 1 kJ/mol of the lowest-energy polymorph, indicating that DFT geometry optimizations and/or inclusion of thermal effects could easily lead to reordering. Finally, charge mobilities[†] were computed for the 10 lowest-energy structures of each of the four helicenes, and two of these compounds were identified as promising targets for future experimental studies.

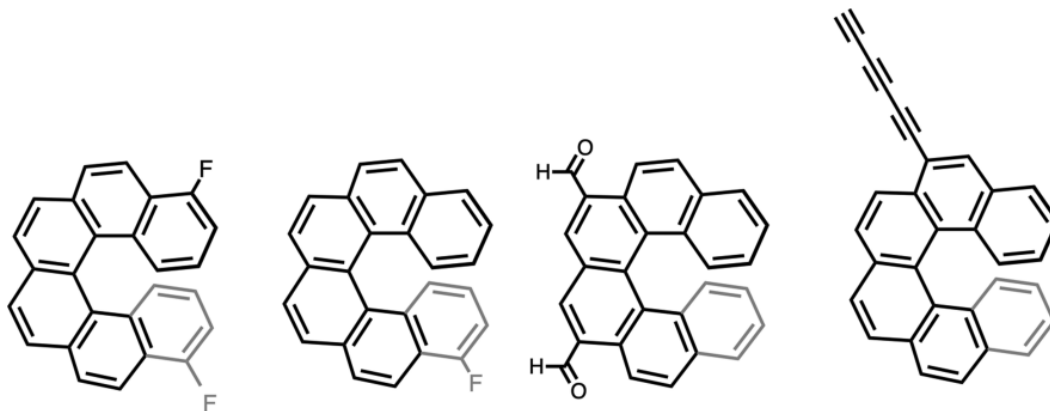
4.2 Computational Methodology

SIESTA Calculations

Density-functional theory (DFT)¹⁰⁻¹² single-point energy calculations on the lower-energy hypothetical polymorphs identified in the CSP searches were performed with an in-house version of the SIESTA^{29,59} code (4.0b-485) to generate crystal-energy landscapes. The functionalized helicenes were selected on the basis of their suitability as organic semiconducting materials as discussed in Ref. 95. The B86bPBE generalized-gradient approximation (GGA) functional was used with the exchange-hole dipole moment (XDM)⁶¹ dispersion model. Parameters for the Becke-Johnson damping function were set to $a_1 = 0.5000$ and $a_2 = 2.5556$ Å. Troullier-Martins norm-conserving pseudopotentials^{44,49} were employed, along with a double-zeta plus polarisation (DZP) basis set of atom-centered numerical orbitals, with the confinement radius set to 0.001 Ry. Additional calculation parameters

[†]Electron mobility calculations were performed by Dr. Julia Schmidt.

Figure 4.3: Four promising OSC molecular materials identified by computational screening. From left to right: 4,13-difluoro[6]helicene, 13-fluoro[6]helicene, 7,10-dialdehyde[6]helicene, and 6-triacetylene[6]helicene.



were a $4 \times 4 \times 4$ Monkhorst-Pack³⁸ \mathbf{k} -point mesh and a real-space integration grid cutoff value of 400 Ry.

4.3 Results and Discussion

From the OSC suitability screening protocol, four promising substituted helicenes were identified and carried forward to CSP analysis. The four compounds are 4,13-difluoro[6]helicene, 13-fluoro[6]helicene, 7,10-dialdehyde[6]helicene, and 6-triacetylene[6]helicene. Their molecular diagrams are shown in Fig. 4.3. The lowest-energy crystal structures within 15 kJ/mol of the DMACRYS energy minimum from the CSP search were carried forward to DFT single-point calculations for higher-accuracy ranking of the candidates. We computed crystal-energy landscapes for each of the four compounds relative to the minimum energy structure and have noted the structures which contain the translational dimer motif, which is postulated to be a favorable configuration for high charge-carrier mobility.

We now examine the DFT crystal-energy landscapes of the four substituted helicene compounds, given in Fig. 4.4, in more detail. Furthermore, the minimum-energy crystal structures are shown in Fig. 4.5. Lastly, electron mobility calculations were carried out for the 10 lowest-energy structures from each landscape and are presented in Fig. 4.6.

Figure 4.4: Crystal-energy landscapes for 13-fluoro[6]helicene (A), 4,13-difluoro[6]helicene (B), 6-triacetylene[6]helicene (C), and 7,10-dialdehyde[6]helicene (D). Energies are plotted relative to the B86bPBE-XDM minimum-energy structure. Data points are shown in blue (red) if they do (do not) contain the translational dimer motif. Low-energy crystals are numbered in order of ascending relative energy from the minimum.

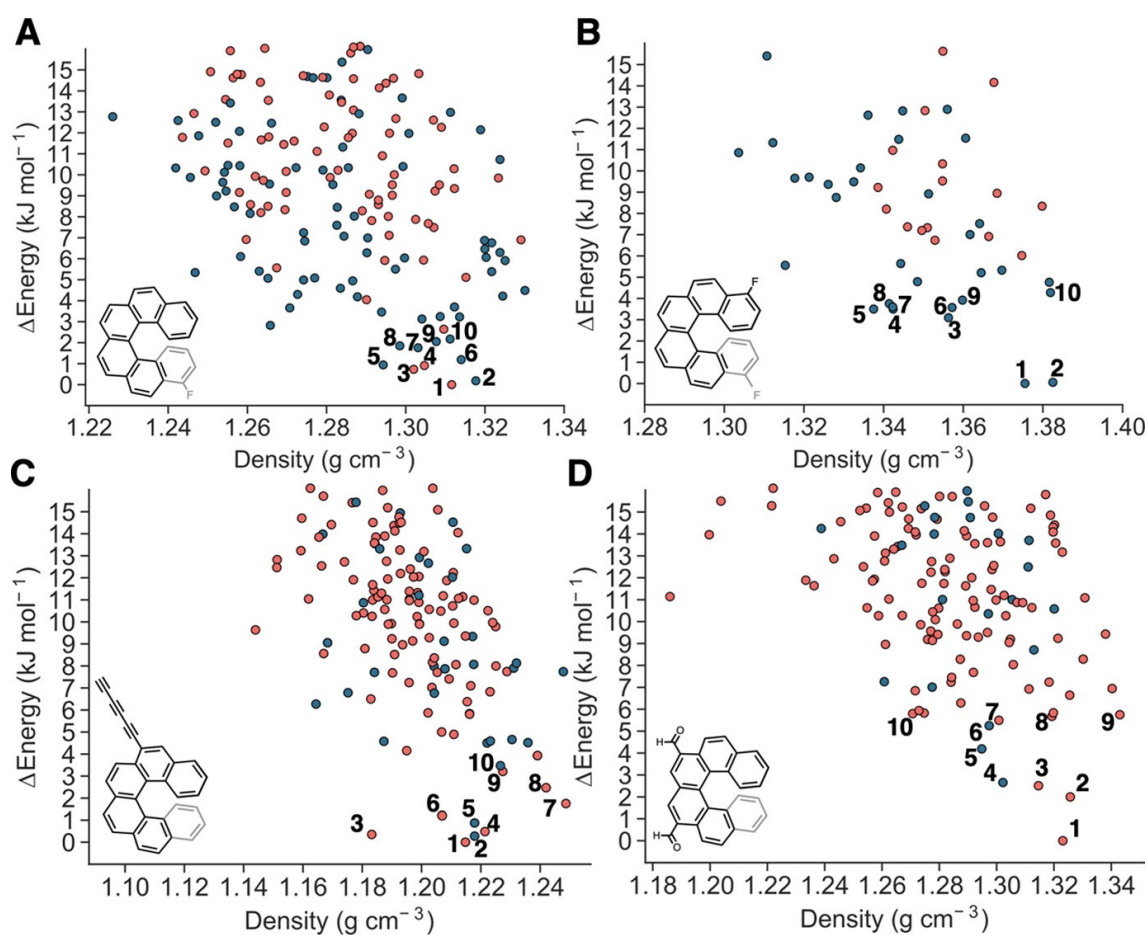


Figure 4.5: Low-energy crystal packing for 13-fluoro[6]helicene (A), 4,13-difluoro[6]helicene (B), 6-triacetylene[6]helicene (C), 7,10-dialdehyde[6]helicene (D). Structures highlighted in blue contain the translational dimer motif while those in red do not. Numbers correspond to candidate crystal structures in Fig. 4.4. Oxygen atoms are shown in red and fluorine in light blue.

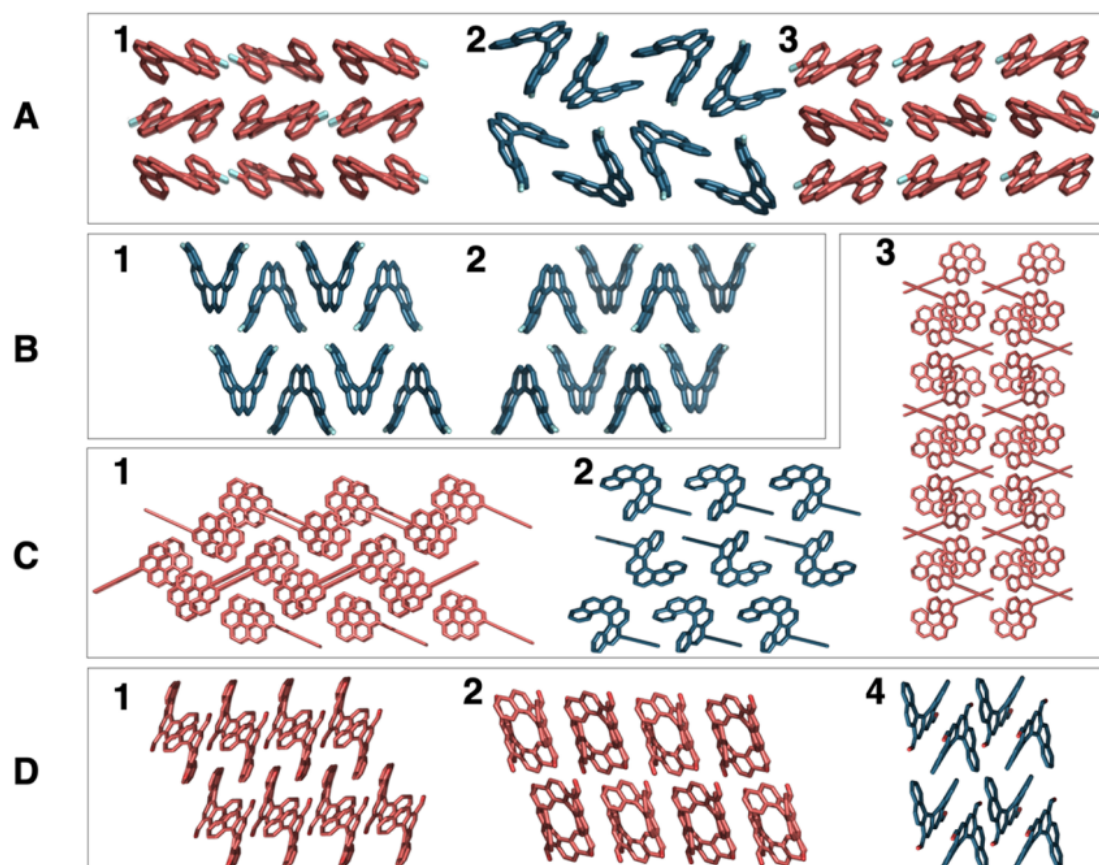
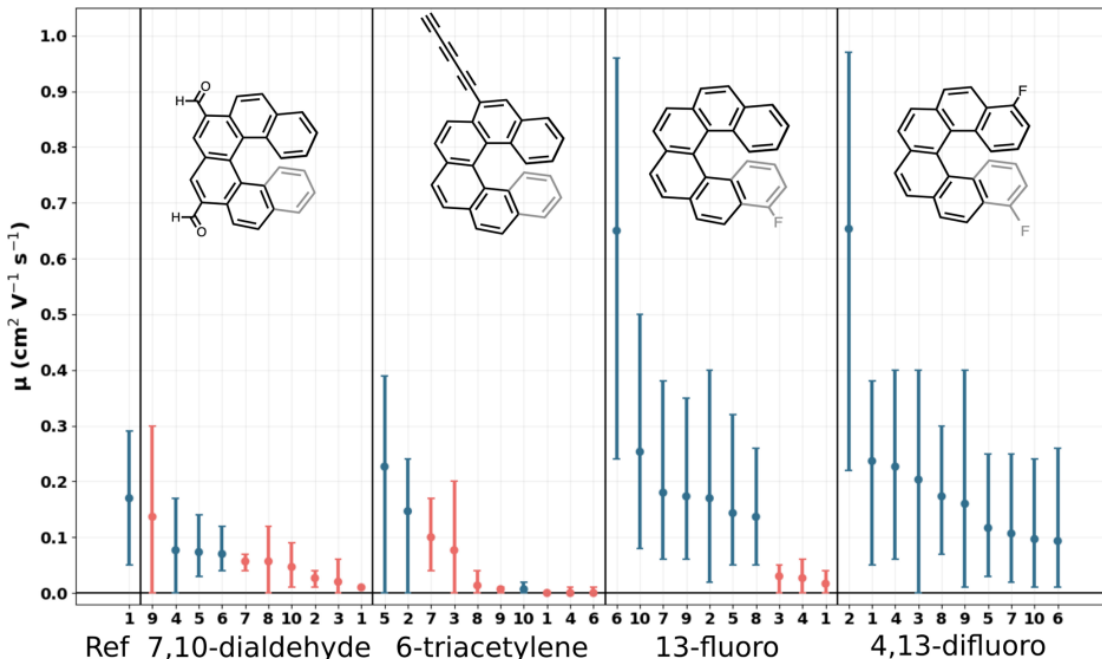


Figure 4.6: Minimum, maximum, and average electron mobilities, μ , computed for each of the 10 lowest-energy polymorphs identified from the CSP searches. Data for structures with (without) the translational motif are shown in blue (red). The result for the reference [6]helicene crystal is shown on the far left.



4.3.1 4,13-difluoro[6]helicene

For 4,13-difluoro[6]helicene, the CSP results found that two polymorphs were significantly lower in energy than the remaining candidates (crystals **1** and **2**, as shown in Fig. 4.5B) and both contain the targeted translational motif. In fact, all 10 lowest-energy crystals contain the translational motif. 4,13-difluoro[6]helicene is a promising OSC candidate as both low-energy polymorphs give electron mobilities which are higher than the [6]helicene reference. Form **2** yields a significantly higher mobility than form **1**. The two structures are nearly degenerate at the presented level of theory. Thus, form **2** could potentially become the most stable form if the crystal structures underwent full geometry optimizations. Furthermore, performing phonon calculations, in addition to using a higher level of theory such as plane-wave DFT, is desirable to determine the likely isolable polymorph.

4.3.2 13-fluoro[6]helicene

In the case of 13-fluoro[6]helicene, we observe more than 10 candidates within 3 kJ/mol of the minimum-energy structure. Furthermore, the majority of crystals contain the translational dimer motif. Mobility calculations reveal that those crystals containing this

favourable dimer arrangement do indeed exhibit desirable OSC properties. The crystal with the highest mobility is form **6**, displaying a 231% increase in charge mobility relative to the reference [6]helicene structure. While form **6** is not the minimum-energy structure at this presented level of theory, it is within 2 kJ/mol of the minimum, suggesting it may be isolable experimentally. The 13-fluoro[6]helicene examined here by our computational screening approach and first-principles CSP highlights the effectiveness of the dimer model for predicting desirable OSC candidates; however, the presented level of theory could be improved upon and reordering of the low-energy structures is possible.

4.3.3 6-triacetylene[6]helicene

Introducing a triacetylene functional group adds a large perturbation to the [6]helicene backbone. As a result, the translation dimer motif is present at a much lower frequency from the CSP results, occurring in only 22% of the crystal structures on which DFT calculations were performed. Furthermore, 7 out of the 10 lowest-energy candidates do not contain the translational dimer motif (Fig. 4.4). Forms **5** and **2** contain the translational dimer and exhibit the highest electron mobilities across the lowest-energy polymorphs. Furthermore, only form **5** yields a maximum mobility higher than that of the reference [6]helicene structure by $0.39 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. This is significantly lower than the mobility results for the mono- and difluoro[6]helicenes. Given the poor mobility results, the fact that CSP suggests multiple plausible low-energy polymorphs that would need to be considered, and other reasons beyond the scope of this thesis⁹⁵ it was determined that it is not a worthwhile effort to study the 6-triacetylene[6]helicene compound further.

4.3.4 7,10-dialdehyde[6]helicene

For the case of the flexible 7,10-dialdehyde[6]helicene, we observe a back-to-back dimer arrangement where the helicene backbones overlap to generate a network-type architecture, as shown in Fig. 4.5D, form **1**. In terms of the crystal-energy landscape, we observe an approximately 2 kJ/mol energy difference between the two lowest-energy structures. Furthermore, the translational dimer motif is observed in only 18% of the total studied crystal structures and in 30% of the 10 lowest-energy candidates. Mobility calculations reveal that the 7,10-dialdehyde[6]helicene does not lead to improved results relative to the reference [6]helicene crystal and performs even worse than the 6-triacetylene[6]helicene crystal.

4.4 Conclusion

In this work, functionalized [6]helicene structures were assessed for their potential as OSC materials. The initial screening approach was based on the assumption that performing computations on a common crystal arrangement of [6]helicenes, the translational dimer motif, is sufficient to obtain a reliable assessment of the OSC properties. Substituting [6]helicene structures with fluorine was found to be the most promising approach to increase electron mobility. CSP was then conducted on several promising candidates, with a combination of a low-level force-field method for geometry optimization and higher-accuracy DFT for single-point energies, to compute the crystal-energy landscapes and energetically rank putative crystal structures.

For the 10 lowest-energy polymorphs appearing in the CSP landscapes of the four most promising systems, the translational motif was found in 100% (4,13-difluoro), 70% (13-fluoro), 30% (6-triacetylene), and 30% (7,10-dialdehyde) of the structures. This is encouraging for the use of the dimer-based model, in that an experimental polymorph screening study could lead to the targeted motif. Furthermore, for each of the two fluorine-substituted [6]helicenes, one low-energy crystal structure predicted with DFT was found to exhibit electron mobilities more than three times the reference value for the unsubstituted [6]helicene. While the crystal-energy landscapes could be improved upon by costly geometry relaxations and phonon calculations, this work presents an advancement in screening of [6]helicene materials. The symmetric 4,13-difluoro[6]helicene is our most promising identified OSC compound, with both low-energy polymorphs identified from CSP possessing high electron mobilities.

CHAPTER 5

A Density-Functional Benchmark of Thermal Free-Energy Corrections for Polymorphic Molecular Crystals

This chapter was adapted from **Weatherby, J. A.**; Rumson, A. F.; Price, A. J. A.; Otero de la Roza, A.; Johnson, E. R. A Density-Functional Benchmark of Thermal Free-Energy Corrections for Polymorphic Molecular Crystals. *J. Chem. Phys.* **2022**, *156*, 114108.

5.1 Introduction

Polymorphism,^{21,22,100} the ability of a substance to exist in more than one crystalline phase, is of great interest in many domains of chemistry and materials science, particularly in drug development.^{101,102} Because polymorphs exhibit different chemical and physical properties, it is often interesting to know *a priori* whether a polymorph with certain desirable properties will be obtained. The field of molecular crystal structure prediction (CSP)^{98,103} aims to use computational methods to predict the thermodynamically stable polymorph, which is often (but not necessarily) also the experimentally observed structure, beginning from the molecular diagram alone. CSP is particularly useful for the elucidation of crystal structures of new molecules, such as a pharmaceutical compound, or when searching for solid-state structures that exhibit specific properties like charge-carrier mobilities.^{8,104}

There is not one unique way to conduct a CSP study. The challenge is in determining a suitable balance between cost and accuracy to predict the likely isolable polymorph(s) and the crystal-energy landscape. Candidate structures are initially generated by sampling the

conformational and solid-form configuration space (usually with restrictions to the most common space groups) of a molecule of interest.^{105,106} While dispersion-corrected density-functional theory (DFT) has shown promise for the subsequent energy ranking,^{5,107-114} it still remains unfeasible to apply DFT to all candidates due to the vast number of structures generated. As a result, CSP studies tend to take the form of a multi-level refinement approach, where several methods are used sequentially to narrow the list of potential candidate structures.^{5,109,115,116} Classical force fields,⁹⁷ density-functional tight binding,^{28,117} or minimal-basis semi-empirical methods^{79,118} may be employed in the early stages of energy ranking to minimize the number of DFT calculations that need to be performed

While there are notable exceptions,¹¹⁰⁻¹¹² most CSP protocols are zeroth-order CSP,⁹⁸ in which only the electronic energies are considered while other contributions to the free energy are neglected. In real molecular crystals, lattice vibrations, known as phonons, contribute in a small but significant way to the free energy. Using a classical force-field approach, it has been shown that the vibrational contributions to the free-energy difference for most organic polymorph pairs are quite small, rarely exceeding 2 kJ/mol.²⁴ Inclusion of vibrational effects still resulted in the reordering of ca. 10-20% of the studied structure pairs,^{23,24} due to the small energy differences between isolable polymorphs. Therefore, an accurate treatment of these vibrational effects is desirable in order to accurately determine the free-energy landscape of a given compound and to find the thermodynamically stable structure. However, due to the expensive nature of phonon frequency calculations, it is at present very computationally demanding to routinely use DFT methods to calculate vibrational free-energy corrections, especially for the large-unit-cell molecular crystals commonly encountered in CSP studies.

As an alternative, approximate methods and models have been proposed to compute the phonon frequencies and vibrational free-energy contributions at reduced cost. When using approximate methods, it is important to consider the trade-off between accuracy and cost. The vibrational free energies calculated using some approximate methods have been compared to converged DFT free energies for a few small molecular crystals.¹¹⁹⁻¹²¹ However, in CSP studies, we are interested in computing the relative free-energy corrections for a crystal structure pair, which may or may not benefit from error cancellation. In this work we propose a new benchmark set, termed the PV17 dataset, for vibrational

free energies and energy differences of molecular crystals. The purpose of this new data set is twofold: i) to provide a reasonably sized benchmark set of high quality (harmonic) vibrational free energy data, and ii) because the dataset contains crystal pairs, it also allows the assessment of approximate methods regarding the calculation of vibrational free-energy differences, and consequently their suitability for incorporation into a CSP protocol.

To build the PV17 dataset, we use the Nyman polymorph library (NPL2016),²³ which contains a large set of molecular crystal structures primarily consisting of two polymorphs for a given organic molecule. A subset of 17 polymorph pairs from this library are identified involving crystals with small unit cells to ensure that the high-level benchmark calculations are feasible. The reference vibrational free-energy data in the PV17 dataset uses dispersion-corrected DFT to evaluate the phonon frequencies and vibrational free-energy corrections within the harmonic approximation. A fully anharmonic or even quasi-harmonic treatment would be desirable, but much more costly, and would complicate the application of the PV17 set to gauge the quality of approximate vibrational models. The B86bPBE-XDM^{30,31,51} functional is used in conjunction with a plane-wave basis set as our high-level method to calculate the reference vibrational free energies because of its high accuracy and reliability for molecular crystals.^{8,32,104,113,114}

Several approximate methods and models are examined in this work using the data in the new PV17 dataset: (i) distributed multipole analysis (DMA) force fields implemented in the DMACRYS package⁹⁷ using the Williams¹²² and FIT¹²³ potentials; (ii) sHF-3c,^{79,118} which is a minimal-basis Hartree–Fock method with added empirical corrections for dispersion and basis-set incompleteness; and (iii) DFTB3-D3(BJ),²⁸ a dispersion-corrected density-functional tight-binding scheme. In addition, we also examine a recent pairing of DFTB3-D3(BJ) with a corrective model, termed the mode-matching¹²¹ approach, which has shown excellent results for computing various thermodynamic quantities. Mode matching serves as an additive correction to the DFTB3-D3(BJ) phonon density of states, although it is easily applicable to other methods.

Our results confirm^{24,119,121} that computing vibrational corrections at the Γ -point only is insufficient to obtain converged values with both high and low levels of theory. The errors yielded by the approximate methods in the calculation of vibrational free energy differences,

including the mode matching approach, are found to be comparable in magnitude to the free-energy differences themselves. This indicates that there are no grounds for preferring these approximate methods over zeroth-order CSP, and that further efforts are needed to develop accurate and cheap vibrational models for routine application in CSP studies. Overall, the converged vibrational corrections to the free-energy differences have values of 2.3 kJ/mol or less, with an average value of 1.0 kJ/mol, confirming the previous force-field result²⁴ that such corrections need only be applied when two (or more) candidate structures are nearly degenerate.

5.2 Background Theory

5.2.1 Phonons and Free-Energy Corrections

Phonon frequencies (ω) are obtained at an arbitrary wavevector \mathbf{q} within the first Brillouin zone as solutions of the secular equation

$$\det \left| \frac{1}{\sqrt{M_m M_n}} \tilde{C}_{mn}^{\xi\eta}(\mathbf{q}) - \omega^2(\mathbf{q}) \right| = 0, \quad (5.1)$$

where M_m is the mass of atom m , ξ and η are the Cartesian directions, and \tilde{C} is the Fourier transform of the force-constant matrix, which is given by

$$C_{mn}^{\xi\eta} \equiv \frac{\partial^2 E}{\partial u_m^\xi \partial u_n^\eta}, \quad (5.2)$$

where u_m^ξ is the displacement of atom m in Cartesian direction ξ . The force-constant matrix is constructed as the second derivative of the potential-energy surface with respect to two nuclear displacements and can be computed via finite-difference methods⁷⁵ or density-functional perturbation theory (DFPT).⁷²

Within the harmonic approximation, the Gibbs free energy is

$$G = E_{\text{static}} + F_{\text{vib}} + pV, \quad (5.3)$$

where E_{static} is the equilibrium, ground-state DFT energy, F_{vib} is the vibrational contribution to the Helmholtz free energy, and pV is the pressure-volume work, which is negligible at ambient pressure. The accurate calculation of F_{vib} and its difference between crystal pairs is the focus of this work. In the harmonic approximation, the vibrational free energy

per unit cell can be computed from the harmonic phonon frequencies as

$$F_{\text{vib}} = \frac{1}{N_{\mathbf{q}}} \sum_{i=1}^{3n} \sum_{\mathbf{q}}^{N_{\mathbf{q}}} \frac{\hbar\omega_{i,\mathbf{q}}}{2} + k_B T \ln \left[1 - \exp \left(-\frac{\hbar\omega_{i,\mathbf{q}}}{k_B T} \right) \right], \quad (5.4)$$

where k_B is Boltzmann’s constant and T is the temperature. The sums run over the $3n$ phonon branches (n is the number of atoms in the unit cell) and the N_q sampled \mathbf{q} -points in the first Brillouin zone. The first term in Eq. 5.4 is the zero-point energy and the second is the temperature-dependent contribution to the F_{vib} .

In this work, the finite-difference approach is used to calculate the force-constant matrix (Eq. 5.2). The convergence of the F_{vib} with respect to the size of the supercell or, equivalently, the \mathbf{q} -point sampling of the Brillouin zone, was studied by systematically varying the supercell according to the formula

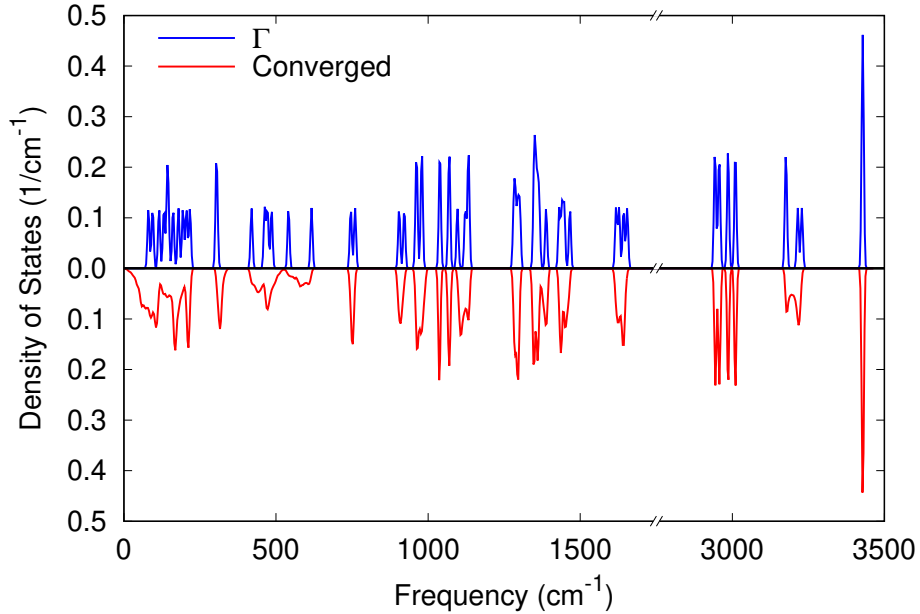
$$n_i = \text{int} \left[\max \left(1, R_k |b_i| + \frac{1}{2} \right) \right], \quad (5.5)$$

where n_i is the size of the supercell in the $i = a, b$, and c directions, b_i is the corresponding reciprocal lattice vector of the primitive unit cell, and R_k is a length parameter. For each individual crystal, the R_k parameter was increased until the F_{vib} was converged to within a threshold of 0.5 kJ/mol per molecule. The phonon frequencies were then Fourier-reinterpolated on a $12 \times 12 \times 12$ Monkhorst-Pack³⁸ mesh, and the final value of F_{vib} was calculated by integration. An example of the difference between the converged phonon density of states (phDOS) and one obtained by reinterpolation using only the Γ -point is shown in Figure 5.1 for the $I\beta$ phase of ethylenediamine. It is clear from this figure that sampling \mathbf{q} points other than Γ is essential to capture the features of the phDOS, particularly at low frequency. It is important to note that the second, temperature-dependent, term in the harmonic free energy (Eq. 5.4) diverges when $\omega \rightarrow 0$. Therefore, the low-frequency region of the phDOS dominates the thermal contribution to the harmonic free energy, for which reason it is essential to model the dispersion of the low-frequency vibrations correctly.

5.2.2 Mode Matching

The mode matching method is a hybrid approximate model recently proposed by Cook and Beran designed to correct the low-level DFTB3-D3(BJ) phonon density of states to yield accurate thermodynamic properties in molecular crystals.¹²¹ The mode matching

Figure 5.1: Overlay of the DFT phDOS for the $I\beta$ phase of ethylenediamine (CCDC code ETDIAM16) computed at Γ and using a converged \mathbf{q} -point grid.

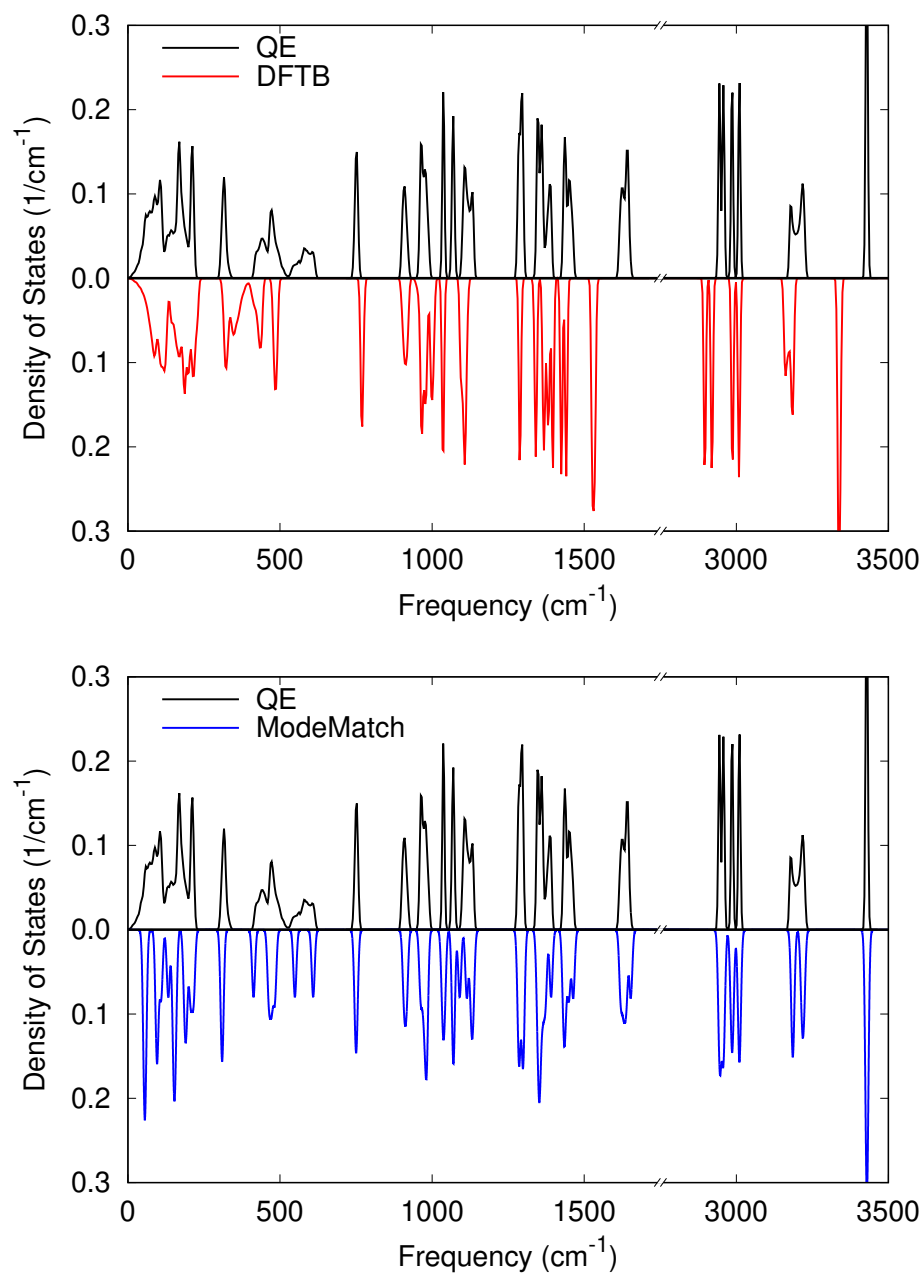


approach calculates the harmonic vibrational frequency for mode i at point \mathbf{q} , $\omega_i(\mathbf{q})$, as

$$\omega_i(\mathbf{q}) \approx \omega_i^{\text{low}}(\mathbf{q}) + \left[\omega_i^{\text{high}}(\Gamma) - \omega_i^{\text{low}}(\Gamma) \right]. \quad (5.6)$$

There are three items required for the evaluation of the mode matching frequencies: i) phonon frequencies at Γ computed with an accurate high-level theoretical method, ii) the frequencies at Γ calculated with a cheaper, less accurate, low-level method, and iii) the converged phDOS from the less accurate method. In Cook and Beran’s work, plane-wave DFT was employed for the high level of theory and DFTB for the low-level method. An example of mode matching is shown in Figure 5.2 for the $I\beta$ phase of ethylenediamine. The advantage of this simple model is that it is no longer necessary to adequately sample the Brillouin zone with DFT, and instead DFTB can be used for this purpose, which leads to a significant reduction in computational cost. Figure 5.2 shows that the mode matching method is successful in reproducing the high-frequency, low-dispersion features of the phDOS. However, the low-frequency acoustic part of the phDOS as well as the region encompassing the intermolecular lattice vibrations, which dominate the thermal contribution to the free energy, are not as well reproduced.

Figure 5.2: Top: Overlay of the phDOS of the $I\beta$ phase of ethylenediamine (ET-DIAM16) computed with QE and DFTB using a converged supercell. Bottom: the same phDOS with QE and the shifted DFTB phDOS calculated using the mode matching method.



5.2.3 Treatment of The Acoustic Modes

Acoustic phonons correspond to low-frequency long-wavelength vibrations of the solid and, hence, are the slowest to converge with respect to \mathbf{q} -point sampling. Due to their low frequencies, the acoustic modes (AM) also have high contributions to the temperature-dependent term of the F_{vib} (Eqn. 5.4). In addition, due to the fact that the acoustic frequencies at Gamma are not zero, Cook and Beran's mode matching method cannot be applied to correct the DFTB acoustic frequencies using the DFT frequencies at Γ .

One way of evaluating the acoustic contribution to the F_{vib} is the Debye model, in which the solid is assumed to behave like a vibrating continuum and the acoustic vibrations are treated as stationary waves spanning the whole crystal. A slightly different version of the Debye model was proposed by Nyman et al.¹¹⁹ and subsequently used by Cook and Beran¹²¹ to treat the vibrational acoustic modes (AM). In this AM model, the phonon dispersion is assumed to be sinusoidal in the wave-vector, as in a one-dimensional atom chain. The acoustic frequencies are given in terms of a Debye-like frequency, ω_{D} , by

$$\omega_{\text{ac}} = \omega_{\text{D}} \sin\left(\frac{\pi}{2} \frac{|\mathbf{q}|}{|\mathbf{q}_{\text{zb}}|}\right), \quad (5.7)$$

where $|\mathbf{q}_{\text{zb}}|$ is the norm of the vector \mathbf{q} at the corresponding Brillouin zone boundary. The Debye-like frequency

$$\omega_{\text{D}} = \frac{2v|\mathbf{q}_{\text{zb}}|}{\pi}, \quad (5.8)$$

is therefore the frequency at the zone boundary for the corresponding direction in reciprocal space (consider $\mathbf{q} = \mathbf{q}_{\text{zb}}$ in Eq. 5.7). In this equation, v is the velocity of sound propagating through the crystal along direction \mathbf{q} obtained by solving the Christoffel equation,

$$\det \left| \Gamma_{ij} - \rho v^2 \delta_{ij} \right| = 0, \quad (5.9)$$

where ρ is the density and δ_{ij} is the Kronecker delta. The Γ_{ij} 's are the 3×3 Christoffel matrices given by

$$\Gamma_{ij} = \sum_{nm} \mathbf{q}_n C_{inmj} \mathbf{q}_m, \quad (5.10)$$

where \mathbf{q}_n are direction vector Cartesian components in reciprocal space.

In their AM model, Nyman et al. chose 13 symmetry-unique directions, corresponding to the simplest Lebedev integration quadrature, and used them to compute an average Debye

frequency, which is ultimately the only parameter in the model.^{119,121} The elastic constant tensor (C_{inmj}) required to calculate the sound velocities along the different propagation directions is computed using the stress-strain relations. DFT calculations for a series of small unit-cell deformations at the equilibrium geometry are used to determine the stress as a function of strain and linear least-squares fits are used to compute the elastic constant tensor.¹²¹

5.3 Computational Methodology

Geometry optimization and phonon frequency calculations were performed for our benchmark set of polymorph pairs in order to calculate the harmonic F_{vib} . In all geometry optimizations, both the atomic positions and lattice vectors were allowed to fully relax, unless otherwise specified. Phonons were always calculated at the same level of theory as the geometry optimization. All free energies were calculated at 300 K. The specific parameters used within each method are given below.

Phonon frequency calculations conducted with the `Quantum ESPRESSO`,³⁹ `CRYSTAL17`,¹²⁴ and `DFTB+`²⁸ packages used the frozen-phonon method as implemented in the `phonopy`⁷⁵ code, v. 2.9.3. The mode matching calculations used the `Modematch` program of Ref. 121. The force-field calculations used the `DMACRYS`⁹⁷ code, v. 2.3.0, together with the `autold` and `autofree` programs of Nyman,²³ which construct a series of linear supercells required to obtain the phonons beyond the Γ -point and subsequently calculate the vibrational free energy.

Plane-wave DFT¹² calculations used periodic boundary conditions and the projector augmented-wave (PAW)⁴⁵ approach implemented in the `Quantum ESPRESSO`³⁹ (QE) software package, v. 6.5. We used the B86bPBE^{30,31} generalized-gradient approximation functional and the exchange-hole dipole moment (XDM) dispersion model.^{51,55} The convergence thresholds were set to 1×10^{-5} and 1×10^{-4} Ry for the energy and forces, respectively. In addition, the convergence in the pressure was set to 1×10^{-2} kbar. Kinetic-energy and charge-density cutoffs of 80 Ry and 800 Ry, respectively, were used along with a Monkhorst-Pack (MP)³⁸ \mathbf{k} -point mesh selected by using an R_k value of 50 bohr in Eq. 5.5.

sHF-3c^{79,118} calculations were performed using the `CRYSTAL17`¹²⁴ code. Full geometry

optimizations (cell and atomic positions) were carried out starting from the B86bPBE-XDM equilibrium geometries. The convergence thresholds on the root-mean-square in the gradient and displacement between subsequent optimization steps were set to 3×10^{-5} and 1.2×10^{-4} a.u., respectively, which are one order of magnitude smaller than the default values. The maximum value for the trust radius was set to 0.25 a.u. and a $4 \times 4 \times 4$ MP \mathbf{k} -point mesh was used. Phonon frequencies were scaled by 0.86, as recommended by Ref. 79.

SCC-DFTB3-D3(BJ)^{28,117} calculations with the 3ob-3-1¹¹⁷ parametrization were performed using the DFTB+²⁸ code, v. 20.2.1. Full geometry optimizations were started from the QE equilibrium geometries. The `MaxForceComponent` parameter, which sets the threshold for convergence in the forces, was set to 1×10^{-5} a.u. Atomic Hubbard derivatives and the parameters used in the D3(BJ) dispersion model were set to those described in Ref. 117.

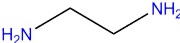
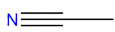
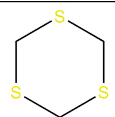
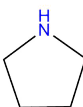
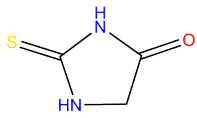
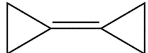
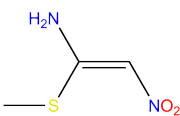
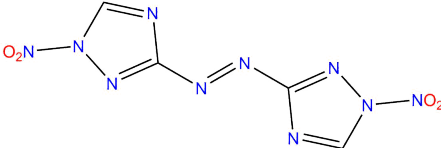
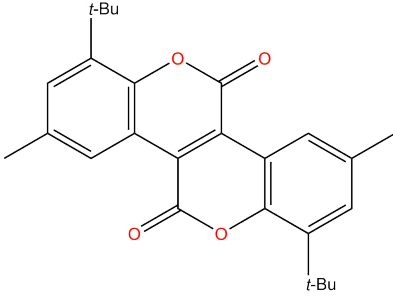
Force-field calculations were carried out under the rigid-body approximation as implemented in the `DMACRYS`⁹⁷ code, which employs distributed atomic multipoles to represent electrostatic contributions. Both the FIT¹²³ and the Williams (W99)¹²² force fields were used, where W99 was supplemented with parameters for sulfur from Ref. 125. Atom-centered multipoles were calculated up to rank 4 (hexadecapole) from a distributed multipole analysis¹²⁶ of the electronic density computed using B3LYP/6-31G** with the `Gaussian09`¹²⁷ program. Rigid-molecule lattice relaxations were then performed, followed by computation of the phonon frequencies¹²⁸ using the `autold` and `autofree` programs of Nyman.¹¹⁹ `autold` generates a series of linear supercells to sample the Brillouin zone beyond the Γ -point;¹¹⁹ here we selected a target \mathbf{q} -point distance of 0.12 \AA^{-1} . `autofree` then collects the frequencies from the supercells and computes the F_{vib} .

5.4 Results

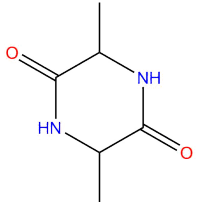
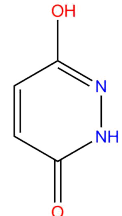
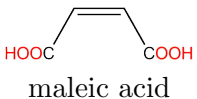
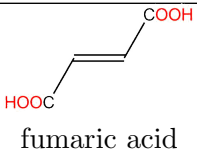
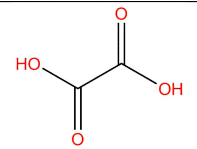
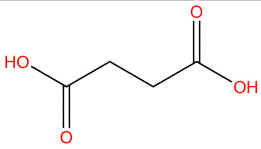
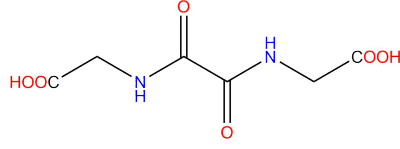
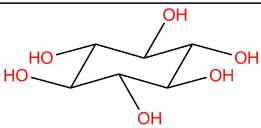
5.4.1 The PV17 Benchmark

The PV17 benchmark was assembled using the Nyman Polymorph Library²⁴ as a starting point. Structures in the library were sorted by cell volume, as this roughly correlates with the numbers of atoms and electrons within the unit cell, and with the computational cost of the eventual phonon frequency calculations. We selected polymorph pairs with both members having volumes less than 600 \AA^3 for further consideration. This choice is necessary to keep the cost of the DFT phonon calculations feasible, although it may introduce a bias in the benchmark set towards small and rigid molecules. We then proceeded to perform

Table 5.1: Molecular structures, Cambridge Crystallographic Data Centre (CCDC)¹³⁰ codes, converged B86bPBE-XDM F_{vib} values at 300 K (in kJ/mol per molecule), and the corresponding supercell sizes, for all crystal structures in the PV17 dataset. ΔF_{vib} values are computed with the lower CCDC number as the reference.

Compound	CSD code	Supercell	F_{vib}	ΔF_{vib}
 ethylenediamine	ETDIAM16 ETDIAM18	2×1×2 2×2×1	274.77 276.21	1.4
 acetonitrile	QQQCIV01 QQQCIV08	3×1×2 3×3×1	105.55 106.45	0.9
 1,3,5-trithiane	TRITAN03 TRITAN10	2×2×1 2×2×2	201.79 201.95	0.2
 pyrrolidine	EFUMAU EFUMAU03	2×2×1 2×3×1	315.50 315.20	-0.3
 2-thiohydantoin	THHYDT THHYDT02	3×2×2 3×2×1	184.47 183.45	-1.0
 bicyclopropylidene	CUMMIG01 CUMMIG02	2×1×3 2×3×2	286.26 285.07	-1.2
 1-nitro-2-methylisothiurea	XOCJEE XOCJEE01	2×1×4 2×2×2	209.12 207.39	-1.7
 1,1'-dinitro-3,3'-azo-1,2,4-triazole	GICTIV GICTIV01	2×1×2 2×2×1	242.64 242.90	0.3
 1,7-di- <i>t</i> -butyl-3,9-dimethyl-dibenzonaphthyrone	DAVVUR DAVVUR01	1×2×2 1×2×2	1163.31 1165.65	2.3

Continued on next page...

Compound	CSD code	Supercell	F_{vib}	ΔF_{vib}
 cyclo-D-alanyl-L-alanyl	TRDMPP01	2×2×2	400.69	
	TRDMPP02	1×2×2	402.83	2.1
 maleic hydrazide	MALEHY10	2×2×2	204.85	
	MALEHY12	3×2×2	205.05	0.2
 maleic acid	MALIAC12	2×1×2	190.94	
	MALIAC13	4×2×1	190.81	-0.1
 fumaric acid	FUMAAC	2×1×2	191.14	
	FUMAAC01	3×2×2	193.02	1.9
 oxalic acid	OXALAC03	2×1×2	113.07	
	OXALAC04	2×1×2	111.49	-1.6
 succinic acid	SUCACB02	2×1×2	251.65	
	SUCACB07	2×2×3	250.32	-1.3
 N,N'-oxalyldiglycine	REKBUE	3×2×1	379.18	
	REKBUE01	3×3×1	378.73	-0.5
 scyllo-inositol	EFURIH	2×2×1	489.43	
	EFURIH04	2×2×2	489.03	-0.4

geometry relaxation and phonon calculations on this subset. Compounds were eliminated from the benchmark if convergence problems were encountered or if the supercell sizes

required for the phonon calculations exceeded our available computational resources. Additionally, compounds were removed from the benchmark if both “polymorphs” converged to the same structure upon relaxation. This can occur if the two reference experimental crystal structures have essentially the same packing, but were determined at different temperatures.¹²⁹

Not all pairs of polymorphs considered have x-ray crystal structures that were obtained at the same temperature. While the two structures forming 12/17 pairs were obtained at the same (room) temperature, those of CUMMIG, EFUMAU, EFURIH, QQQCIV, and THHYDT were obtained at different temperatures. Both MALIAC crystal structures were obtained lower than room temperature (180 K). This poses the issue of whether structures determined at different temperatures are true polymorphs, or if structural differences are simply a result of thermal expansion, which is a difficult distinction to make.¹²⁹ To ensure all structures in our set were in fact distinct polymorphs, we computed the powder-pattern difference (POWDIFF) values using the critic2 program.¹³¹ Specifically, we used the volume-corrected POWDIFF proposed by Mayo and Johnson.¹³² To evaluate the POWDIFF, the x-ray diffraction pattern is computed for each of a pair of given crystal structures and their similarity is assessed using de Gelder’s cross-correlation function.¹³³ The powder-pattern difference is one minus the similarity. Further, in situations where two polymorphs are in fact the same structure, DFT minimization will reveal this as both will converge to the same geometry since the calculations are conducted at effectively zero Kelvin and do not account for temperature-dependent volume changes. Our POWDIFF assessment revealed that two crystal pairs in our initial set (pyrazine and benzoquinone) were actually the same polymorph. This was then confirmed through DFT minimization, which yielded equivalent volumes and electronic energies for the resulting structures.

The resulting benchmark set of 17 polymorph pairs is shown in Table 5.1, along with the converged F_{vib} values and corresponding supercell sizes for each crystal. The final results for the ΔF_{vib} between each polymorph pair are shown in Table 5.2. By convention, we take ΔF_{vib} as the F_{vib} value obtained for the structure with the higher number in its assigned CCDC code minus that for the structure with the lower numbered CCDC code.

In a previous study using a distributed-multipole force field approach,²⁴ Nyman and Day found the harmonic $|\Delta F_{\text{vib}}|$ to be less than 1 kJ/mol for more than 70% of the polymorph

pairs studied in their work, and greater than 2 kJ/mol in fewer than 6% of cases. Table 5.2 shows a summary of our high-level B86bPBE-XDM results, giving the electronic energy difference between each polymorph pair, the converged ΔF_{vib} , and the resulting free-energy difference. Compared to Nyman and Day’s results, the DFT values are similar, although somewhat larger than the force-field results. The average $|\Delta F_{\text{vib}}|$ is 1.0 kJ/mol and values greater than 2 kJ/mol occur for 2/17 (12%) of the compounds considered, although no values exceeded 2.5 kJ/mol for our limited dataset. However, we expect greater differences in the optical region of the phDOS, and therefore greater $|\Delta F_{\text{vib}}|$, if the molecules under study are larger and more flexible than those present in the PV17 set, particularly if polymorphs present different molecular conformations.

Nyman and Day also determined that vibrational free-energy corrections altered the stability ordering for roughly 9% of polymorph pairs when employing the harmonic approximation.²⁴ Our results in Table 5.2 show that the ordering is reversed for 3/17 pairs (18%). We denote reordering of a pair when there is a change in sign from the electronic energy difference (ΔE) to the Gibbs free-energy difference (ΔG). There were a further two pairs that were predicted to be degenerate based on the electronic energy alone, so the F_{vib} entirely determines the stability ranking.

5.4.2 Assessment of Low-Level Methods

With our benchmark values in hand, we proceed to assess the performance of selected low-cost methods for prediction of F_{vib} and ΔF_{vib} for our set of 17 polymorph pairs. The results are collected in Table 5.3 for F_{vib} and Table 5.4 for ΔF_{vib} . Γ indicates results obtained using only the Γ point, while “Conv” indicates results obtained using a converged supercell. Specific to the plane-wave DFT data, we also include the free-energy results calculated using the Γ -point frequencies for the optical contribution and the AM model for the acoustic contribution, as this is much more computationally expedient than fully converging the F_{vib} . These latter results allow us to determine the relative importance of converging the optical and acoustic modes for computation of ΔF_{vib} .

Considering the absolute F_{vib} values in Table 5.3, Γ -point-only calculations give fairly large errors, following the order of HF-3c>DFTB+>DFT. In contrast, using the AM model for the acoustic contribution and calculating the optical contribution to the free energy with the DFT Γ -point frequencies results in quite good performance, with a mean

Table 5.2: Relative B86bPBE-XDM electronic energies (ΔE), converged vibrational free-energy corrections (ΔF_{vib}), and relative free energies (ΔG), for the PV17 benchmark. All values are in units of kJ/mol per molecule. The last row shows the average magnitude of each quantity.

Polymorph Pair	ΔE	ΔF_{vib}	ΔG
ETDIAM	-0.0	1.4	1.4
QQQCIV	-0.6	0.9	0.3
TRITAN	-0.3	0.2	-0.1
EFUMAU	1.0	-0.3	0.7
THHYDT	-1.2	-1.0	-2.2
CUMMIG	4.9	-1.2	3.7
XOCJEE	0.8	-1.7	-0.9
GICTIV	-6.7	0.3	-6.5
DAVVUR	-4.6	2.3	-2.3
TRDMPP	-2.0	2.1	0.1
MALEHY	-0.2	0.2	0.0
MALIAC	0.3	-0.1	0.2
FUMAAC	-7.4	1.9	-5.5
OXALAC	-2.4	-1.6	-4.0
SUCACB	-0.0	-1.3	-1.4
REKBUE	-3.5	-0.5	-4.0
EFURIH	-6.0	-0.4	-6.4
Avg.	2.5	1.0	2.3

absolute error (MAE) of only 2.0 kJ/mol. This error can be reduced even further, to 1.7 kJ/mol, by accounting for \mathbf{q} -point dependence of the optical modes using a converged DFTB calculation in combination with the mode matching approach. This confirms the high accuracy of the mode matching method seen for F_{vib} in Cook and Beran’s study.¹²¹

Table 5.4 compares the performance of the various examined methods in the calculation of vibrational free energy differences for polymorph pairs. In this table, and unlike Table 5.3, we have also included DMACRYS results with the FIT and W99 force fields. The reason for this difference is that, because the DMA force fields use rigid molecules, it is not possible to calculate the intramolecular contribution to F_{vib} . The DMA force field results in Table 5.4 assume implicitly that the intramolecular contribution to the vibrational free energy is the same for both polymorphs.

In contrast to the results for the absolute free energies, we no longer see significant differences in performance between the three QM methods when considering free-energy

Table 5.3: Absolute F_{vib} (kJ/mol per molecule) values for all crystals in our PV17 benchmark set. Results are shown for three electronic structure methods (sHF-3c, DFTB3-D3(BJ), and B86bPBE-XDM) and the mode-matching approach. Free energies were calculated using either the Γ point only or with a converged (Conv) supercell. For B86bPBE-XDM, results are also given for the combination of Γ point frequencies for the optical and the AM model for the acoustic modes (Γ +AM). MAE: Mean absolute error; MAX: Maximum absolute error.

Polymorph	sHF-3c		DFTB3-D3(BJ)		Mode Match	B86bPBE-XDM		
	Γ	Conv	Γ	Conv		Γ	Γ +AM	Conv
ETDIAM16	285.5	281.5	275.8	272.3	273.6	279.3	274.3	274.8
ETDIAM18	287.1	282.4	275.8	272.6	275.1	280.7	275.7	276.2
QQQCIV01	108.6	106.0	107.4	105.1	104.7	107.9	105.0	105.5
QQQCIV08	109.0	106.3	110.7	107.3	105.7	110.3	105.9	106.4
TRITAN03	208.3	205.4	205.5	203.6	200.9	204.3	200.7	201.8
TRITAN10	210.9	204.9	207.2	203.2	201.0	207.3	200.8	202.0
EFUMAU	327.3	325.4	317.3	316.2	314.7	317.3	313.5	315.5
EFUMAU03	327.4	325.3	318.2	316.6	313.2	316.7	313.1	315.2
THHYDT	185.5	182.3	189.0	186.7	184.4	187.2	184.5	184.5
THHYDT02	185.7	179.7	189.7	185.6	181.9	188.4	182.1	183.4
CUMMIG01	297.0	293.4	289.5	287.8	286.5	289.5	283.8	286.3
CUMMIG02	297.6	293.5	289.1	287.5	283.8	288.5	281.7	285.1
XOCJEE	208.8	204.0	214.9	212.1	205.9	212.4	205.7	209.1
XOCJEE01	208.6	203.2	212.7	208.7	206.1	212.6	206.0	207.4
GICTIV	231.8	227.2	260.1	258.4	240.3	243.3	236.5	242.6
GICTIV01	234.6	227.3	267.7	265.1	238.7	247.8	238.7	242.9
DAVVUR	1198.5	1185.1	1182.5	1171.7	1159.0	1175.5	1160.9	1163.3
DAVVUR01	1202.7	1188.7	1187.4	1176.1	1161.8	1179.0	1163.8	1165.7
TRDMPP01	405.8	399.9	404.4	399.8	399.6	406.2	399.8	400.7
TRDMPP02	409.1	404.0	403.9	399.6	400.4	407.0	400.0	402.8
MALEHY10	208.8	205.0	211.1	207.4	203.5	209.5	203.6	204.8
MALEHY12	205.5	203.2	207.3	207.1	204.0	206.1	202.1	205.1
MALIAC12	189.8	187.8	192.4	192.0	190.0	192.6	189.3	190.9
MALIAC13	193.1	187.6	194.2	189.9	189.4	194.8	189.0	190.8
FUMAAC	187.0	185.7	192.3	191.6	188.3	191.3	187.4	191.1
FUMAAC01	197.3	188.9	198.9	193.3	189.9	200.6	190.0	193.0
OXALAC03	109.8	107.6	113.4	111.5	111.8	114.9	111.8	113.1
OXALAC04	111.0	107.6	113.9	111.7	110.3	114.5	109.4	111.5
SUCACB02	255.6	250.9	254.9	253.0	250.8	255.2	250.0	251.6
SUCACB07	253.5	249.0	253.1	250.2	248.1	253.8	247.5	250.3
REKBUE	376.3	370.1	379.9	375.1	377.9	384.6	378.3	379.2
REKBUE01	381.1	371.3	381.4	375.5	376.9	387.8	375.7	378.7
EFURIH	497.8	494.4	479.9	476.9	488.0	492.9	487.3	489.4
EFURIH04	498.0	493.4	479.5	475.1	486.5	493.6	487.1	489.0
MAE	7.3	6.0	5.6	3.8	1.7	4.2	2.0	–
MAX	37.1	23.0	24.8	22.2	4.4	13.3	6.1	–

Table 5.4: Relative F_{vib} (kJ/mol per molecule) values for all polymorph pairs in the PV17 benchmark set. Results are shown for two DMA force fields (DMACRYS), three electronic structure methods (sHF-3c, DFTB3-D3(BJ), and B86bPBE-XDM), and the mode-matching approach. Results are reported using either the Γ point or with a converged (Conv) supercell. Only converged results are shown for the DMA calculations. For B86bPBE-XDM, results are also given using the Γ point with the AM mode correction (Γ +AM). MAE: Mean absolute error; MAX: Maximum absolute error.

Polymorph Pair	DMACRYS		sHF-3c		DFTB3-D3(BJ)		Mode Match	B86bPBE-XDM		
	W99	FIT	Γ	Conv	Γ	Conv		Γ	Γ +AM	Conv
ETDIAM	0.0	0.5	1.6	0.9	0.1	0.2	1.4	1.4	1.4	1.4
QQQCIV	0.3	0.3	0.4	0.3	3.3	2.2	1.0	2.3	0.9	0.9
TRITAN	1.4	-0.1	2.6	-0.5	1.7	-0.5	0.1	3.1	0.1	0.2
EFUMAU	0.6	0.3	0.1	0.0	0.9	0.4	-1.5	-0.6	-0.4	-0.3
THHYDT	0.5	0.7	0.2	-2.6	0.6	-1.1	-2.5	0.8	-2.4	-1.0
CUMMIG	0.3	0.5	0.7	0.1	-0.4	-0.3	-2.7	-1.0	-2.1	-1.2
XOCJEE	-0.4	-0.3	-0.3	-0.8	-2.2	-3.4	0.2	0.2	0.3	-1.7
GICTIV	0.3	0.7	2.8	0.1	7.6	6.7	-1.6	4.5	2.2	0.3
DAVVUR	1.1	1.1	4.2	3.6	4.9	4.4	2.8	3.4	2.9	2.3
TRDMPP	1.4	1.3	3.3	4.1	-0.4	-0.1	0.8	0.8	0.2	2.1
MALEHY	-1.4	-1.6	-3.3	-1.8	-3.8	-0.2	0.5	-3.3	-1.5	0.2
MALIAC	-0.1	-0.2	3.2	-0.2	1.7	-2.0	-0.6	2.3	-0.3	-0.1
FUMAAC	2.0	2.2	10.3	3.2	6.5	1.8	1.6	9.3	2.6	1.9
OXALAC	0.7	-0.2	1.3	0.0	0.5	0.1	-1.5	-0.4	-2.4	-1.6
SUCACB	-0.7	-1.6	-2.0	-1.9	-1.8	-2.8	-2.8	-1.5	-2.5	-1.3
REKBUE	0.8	-0.4	4.9	1.2	1.5	0.4	-1.0	3.3	-2.6	-0.5
EFURIH	0.9	-0.7	0.2	-1.1	-0.4	-1.7	-1.5	0.8	-0.2	-0.4
MAE	1.1	0.8	2.2	1.0	2.1	1.5	0.8	2.0	0.9	-
MAX	2.3	1.8	8.4	2.0	7.3	6.4	1.9	7.4	2.2	-

differences. Γ -point only calculations with sHF-3c, DFTB3-D3(BJ), and B86bPBE-XDM all yield mean absolute errors (MAEs) of nearly 2 kJ/mol and maximum errors (MAX) of 7.3-8.4 kJ/mol. In some cases, most notably for fumaric acid (FUMAAC), the ΔF_{vib} computed using only the Γ point can be much higher than the converged values, emphasizing the importance of properly accounting for the low-frequency phonon dispersion and the acoustic contributions. Thus, using only the Γ -point frequencies to assess the magnitude of the free-energy correction in CSP studies may be quite misleading, and is not recommended.

Comparing the force-field, converged low-level QM, and mode matching results with the converged B86bPBE-XDM reference values, we see that all give fairly equivalent error statistics, with MAEs of 0.8-1.5 kJ/mol and maximum errors of ca. 2-7 kJ/mol. Of these approaches, DFTB3-D3(BJ) gives the largest error for 1,1'-dinitro-3,3'-azo-1,2,4-triazole (GICTIV), indicating that it might not perform well for azo compounds. Indeed, from Table 5.3, GICTIV shows the largest errors in absolute F_{vib} for DFTB3-D3(BJ) as well. Mode matching tends to give lower errors than the converged DFTB results, particularly for acetonitrile (QQQCIV), 1,1'-dinitro-3,3'-azo-1,2,4-triazole (GICTIV), 1,7-di-*t*-butyl-3,9-dimethyl-dibenzonaphthyrone (DAVVUR), and most of the systems with dimeric H-bonds involving either COOH or CONH groups. However, it is notable that the DMA force fields give comparable overall errors to the semi-empirical QM methods, and mode matching, with a much lower computational cost.

Finally, we consider the performance of using the B86bPBE-XDM optical contribution estimated from the Γ point frequencies coupled with AM model for the acoustic contribution. The rationale beyond this combination is that the low-frequency, long-wavelength acoustic modes are the slowest to converge with respect to \mathbf{q} -points and are missing from the Γ -point treatment entirely. We can therefore add the acoustic contribution to F_{vib} from the approximate AM model to the Γ -point results as a correction and, because the optical modes typically have a weaker dependence on \mathbf{q} -points, this may be a good approximation to the converged result. The AM approximation has an additional set of calculations on top of the Γ -point frequency calculations, since it requires evaluation of the elastic constants.

Inclusion of the AM model correction more than halves the MAE relative to the Γ -point-only results with B86bPBE-XDM. Notably the Γ +AM approach gives significantly better agreement with the converged reference values for many of the polymorph pairs, such

as acetonitrile (QQQCIV), 1,3,5-trithiane (TRITAN), 1,1'-dinitro-3,3'-azo-1,2,4-triazole (GICTIV), maleic hydrazide (MALEHY), maleic acid (MALIAC), N,N'-oxalyldiglycine (REKBUE), and most significantly fumaric acid (FUMAAC), which consistently has the largest error in the Γ -point calculations. Once more, this result highlights the importance of properly accounting for the acoustic contributions to ΔF_{vib} .

The cases where Γ +AM still gives errors in excess of 1 kJ/mol are 2-thiohydantoin (THHYDT), 1,1'-dinitro-3,3'-azo-1,2,4-triazole (GICTIV), maleic hydrazide (MALEHY), cyclo-D-alanyl-L-alanyl (TRDMPP), succinic acid (SUCACB), and N,N'-oxalyldiglycine (REKBUE). With one exception, all of these compounds form dimeric H-bonds involving two COOH or CONH groups. This suggests that optical modes involving strong, cooperative H-bonding networks require more extensive \mathbf{q} -point sampling and that changes in these H-bonding motifs between polymorphs have significant contributions to ΔF_{vib} . The remaining errors in the Γ +Debye results may be due to either neglect of the dependence of the optical modes on \mathbf{q} -points, as seen for the dimeric H-bonds, or to breakdowns of the AM model approximation.

5.5 Discussion and Conclusions

This work presented the new PV17 benchmark, containing DFT absolute (F_{vib}) and relative (ΔF_{vib}) harmonic vibrational free energy data for pairs of crystalline organic polymorphs. This benchmark was used to assess the performance of several force-field and semi-empirical QM methods for prediction of relative free-energy corrections in molecular crystal polymorph pairs. Both planewave DFT frequency calculations at the Γ point, augmented with treatment of the acoustic modes using a Debye-like model, and the recent mode-matching approach¹²¹ showed good performance for prediction of the absolute F_{vib} for the individual molecular crystals.

However, none of the methods studied were able to benefit significantly from error cancellation in the evaluation of ΔF_{vib} for the polymorph pairs. Overall, the MAEs obtained with all examined methods were comparable to the average magnitude of ΔF_{vib} itself. As an illustration, a method that consistently gives a ΔF_{vib} of zero, which is equivalent to a zeroth-order CSP protocol, would yield a MAE of 1.0 kJ/mol and MAX error of 2.3 kJ/mol, which are on par with, or better than, the results obtained with all of the methods considered in this study. Thus, although free energies calculated at a low-level

of theory may be useful in other contexts, our recommendation at present is to neglect thermal free-energy effects on CSP ranking entirely, rather than to calculate them with a low level of theory and introduce an additional uncontrolled error. However, some of the approaches examined, like Cook and Beran’s mode matching approach and also combining a DFT Γ -point frequency calculation with a model for the acoustic contribution do show promise, but more work is required until practical CSP protocols can benefit from these methods.

While the magnitudes of our predicted thermal corrections were slightly larger than those obtained previously, we confirm the finding of Nyman and Day²⁴ that ΔF_{vib} is generally small in magnitude. This means that thermal effects need only be taken into account at the conclusion of a CSP protocol if two or more polymorphs are nearly degenerate, with a relative energy difference of less than ca. 2.5 kJ/mol.

Calculating ΔF_{vib} for polymorph pairs is very challenging because it is a small difference between large numbers. The individual F_{vib} values have typical magnitudes of ca. 100-500 kJ/mol per molecule for the molecular crystals considered here, while the differences between polymorphs of the same compound have magnitudes under 2.5 kJ/mol. Thus, it is reasonable that converging the total F_{vib} to within 0.5 kJ/mol, which requires relatively dense \mathbf{q} -point grids in some cases, may not be entirely satisfactory for the purpose of CSP candidate ranking based on the free energy. In addition, we have not considered quasi-harmonic or anharmonic effects, whose inclusion would complicate matters even further. While the present work represents an advance over the previous assessment of ΔF_{vib} using DMA potentials, even more precise benchmarks for ΔF_{vib} are highly desirable, as well as the development of more accurate vibrational models for organic molecular crystals.

CHAPTER 6

Conclusions and Future Work

6.1 Summary of Chapters

The work presented in this thesis applied high-accuracy plane-wave density-functional theory (DFT) to allotropes and polymorphs to determine relative energetic stabilities. Furthermore, we included temperature-dependant vibrational effects to compute the vibrational free-energy corrections, F_{vib} , in several chapters. The work presented here has applications in molecular crystal structure prediction (CSP), and specifically polymorphism. Our work began with investigating the long-standing question of which allotrope of carbon, diamond or graphite, is more thermodynamically stable at room temperature. Next, with the help of first-principles CSP, we explored functionalized [6]helicene systems for their suitability as organic semi-conducting materials. Finally, we proposed a new benchmark of thermal free-energy corrections for a set of 17 organic molecular polymorph pairs.

Our first study in Ch. 3 was inspired by previous computational work, which proposed that diamond and graphite become degenerate as the temperature approaches 0 K.⁸¹ We used plane-wave DFT with the B86bPBE-25X//B86bPBE-XDM and PBE0-XDM//PBE-XDM density functionals to compute the electronic energies of both allotropes. Accounting for temperature effects by performing phonon calculations allowed for the computation of the vibrational contribution to the free energy, F_{vib} . The resulting ΔG values at 298 K relative to diamond were predicted to be -3.65 and -2.96 kJ/mol per C atom with the B86bPBE- and PBE-based models listed above. The experimental value obtained from high-accuracy thermochemical measurements was determined to be -3.17 kJ/mol per C atom.⁸² Not only are the DFT models presented here highly accurate in reproducing the experimental value, we see that graphite is more thermodynamically stable than diamond.

In a collaborative project, functionalized [6]helicene structures were screened for their suitability as organic semi-conducting materials.⁹⁵ The functional group landscape was surveyed using a translational dimer as a simple model of a frequently recurring packing motif in helicene crystal structures. The dimer model was used to compute several parameters to assess OSC suitability. Following this screening procedure, a full CSP study was undertaken for four promising structures to determine the likely isolable polymorph(s). The SIESTA program was used to compute the crystal-energy landscapes and several low-energy structures were identified, as described in Ch. 4. Furthermore, the translational dimer model was validated as many of the low-energy crystal structures contained this packing motif. Fluorine substitution was found to significantly increase the charge-carrier mobility of the system relative to the unsubstituted [6]helicene crystal. We propose that the symmetrically substituted 4,13-difluoro[6]helicene structure is the most promising material for future synthesis, with both low-energy polymorphs possessing high electron mobilities.

The final project we undertook for this thesis consisted of generating accurate plane-wave DFT benchmark data for relative vibrational contributions to the free energy (F_{vib}), specifically for polymorphic pairs of organic molecular crystals. In Ch. 5, we assembled the PV17 benchmark set, which was derived from the Nyman Polymorph Library – a large library of polymorphic molecular crystals.²³ We found that the magnitude of relative F_{vib} values were $\sim 0 - 2$ kJ/mol per molecule, in agreement with previous force field studies.²⁴ Thus, applying thermal corrections will only lead to reordering when polymorphs are nearly degenerate. Given our benchmark set of small organic molecules, with minimal structural differences between polymorphs, this result is unsurprising. There would need to be large differences in the phDOS, specifically in the vibrations corresponding to the low-frequency and acoustical modes, to cause large enough ΔF_{vib} values to result in energetic reordering. It is expected that larger ΔF_{vib} values will result when there are drastic bonding or conformational differences. We then assessed the accuracy of several low-cost computational models, namely sHF-3c,¹¹⁸ DFTB3-D3(BJ),^{28,117} distributed multipole force fields within the DMACRYS⁹⁷ framework, and lastly the mode-matching approach of Cook and Beran.¹²¹ Disappointingly, none of these methods were satisfactory in reproducing the benchmark data. Furthermore, the MAEs of most models were as large as the relative thermal corrections themselves, meaning that neglecting the thermal free-energy corrections entirely may be a better approximation than using low-level results in CSP.

6.2 Future Work

There are several clear advancements that have been accomplished in this work, specifically with regards to computing accurate relative (free) energies of molecular crystals and allotropes of carbon. We now propose some ideas for extending the present work.

In Ch. 4, one main drawback of the computational methodology used is the omission of full DFT geometry relaxations (to obtain the minimum-energy crystal structures) and phonon calculations to compute thermal effects. While the SIESTA method could be used for the geometry optimizations, and has been shown to be accurate in predicting isolable polymorphs for several organic molecular crystals,⁵ full plane-wave DFT using the QUANTUM ESPRESSO package would be highly desirable. However, plane-wave DFT is computationally expensive due to the delocalized basis set. Use of numerical atom-centered orbitals with the FHIAIMS software package,¹³⁴ which has been shown to provide an accuracy comparable to that of plane-wave codes, could be a logical next step for geometry optimizations and phonon calculations.

We saw in Ch. 5 that none of the low-cost methods considered were capable of reproducing the benchmark data. Therefore, identifying or developing new, accurate, low-cost models is highly desirable. The force fields employed in this work use the DMACRYS code, which operates under the rigid-body approximation. As a result, intermolecular interactions (and vibrational modes) are well modeled, but we are unable to describe the intramolecular changes in geometry and contributions from those vibrational modes. Intramolecular vibrations, such as OH and CO stretches, will be impacted by hydrogen bonding, but these contributions are not captured by the DMA potentials. Tailor-made flexible-molecule force fields could be generated for each compound; however, high-quality reference data would be required, in addition to the force field being non-transferable. Next, with regards to the treatment of the acoustic modes, we assumed that the lowest three phonon frequencies at each \mathbf{q} -point were the acoustic modes, which is not necessarily the case. For a more accurate treatment, we would need to track the acoustic phonon dispersion as a function of \mathbf{q} -point. Lastly, the reference data was computed using the plane-wave formalism in QUANTUM ESPRESSO, causing the phonon calculations to be very expensive in both time and memory. Future work could focus on using the FHIAIMS code, as calculations using numerical atom-centered orbitals are significantly faster and require less memory. Therefore, phonons could be computed more efficiently for molecular crystals with large

unit-cell sizes and many atoms. The first step would be to apply the new implementation of the B86bPBE-XDM functional in FHIAIMS to the PV17 benchmark and assess the quality of the results and the cost of the calculations.

Bibliography

- [1] MacMillan, J. W.; Marczenko, K. M.; Johnson, E. R.; Chitnis, S. S. Hydrostibination of alkynes: A radical mechanism. *Chem. Eur. J.* **2020**, *26*, 17134–17142.
- [2] Lazarou, Y. G.; Prosmiris, A. V.; Papadimitriou, V. C.; Papagiannakopoulos, P. Theoretical calculation of bond dissociation energies and enthalpies of formation for halogenated molecules. *J. Phys. Chem. A* **2001**, *105*, 6729–6742.
- [3] Price, A. J.; Johnson, E. R. Theoretical investigation of amino-acid adsorption on hydroxylated quartz surfaces: dispersion can determine enantioselectivity. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16571–16578.
- [4] Otero-de-la Roza, A.; Abbasi-Pérez, D.; Luaña, V. Gibbs2: A new version of the quasiharmonic model code. II. Models for solid-state thermodynamics, features and implementation. *Comput. Phys. Commun.* **2011**, *182*, 2232–2248.
- [5] LeBlanc, L. M.; Johnson, E. R. Crystal-energy landscapes of active pharmaceutical ingredients using composite approaches. *CrystEngComm* **2019**, *21*, 5995–6009.
- [6] Isla, H.; Crassous, J. Helicene-based chiroptical switches. *C. R. Chimie* **2016**, *19*, 39–49.
- [7] Yang, Y.; da Costa, R. C.; Fuchter, M. J.; Campbell, A. J. Circularly polarized light detection by a chiral organic semiconductor transistor. *Nat. Photonics* **2013**, *7*, 634–638.
- [8] Yang, Y. et al. Emergent properties of an organic semiconductor driven by its molecular chirality. *ACS Nano* **2017**, *11*, 8329–8338.
- [9] Jhulki, S.; Mishra, A. K.; Chow, T. J.; Moorthy, J. N. Helicenes as All-in-One Organic Materials for Application in OLEDs: Synthesis and Diverse Applications of Carbo- and Aza[5]helical Diamines. *Chem. Eur. J* **2016**, *22*, 9375–9386.
- [10] Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864.
- [11] Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.
- [12] Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **2014**, *140*, 18A301.

- [13] Mazurek, A. H.; Szeleszczuk, L.; Pisklak, D. M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415.
- [14] Makkar, P.; Ghosh, N. N. A review on the use of DFT for the prediction of the properties of nanomaterials. *RSC Adv.* **2021**, *11*, 27897–27924.
- [15] <https://en.wikipedia.org/wiki/Phonon>.
- [16] Erba, A.; Maul, J.; Civalleri, B. Thermal properties of molecular crystals through dispersion-corrected quasi-harmonic ab initio calculations: the case of urea. *ChemComm* **2016**, *52*, 1820–1823.
- [17] Heit, Y. N.; Nanda, K. D.; Beran, G. J. Predicting finite-temperature properties of crystalline carbon dioxide from first principles with quantitative accuracy. *Chem. Sci.* **2016**, *7*, 246–255.
- [18] Salim, M. A.; Willow, S. Y.; Hirata, S. Ice Ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization. *J. Chem. Phys.* **2016**, *144*, 204503.
- [19] Ruggiero, M. T.; Zeitler, J. A.; Erba, A. Intermolecular anharmonicity in molecular crystals: interplay between experimental low-frequency dynamics and quantum quasi-harmonic simulations of solid purine. *ChemComm* **2017**, *53*, 3781–3784.
- [20] McKinley, J. L.; Beran, G. J. Improving Predicted Nuclear Magnetic Resonance Chemical Shifts Using the Quasi-Harmonic Approximation. *J. Chem. Theory Comput.* **2019**, *15*, 5259–5274.
- [21] Bernstein, J. *Polymorphism in Molecular Crystals*; Oxford University Press, 2020.
- [22] Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **2015**, *44*, 8619–8635.
- [23] Nyman, J.; Day, G. M. Modelling temperature-dependent properties of polymorphic organic molecular crystals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31132–31143.
- [24] Nyman, J.; Day, G. M. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm* **2015**, *17*, 5154–5165.
- [25] Ashcroft, N.; Mermin, N. Solid state physics; thomson learning, inc. *Stamford, CT* **1976**,
- [26] Červinka, C.; Beran, G. J. Ab initio prediction of the polymorph phase diagram for crystalline methanol. *Chem. Sci.* **2018**, *9*, 4622–4629.

- [27] Brandenburg, J. G.; Grimme, S. Dispersion corrected Hartree–Fock and density functional theory for organic crystal structure prediction. *Prediction and Calculation of Crystal Structures* **2013**, 1–23.
- [28] Hourahine, B. et al. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101.
- [29] García, A. et al. Siesta: Recent developments and applications. *J. Chem. Phys.* **2020**, *152*, 204108.
- [30] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- [31] Becke, A. On the large-gradient behavior of the density functional exchange energy. *J. Chem. Phys.* **1986**, *85*, 7184–7187.
- [32] Otero-de-la Roza, A.; Johnson, E. R. A benchmark for non-covalent interactions in solids. *J. Chem. Phys.* **2012**, *137*, 054103.
- [33] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- [34] Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- [35] Otero-de-la Roza, A.; LeBlanc, L. M.; Johnson, E. R. Dispersion XDM with hybrid functionals: Delocalization error and halogen bonding in molecular crystals. *J. Chem. Theory Comput.* **2019**, *15*, 4933–4944.
- [36] Lin, L. Adaptively compressed exchange operator. *J. Chem. Theory Comput.* **2016**, *12*, 2242–2249.
- [37] Bloch, F. Quantum mechanics of electrons in crystal lattices. *Z. Phys.* **1928**, *52*, 555–600.
- [38] Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188.
- [39] Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Condens. Matter Phys.* **2009**, *21*, 395502.
- [40] Giannozzi, P. et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Condens. Matter Phys.* **2017**, *29*, 465901.
- [41] Gonze, X. et al. The Abinit project: Impact, environment and recent developments. *Comput. Phys. Commun.* **2020**, *248*, 107042.

- [42] Romero, A. H. et al. ABINIT: Overview, and focus on selected capabilities. *J. Chem. Phys.* **2020**, *152*, 124102.
- [43] Schwerdtfeger, P. The pseudopotential approximation in electronic structure theory. *Chemphyschem* **2011**, *12*, 3143–3155.
- [44] Troullier, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **1991**, *43*, 1993.
- [45] Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **1994**, *50*, 17953.
- [46] Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **1990**, *41*, 7892.
- [47] Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758.
- [48] Hamann, D.; Schlüter, M.; Chiang, C. Norm-conserving pseudopotentials. *Phys. Rev. Lett.* **1979**, *43*, 1494.
- [49] Troullier, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculations. II. Operators for fast iterative diagonalization. *Phys. Rev. B* **1991**, *43*, 8861.
- [50] Andersen, O. K. Linear methods in band theory. *Phys. Rev. B* **1975**, *12*, 3060.
- [51] Otero-de-la Roza, A.; Johnson, E. R. Van der Waals interactions in solids using the exchange-hole dipole moment model. *J. Chem. Phys.* **2012**, *136*, 174109.
- [52] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105–5154.
- [53] Berland, K.; Cooper, V. R.; Lee, K.; Schröder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B. I. van der Waals forces in density functional theory: a review of the vdW-DF method. *Rep. Prog. Phys.* **2015**, *78*, 066501.
- [54] Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction revisited. *J. Chem. Phys.* **2007**, *127*, 154108.
- [55] Johnson, E. R. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; Otero-de-la-Roza, A., DiLabio, G., Eds.; Elsevier: Amsterdam, Netherlands, 2017; Chapter 5, pp 169–194.
- [56] Otero-de-la Roza, A.; Johnson, E. R. Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

- [57] Christian, M. S.; Otero-de-la Roza, A.; Johnson, E. R. Surface adsorption from the exchange-hole dipole moment dispersion model. *J. Chem. Theory Comput.* **2016**, *12*, 3305–3315.
- [58] Otero-de-la Roza, A.; LeBlanc, L. M.; Johnson, E. R. Asymptotic Pairwise Dispersion Corrections Can Describe Layered Materials Accurately. *J. Phys. Chem. Lett.* **2020**, *11*, 2298–2302.
- [59] Soler, J. M.; Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. The SIESTA method for ab initio order-N materials simulation. *J. Condens. Matter Phys.* **2002**, *14*, 2745.
- [60] Artacho, E. et al. The SIESTA method; developments and applicability. *J. Condens. Matter Phys.* **2008**, *20*, 064208.
- [61] LeBlanc, L. M.; Weatherby, J. A.; Otero-de-la Roza, A.; Johnson, E. R. Non-covalent interactions in molecular crystals: Exploring the accuracy of the exchange-hole dipole moment model with local orbitals. *J. Chem. Theory Comput.* **2018**, *14*, 5715–5724.
- [62] Frisch, M. J. et al. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- [63] The postg program can be obtained from <http://schooner.chem.dal.ca>.
- [64] Axilrod, B.; Teller, E. Interaction of the van der Waals type between three atoms. *J. Chem. Phys.* **1943**, *11*, 299–300.
- [65] Muto, Y. Force between nonpolar molecules. *J. Phys. Math. Soc. Jpn.* **1943**, *17*, 629–631.
- [66] Otero-de-la Roza, A.; LeBlanc, L. M.; Johnson, E. R. What is “many-body” dispersion and should I worry about it? *Phys. Chem. Chem. Phys.* **2020**, *22*, 8266–8276.
- [67] Salem, L. The calculation of dispersion forces. *Mol. Phys.* **1960**, *3*, 441–452.
- [68] Becke, A. D.; Roussel, M. R. Exchange holes in inhomogeneous systems: A coordinate-space model. *Phys. Rev. A* **1989**, *39*, 3761.
- [69] Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, *44*, 129–138.
- [70] Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-theoretic approaches to atoms-in-molecules: Hirshfeld family of partitioning schemes. *J. Phys. Chem. A* **2017**, *122*, 4219–4245.

- [71] Kannemann, F. O.; Becke, A. D. van der Waals interactions in density-functional theory: intermolecular complexes. *J. Chem. Theory Comput.* **2010**, *6*, 1081–1088.
- [72] Baroni, S.; De Gironcoli, S.; Dal Corso, A.; Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *RMP* **2001**, *73*, 515.
- [73] Hellman, H. Einführung in die Quantenchemie. *Franz Deuticke, Leipzig* **1937**, 285.
- [74] Feynman, R. P. Forces in molecules. *Phys. Rev.* **1939**, *56*, 340.
- [75] Togo, A.; Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **2015**, *108*, 1–5.
- [76] Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B* **1995**, *51*, 12947.
- [77] Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. Automated parametrization of SCC-DFTB repulsive potentials: Application to hydrocarbons. *J. Phys. Chem. A* **2009**, *113*, 11866–11881.
- [78] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. 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.
- [79] Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- [80] Brandenburg, J. G.; Alessio, M.; Civalleri, B.; Peintinger, M. F.; Bredow, T.; Grimme, S. Geometrical correction for the inter-and intramolecular basis set superposition error in periodic density functional theory calculations. *J. Phys. Chem. A* **2013**, *117*, 9282–9292.
- [81] Grochala, W. Diamond: electronic ground state of carbon at temperatures approaching 0 K. *Angew. Chem. Int. Ed.* **2014**, *53*, 3680–3683.
- [82] White, M. A.; Kahwaji, S.; Freitas, V. L.; Siewert, R.; Weatherby, J. A.; Ribeiro da Silva, M. D.; Verevkin, S. P.; Johnson, E. R.; Zwanziger, J. W. The Relative Thermodynamic Stability of Diamond and Graphite. *Angew. Chem. Int. Ed.* **2021**, *60*, 1546–1549.
- [83] Gražulis, S.; Chateigner, D.; Downs, R. T.; Yokochi, A.; Quirós, M.; Lutterotti, L.; Manakova, E.; Butkus, J.; Moeck, P.; Le Bail, A. Crystallography Open Database—an open-access collection of crystal structures. *J. Appl. Crystallogr.* **2009**, *42*, 726–729.

- [84] Trucano, P.; Chen, R. Structure of graphite by neutron diffraction. *Nature* **1975**, *258*, 136–137.
- [85] Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **2006**, *125*, 224106.
- [86] Popov, I. V.; Görne, A. L.; Tchougréeff, A. L.; Dronskowski, R. Relative stability of diamond and graphite as seen through bonds and hybridizations. *Phys. Chem. Chem. Phys.* **2019**, *21*, 10961–10969.
- [87] Kim, M.-C.; Sim, E.; Burke, K. Understanding and reducing errors in density functional calculations. *Phys. Rev. Lett.* **2013**, *111*, 073003.
- [88] Woodcock, H. L.; Schaefer, H. F.; Schreiner, P. R. Problematic energy differences between cumulenes and poly-yne: Does this point to a systematic improvement of density functional theory? *J. Phys. Chem. A* **2002**, *106*, 11923–11931.
- [89] Heaton-Burgess, T.; Yang, W. Structural manifestation of the delocalization error of density functional approximations: C₄N₂ rings and C₂₀ bowl, cage, and ring isomers. *J. Chem. Phys.* **2010**, *132*, 234113.
- [90] Johnson, E. R.; Becke, A. D. A unified density-functional treatment of dynamical, nondynamical, and dispersion correlations. II. Thermochemical and kinetic benchmarks. *J. Chem. Phys.* **2008**, *128*, 124105.
- [91] Johnson, E. R.; Clarkin, O. J.; DiLabio, G. A. Density Functional Theory Based Model Calculations for Accurate Bond Dissociation Enthalpies. 3. A Single Approach for X-H, X-X, and X-Y (X, Y = C, N, O, S, Halogen) Bonds. *J. Phys. Chem. A* **2003**, *107*, 9953–9963.
- [92] Hafner, J. Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *J. Comput. Chem.* **2008**, *29*, 2044–2078.
- [93] Gruber, T.; Grüneis, A. Ab initio calculations of carbon and boron nitride allotropes and their structural phase transitions using periodic coupled cluster theory. *Phys. Rev. B* **2018**, *98*, 134108.
- [94] Rice, B.; LeBlanc, L. M.; de-la Roza, A. O.; Fuchter, M. J.; Johnson, E. R.; Nelson, J.; Jelfs, K. E. A computational exploration of the crystal energy and charge-carrier mobility landscapes of the chiral [6]helicene molecule. *Nanoscale* **2018**, *10*, 1865.
- [95] Schmidt, J. A.; Weatherby, J. A.; Sugden, I. J.; Santana-Bonilla, A.; Salerno, F.; Fuchter, M. J.; Johnson, E. R.; Nelson, J.; Jelfs, K. E. Computational Screening of Chiral Organic Semiconductors: Exploring Side-Group Functionalization and Assembly to Optimize Charge Transport. *Cryst. Growth Des.* **2021**, <https://doi.org/10.1021/acs.cgd.1c00473>.

- [96] Karamertzanis, P. G.; Pantelides, C. C. Ab initio crystal structure prediction. II. Flexible molecules. *Mol. Phys.* **2006**, *105*, 273–291.
- [97] Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.
- [98] Price, S. L. 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.
- [99] Weatherby, J. A.; Rumson, A. F.; Price, A. J. A.; Otero de la Roza, A.; Johnson, E. R. A Density-Functional Benchmark of Thermal Free-Energy Corrections for Polymorphic Molecular Crystals. *J. Chem. Phys.* **2022**, *156*, 114108.
- [100] Cruz-Cabeza, A. J.; Feeder, N.; Davey, R. J. Open questions in organic crystal polymorphism. *Commun. Chem.* **2020**, *3*, 142.
- [101] Hilfiker, R.; von Raumer, M. *Polymorphism in the Pharmaceutical Industry: Solid Form and Drug Development*; John Wiley & Sons, 2018.
- [102] Gentili, D.; Gazzano, M.; Jones, D.; Cavallini, M. Polymorphism as an additional functionality of materials for technological applications at surfaces and interfaces. *Chem. Soc. Rev.* **2019**, *48*, 2502–2517.
- [103] Oganov, A. R. Crystal structure prediction: reflections on present status and challenges. *Faraday Discuss.* **2018**, *211*, 643–660.
- [104] Rice, B.; LeBlanc, L. M.; Otero-de-la-Roza, A.; Fuchter, M. J.; Johnson, E. R.; Nelson, J.; Jelfs, K. E. A computational exploration of the crystal energy and charge carrier mobility landscapes of the chiral [6]helicene molecule. *Nanoscale* **2018**, *10*, 1865–1876.
- [105] Glass, C. W.; Oganov, A. R.; Hansen, N. USPEX—Evolutionary crystal structure prediction. *Comput. Phys. Commun.* **2006**, *175*, 713–720.
- [106] Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. CALYPSO: A method for crystal structure prediction. *Comput. Phys. Commun.* **2012**, *183*, 2063–2070.
- [107] Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem. Int. Ed.* **2008**, *47*, 2427–2430.
- [108] Reilly, A. M. et al. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Cryst. B* **2016**, *72*, 439–459.

- [109] Iuzzolino, L.; McCabe, P.; Price, S. L.; Brandenburg, J. G. 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.
- [110] Hoja, J.; Tkatchenko, A. First-principles stability ranking of molecular crystal polymorphs with the DFT+MBD approach. *Faraday Discuss.* **2018**, *211*, 253–274.
- [111] Hoja, J.; Ko, H. Y.; Neumann, M. A.; Car, R.; Distasio, R. A.; Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Sci. Adv.* **2019**, *5*.
- [112] Mortazavi, M.; Hoja, J.; Aerts, L.; Quéré, L.; van de Streek, J.; Neumann, M. A.; Tkatchenko, A. Computational polymorph screening reveals late-appearing and poorly-soluble form of rotigotine. *Commun. Chem.* **2019**, *2*, 1–7.
- [113] Whittleton, S. R.; Otero-de-la-Roza, A.; Johnson, E. R. The exchange-hole dipole dispersion model for accurate energy ranking in molecular crystal structure prediction. *J. Chem. Theory Comput.* **2017**, *13*, 441–450.
- [114] Whittleton, S. R.; Otero-de-la-Roza, A.; Johnson, E. R. The exchange-hole dipole dispersion model for accurate energy ranking in molecular crystal structure prediction II: Non-planar molecules. *J. Chem. Theory Comput.* **2017**, *13*, 5332–5342.
- [115] Leblanc, L. M.; Otero-de-la-Roza, A.; Johnson, E. R. Composite and Low-Cost Approaches for Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2018**, *14*, 2265–2276.
- [116] Taylor, C. R.; Mulvee, M. T.; Perenyi, D. S.; Probert, M. R.; Day, G. M.; Steed, J. W. Minimizing Polymorphic Risk through Cooperative Computational and Experimental Exploration. *J. Am. Chem. Soc.* **2020**, *142*, 16668–16680.
- [117] Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M. Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. *J. Chem. Theory Comput.* **2015**, *11*, 332–342.
- [118] Cutini, M.; Civalleri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. Assessment of Different Quantum Mechanical Methods for the Prediction of Structure and Cohesive Energy of Molecular Crystals. *J. Chem. Theory Comput.* **2016**, *12*, 3340–3352.
- [119] Nyman, J.; Pundyke, O. S.; Day, G. M. Accurate force fields and methods for modelling organic molecular crystals at finite temperatures. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15828.

- [120] Heit, Y. N.; Beran, G. J. O. How important is thermal expansion for predicting molecular crystal structures and thermochemistry at finite temperatures? *Acta Cryst.* **2016**, *B72*, 514–529.
- [121] Cook, C.; Beran, G. J. O. Reduced-cost supercell approach for computing accurate phonon density of states in organic crystals. *J. Chem. Phys.* **2020**, *153*, 224105.
- [122] Williams, D. E. Improved intermolecular force field for molecules containing H, C, N, and O atoms, with application to nucleoside and peptide crystals. *J. Comput. Chem.* **2001**, *22*, 1154–1166.
- [123] Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. Role of Electrostatic Interactions in Determining the Crystal Structures of Polar Organic Molecules. A Distributed Multipole Study. *J. Phys. Chem.* **1996**, *100*, 7352–7360.
- [124] Dovesi, R. et al. Quantum-mechanical condensed matter simulations with CRYSTAL. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1360.
- [125] Abraha, A.; William, D. E. Spherical and Aspherical Intermolecular Force Fields for Sulfur Allotropes. *Inorg. Chem.* **1999**, *38*, 4224–4228.
- [126] Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.
- [127] Frisch, M. J. et al. Gaussian 09 Revision E.01. Gaussian Inc. Wallingford CT 2009.
- [128] Day, G. M.; Price, S. L.; Leslie, M. Atomistic calculations of phonon frequencies and thermodynamic quantities for crystals of rigid organic molecules. *J. Phys. Chem. B* **2003**, *107*, 10919–10933.
- [129] Sacchi, P.; Lusi, M.; Cruz-Cabeza, A. J.; Nauha, E.; Bernstein, J. Same or different – that is the question: identification of crystal forms from crystal structure data. *CrystEngComm* **2020**, *22*, 7170–7185.
- [130] Allen, F. H. The Cambridge Structural Database: a Quarter of a Million Crystal Structures and Rising. *Acta Cryst. B* **2002**, *58*, 380–388, Crystal structures can be obtained free of charge via http://www.ccdc.cam.ac.uk/data_request/cif.
- [131] Otero-de-la Roza, A.; Johnson, E. R.; Luaña, V. Critic2: A program for real-space analysis of quantum chemical interactions in solids. *Comput. Phys. Commun.* **2014**, *185*, 1007–1018.
- [132] Mayo, R. A.; Johnson, E. R. Improved Quantitative Crystal-Structure Comparison using Powder Diffractograms via Anisotropic Volume Correction. *CrystEngComm* **2021**, *23*, 7118–7131.

- [133] de Gelder, R.; Wehrens, R.; Hageman, J. A. A generalized expression for the similarity of spectra: Application to powder diffraction pattern classification. *J. Comput. Chem.* **2001**, *22*, 273–289.
- [134] Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. *Ab initio* molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Comm.* **2009**, *180*, 2175–2196.