Method Development for Benchmarking Key Interactions in Quantum Crystallography

Dissertation

for the award of the degree
"Doctor rerum naturalium" (Dr. rer. nat.)
of the Georg-August-Universität Göttingen

UNIVERSITÄT
GÖTTINGEN

within the doctoral program 'Chemistry'
of the Georg-August University School of Science (GAUSS)

submitted by
Paul Niklas Ruth
from München

Göttingen, 2022

Thesis Committee

Prof. Dr. Dietmar Stalke, Fakultät für Chemie, Institut für Anorganische Chemie, Georg-August-Universität Göttingen

Prof. Dr. Martin Suhm, Fakultät für Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Dr. Carlo Gatti, Consigle Nazionale delle Ricerche SCITEC, Istituto di Scienze e Tecnologie Chimiche, Milan

Members of the Examination Board

Reviewer: Prof. Dr. Dietmar Stalke, Fakultät für Chemie, Institut für Anorganische Chemie, Georg-August-Universität Göttingen

Second Reviewer: Prof. Dr. Martin Suhm, Fakultät für Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Further members of the Examination Board:

Prof. Dr. Burkhard Geil, Fakultät für Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Prof. Dr. Ricardo Mata, Fakultät für Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Jun.-Prof. Dr. Anna Krawczuk, Institut für Anorganische Chemie, Georg-August-Universität Göttingen

Dr. Lisa Vondung, Institut für Anorganische Chemie, Georg-August-Universität Göttingen

Date of the oral examination: 23.02.23

Contents

Acronyms

ADP Atomic Displacement Parameter

DFT Density Functional Theory

ELMO Extremely Localised Molecular Orbital

FFT Fast Fourier Transform

HAR Hirshfeld Atom Refinement

IAM Independent Atom Model

LCAO Linear Combination of Atomic Orbitals

NMR Nuclear Magnetic Resonance

PAW Projector Augmented Wave

1. Introduction

The crystal structure of a material can be paramount to its properties. It can induce optical properties, determine catalytic activity, and can be connected to myriad other properties of interest for industrial applications or just for human curiosity. The crystalline state is unique in that the relative position between molecules is usually more stable and localised in comparison to other phases. The ordered aggregation is governed by localised interactions such as covalent and hydrogen bonds, dipole-dipole interactions, halogen and chalcogen bonding and less localised interactions such as ionic bonds and van-der-Waals interactions.

As these interactions determine the crystal properties, their modelling is of constant interest. From quantum chemical methods, Density Functional Theory (DFT) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965) can certainly be called the most successful method for the description of crystalline materials. In order to verify the performance of a given method within this framework, various benchmark datasets have been provided.

The benchmarks for molecular structures are often based on calorimetric measurements, such as the X23 dataset (Otero-de-la Roza and Johnson, 2012*a*; Dolgonos *et al.*, 2019), ICE10 (Brandenburg *et al.*, 2015), POLY59 (Brandenburg and Grimme, 2016) and the most recent example by Díaz Mirón and Stein (2022), although other datasets probing different properties have also been provided, such as the use of ^{13}C shielding tensors from the solid structures of peptides (Czernek *et al.*, 2012).

On the other hand, most interactions in the crystal (with the exclusion of core-core Coulomb repulsion) are governed by and influence the spatial distribution of electrons within a crystal and as such the study of this distribution within atoms and crystals has always been a topic of interest. The study of electron density can be separated along the usual lines of theory and experiment.

The study of the crystalline state is the field of crystallography and one of its most successful experimental methods is X-ray diffractometry. As Peter Debye wrote in 1915, albeit about the study of amorphous matter:

> *[...] then it seems to me that the experimental investigation of diffracted radiation, particularly for light atoms, warrants raised interest, as this should be the way to experimentally discern the arrangement of electrons within the atom. Accordingly, such an investigation has the significance of an ultra-microscopy of the inner atom.* (translated from the German original from Debye (1915))

In the same year, Compton (1915) successfully excluded models of electron density distributions of the atoms in rock salt, based on the X-ray diffraction experiments of Bragg and Bragg (1913). An early attempt to derive a radial electron density distribution from X-ray diffraction dates back to the 1920s (Bragg *et al.*, 1922). However, due to experimental uncertainties, the determined electron distribution does not match the correct number of electron shells for the sodium and chlorine atoms in the crystal.

The study of the behaviour of electrons from the theoretical side is of course described by quantum mechanics. The field of crystal structure analysis was able to profit significantly from the atomic quantum electron densities calculated by methodologies from quantum mechanics (Hartree and Waller, 1929) and atomic densities calculated by quantum chemical methods have since become the basis of the ubiquitous independent atom model (Doyle and Turner, 1968).

Starting from theoretical atomic densities a field emerged in crystallography, where these densities were modified to fit experimental intensities from the X-ray diffraction experiment, the latest approach of which is the model by Hansen and Coppens (1978) now ubiquitous in charge density refinement. The obtained densities were approximately deconvoluted from the vibrational influence and could subsequently be compared to DFT calculations (Lippmann *et al.*, 2003; Volkov *et al.*, 2000). A more recent example postulated an influence of dispersion interaction, which a recent van-der-Waals functional was not able to reproduce (Zhang *et al.*, 2018).

The final set of methods tries to combine wavefunction calculations and intensities from X-ray diffractometry during the refinement of the model itself, effectively trying to reach a compromise of the energy of the wavefunction and the agreement to the experimental values. X-ray restrained wavefunction fitting introduced by Jayatilaka and Grimwood (2001) is currently the most prominent of these methods. These methods are occupied with getting a more accurate wavefunction with some impressive results (Genoni *et al.*, 2018). However, to this date calculations do not include periodic boundary conditions and therefore the difference between the theoretical wavefunction and the wavefunction including experimental influences cannot be attributed to inaccuracies in the theoretical method of choice. This makes its use limited for the benchmarking purposes pursued in this work until this drawback is remedied.

The final approach is splitting up the theoretically calculated density into atomic contributions for further use. The most prominent of these methods is Hirshfeld Atom Refinement (HAR) (Jayatilaka and Dittrich, 2008; Capelli *et al.*, 2014) and its improvement and use will be explored in Chapter 6.

The marriage of crystallography and quantum mechanics has been coined *quantum crystallography*. While this term was originally coined to only describe the combination of crystallographic data and quantum mechanical techniques to gain additional or more accurate information (Huang *et al.*, 1999), the term has recently been used in a broader sense, for example also including the aforementioned charge density refinements without

a direct quantum mechanical calculation. The more recent definition used in this work is given by Macchi (2020):

> quantum crystallography includes the study of all those features of a crystal, which cannot be described only by classical or semi-classical physics.

Using this definition this work will try to evaluate different approaches for a benchmarking approach of the solid-state system. For every chapter, it will rely on approaches which can be included in the definition of quantum crystallography. However, each chapter will use varying broadness of the used term.

2. Kinematic theory of X-ray diffraction

This chapter is meant as a review of the theory of X-ray diffraction employed within the following chapters including the connection of X-ray scattering and electron density. It will stay within the *kinematical* approximation, *i. e.* all presented derivations work under the assumption that X-ray radiation is scattered only or mainly elastically by crystals once, while non-elastic contributions are very small in comparison.

This approximation is very successful, as can be seen by the over 1 200 000 structures in the Cambridge Structural Database,[1] which are almost exclusively refined using kinematical models. In order to include essential non-kinematical effects, the crystallographic community has developed corrections, as will be discussed in section 2.7.

The explanation will start from the most simple system imaginable, namely a one-dimensional grating. From there on, the diffraction of X-ray radiation at a crystal will be developed. In contrast, the final two sections will deal with how the model for the intensity of the X-ray experiment can be fitted to the recorded intensity. The final section will tackle how the quality of the fitted model can be assessed.

2.1. Diffraction from a one-dimensional grating

This section follows the derivation as presented by Schwarzenbach (1996). The basis of all diffraction experiments is a wavefront of electromagnetic radiation. The direction and wavelength of this wavefront are given by the wave vector of the incident wave $\mathbf{s_0}$. The direction of the vector is the direction of propagation of the electromagnetic radiation. The magnitude of the vector is given as $|\mathbf{s_0}| = 1/\lambda$. For this wave vector, the displacement of the electric field ψ at position $\mathbf{r}$ at time t is given as:

$$\psi(\mathbf{r}, t) = Ae^{2\pi i(\mathbf{s_0}\cdot\mathbf{r}-\nu t+\phi)} \tag{2.1}$$

here A denotes the amplitude and λ the wavelength. ν is the frequency and ϕ is the phase of the wave at $|\mathbf{r}| = 0$ and $t = 0$. If a second wavefront of light is added, the superposition of the two vectors will be obtained. With the phases of the two waves ϕ_1 and ϕ_2 and the amplitudes A_1 and A_2 and the same direction of propagation, the superposition can be described as:

[1] The number is taken from the CCDC website in December 2022

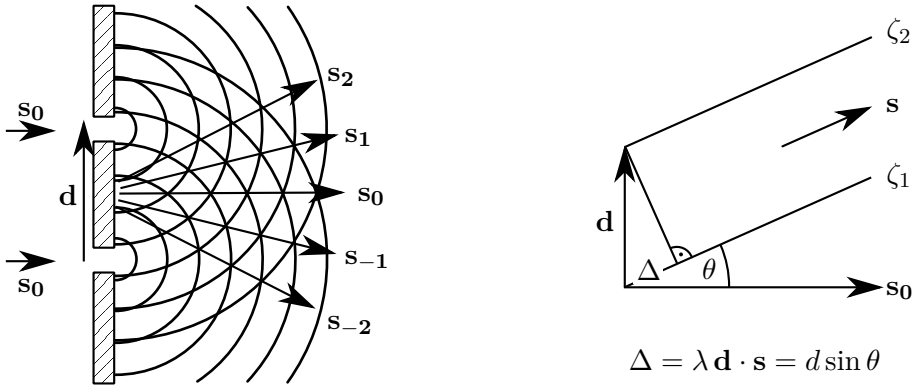


Figure 2.1.: Diffraction of electromagnetic radiation in the double-slit experiment. $\mathbf{d}$ is the distance between the two slits. $\mathbf{s}$ is a wave vector in the direction of diffraction. $\mathbf{s_0}$ is the wave vector of the incident wave. Reproduced after Schwarzenbach (1996).

$$\zeta = \zeta_1 + \zeta_2 = A_1 e^{2\pi i \phi_1} + A_2 e^{2\pi i \phi_2} \tag{2.2}$$

In the simplest diffraction experiment, a parallel wavefront hits a wall that is impenetrable, except for two small punctures (see Figure 2.1). From Huygens' principle follows, that these two punctures constitute two new sources of electromagnetic waves, which are spherically emitted from the punctures. Assuming the distance of observation is large compared to the spacing between these two punctures $d = |\mathbf{d}|$, the product of the spherical waves is a new plane wave with the new wave vector $\mathbf{s}$ (Fraunhofer's approximation). The path difference between these two waves ζ_1 and ζ_2 can be deduced from Figure 2.1 to be:

$$\Delta = d \sin\theta = \lambda\, \mathbf{d} \cdot \mathbf{s} \tag{2.3}$$

The relative phase difference is constant and therefore the first phase can be set to $\phi(\zeta_1) = 0$. Due to the path difference, the relative phase of the second wave has to be $\phi(\zeta_2) = \Delta/\lambda$ and according to Equation 2.2 the result is:

$$\zeta = \zeta_1 + \zeta_2 = A(e^{2\pi i 0} + e^{2\pi i \Delta/\lambda}) = A(1 + e^{2\pi i \mathbf{d}\cdot\mathbf{s}}) \tag{2.4}$$

The intensity $I(\mathbf{s})$ in direction $\mathbf{s}$ is proportional to the squared absolute value of ζ:

$$I(\mathbf{s}) = |\zeta|^2 = 4A^2 \cos^2(\pi \mathbf{d} \cdot \mathbf{s}) = 4A^2 \cos^2\left(\pi \frac{d}{\lambda} \sin\theta\right) \tag{2.5}$$

Now, this expression needs to be generalised for a linear arrangement of N punctures in a row. In addition to the point sources at 0 and $\mathbf{d}$, point sources are located at $2\mathbf{d}$, $3\mathbf{d}$, $\ldots$, $(N-1)\mathbf{d}$. Accordingly, the resulting phase differences are $\mathbf{s}\cdot\mathbf{d}$, $2\mathbf{s}\cdot\mathbf{d}$, $\ldots$, $(N-1)\mathbf{s}\cdot\mathbf{d}$. Combining the sum of geometric progression to evaluate the sum of exponential functions with Equation 2.2 yields:

$$\zeta = \sum_{n=1}^{N} \zeta_n = A \sum_{n=0}^{N-1} e^{2\pi i n \mathbf{s}\cdot\mathbf{d}} = A \frac{1 - e^{2\pi i N \mathbf{s}\cdot\mathbf{d}}}{1 - e^{2\pi i \mathbf{s}\cdot\mathbf{d}}} \tag{2.6}$$

$$I = |\zeta|^2 = \zeta\zeta^* = A^2 \frac{1 - \cos(2\pi N \mathbf{s}\cdot\mathbf{d})}{1 - \cos(2\pi \mathbf{s}\cdot\mathbf{d})} = A^2 \frac{\sin^2(\pi N \mathbf{s}\cdot\mathbf{d})}{\sin^2(\pi \mathbf{s}\cdot\mathbf{d})} = A^2 J_N^2(\mathbf{s}\cdot\mathbf{d}) \tag{2.7}$$

$$J_N^2(\mathbf{s}\cdot\mathbf{d}) = \frac{\sin^2(\pi N \mathbf{s}\cdot\mathbf{d})}{\sin^2(\pi \mathbf{s}\cdot\mathbf{d})} \tag{2.8}$$

The function $J_N^2(\mathbf{s}\cdot\mathbf{d})$ is the interference function of a row of evenly spaced scattering centres, which could be viewed as a one-dimensional crystal. Its root $J_N(\mathbf{s}\cdot\mathbf{d})$ is the one-dimensional equivalent of the so-called structure factor for this specific system of equal scatterers.

The structure factor describes the scattering of an object and is defined in reciprocal space, which is spanned by possible diffraction vectors $\mathbf{S} = \mathbf{s} - \mathbf{s_0}$. For any scattering system, it can be written as:

$$I(\mathbf{S}) \propto |F(\mathbf{S})|^2 \tag{2.9}$$

2.2. Fourier transform and convolution in crystallography

In the next section, a three-dimensional lattice of electron density will be constructed to describe the crystal. However, two mathematical tools are needed for that purpose, which need to be described first. Namely, this section will introduce the concepts of Fourier Transformation and convolution. This will not be done in a mathematically rigorous way but will take more of a picture-book approach. For a rigorous treatment, the reader might consult the established literature (Bricogne, 2010).

Fundamentally, the Fourier Transform extracts the wave component from a function in time or space. The latter will be used in this section as that is the type of waves which are usually investigated in crystallography. For simplicity reasons, this section will start with a sum of two Gaussian functions as the target of the example investigation. Starting with the one-dimensional case, the formula for the Fourier Transform is:

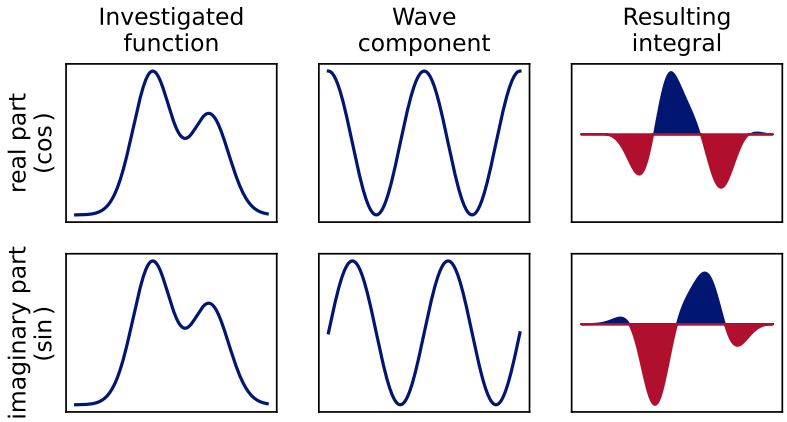


Figure 2.2.: Depiction of the calculation of Fourier transform for one point in reciprocal space for the real-valued function. The plots on the right are the product of the other two plots. The resulting imaginary and real space components are the difference between the blue and red areas of these plots.

$$\mathcal{F}[f(x)](\zeta) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \zeta x} dx \tag{2.10}$$

It is illustrative to visualise the integral of the product of these two functions for a given wavelength ζ, but first, the real and imaginary parts of this integration need to be separated using Euler's formula:

$$\mathcal{F}[f(x)](\zeta) = \int_{-\infty}^{\infty} f(x) \left[\cos\left(-2\pi\zeta x\right) + i \sin\left(-2\pi\zeta x\right)\right] dx \tag{2.11}$$

Now, the integrals for calculating the imaginary and real parts can be visualised separately. As can be seen in Figure 2.2 the calculation of the Fourier transform for a given function and wave vector is relatively straightforward. The real part is related to the cosine of the wave with the given wavelength. It is calculated by multiplying with a cosine function constructed with the wave vector and subsequently determining the integral. The imaginary part is constructed with the sine function accordingly. Following this recipe for each wavelength gives the reciprocal spectrum or the Fourier transform of a given function constructed in real space.

Of course, there is an analytical solution for the investigated function, as the Fourier transform is a linear transformation, which means that both Gaussian functions can be investigated separately. However, the graphical illustration shall suffice for the purpose of this work.

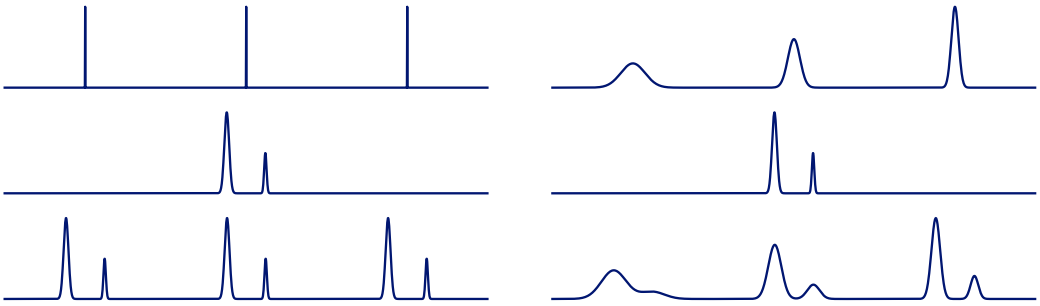


Figure 2.3.: Illustration of the effect of convolution. The first two rows are convoluted with each other to yield the two functions in the lower row.

The inverse operation is called the inverse Fourier transform, so for a function $F(\zeta)$:

$$\mathcal{F}^{-1}F(\zeta) = \int_{-\infty}^{\infty} F(\zeta)e^{2\pi i\zeta x}d\zeta \tag{2.12}$$

A normalisation for either the inverse Fourier transform or both directions was omitted, as the convention for this is not consistent in literature.

For the three-dimensional case, the integral is naturally done over all three dimensions. In preparation for later discussion, the inverse Fourier transform is applied to an arbitrary density $\rho(\mathbf{x})$. The resulting function is the structure factor $F(\mathbf{S})$ from the last section.

$$F(\mathbf{S}) = \mathcal{F}\rho(\mathbf{x}) = \int_{\mathbb{R}^3} \rho(\mathbf{x})e^{2\pi i\mathbf{S}\cdot\mathbf{x}}d\mathbf{x} \tag{2.13}$$

The structure factor is a complex number and determines the relative scattering intensity of reflections in crystallography.

For the second part of this section, the concept of convolution shall be investigated in the same manner. Imagine two functions $a(x)$ and $b(x)$. For each point of the first function $a(x)$, the second function $b(x)$ is multiplied by the value of the first function and then the function $b(x)$ is drawn around the point of the first function $a(x)$. The integral obtained by applying this recipe to all functions is the resulting convolution. This concept is graphically illustrated in Figure 2.3.

The convolution theorem connects the Fourier transform of the resulting convolution to the product of the Fourier transforms of the individual functions.

$$\mathcal{F}\left[a(\mathbf{x}) * b(\mathbf{x})\right] = \mathcal{F}\left(a(\mathbf{x})\right) \cdot \mathcal{F}\left(b(\mathbf{x})\right) \tag{2.14}$$

This equation will be used in the next sections to construct the structure factor of our complete lattice from individual components.

2.3. Constructing the electron density of an infinite lattice and its (inverse) Fourier transform

The crystal lattice is defined by the unit cell which is composed of a parallelepiped spanned by three unit vectors **a**, **b** and **c**. These parallelepipeds form a precise lattice, which is assumed to be perfect and infinite in all three directions. The electron density within one of these cells is given by $\rho_{\text{cell}}(\mathbf{r})$. How this density can be described will be the subject of the next two sections. For now, it just needs to be accepted that it exists.

There are two conventions of coordinates to use within one cell of the lattice. The vector **r** shall denote a position in cartesian coordinates. In contrast, crystallography usually uses *fractional coordinates* for atomic positions with the associated vector **x**. The fractional coordinates are only unique within the range $0 \leq x_i \leq 1$. Larger or smaller values correspond to equivalent positions in neighbouring unit cells. The coordinates can easily be transformed from one convention into the other by using.

$$\mathbf{r} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \mathbf{x} = \mathbf{M}\mathbf{x} \tag{2.15}$$

where a_1, a_2 and a_3 denote the components of the unit vector **a** and b_i and c_i are the components of **b** and **c**, respectively.

The lattice itself can be defined as an infinite sum of delta functions, where each delta function describes one point of the lattice. A lattice point is only found if there is a linear combination of the unit vectors with integer factors which gives the coordinate. Let u, v, w be integer values. Then a coordinate r_{latt} only belongs to the lattice if there is a combination of u, v, w which yields:

$$r_{\text{latt}} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c} \tag{2.16}$$

The function describing the lattice $\delta_{\text{latt}}(\mathbf{r})$ can be constructed as the sum of all possible δ functions that fulfil this condition. One can simply write:

$$\delta_{\text{latt}}(\mathbf{r}) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} \delta(\mathbf{r} - u\mathbf{a} - v\mathbf{b} - w\mathbf{c}) \tag{2.17}$$

The overall density of the lattice can now be described as a convolution between this function and the density of an individual cell:

$$\rho_{\text{latt}}(\mathbf{r}) = \delta_{\text{latt}}(\mathbf{r}) * \rho_{\text{cell}}(\mathbf{r}) \tag{2.18}$$

Using Equation 2.14 the structure factor can be calculated:

$$F_{\text{latt}}(\mathbf{S}) = \mathcal{F}\left[\delta_{\text{latt}}(\mathbf{r}) * \rho_{\text{cell}}(\mathbf{r})\right] \tag{2.19}$$
$$= \mathcal{F}[\delta_{\text{latt}}(\mathbf{r})] \cdot \mathcal{F}[\rho_{\text{cell}}(\mathbf{r})] \tag{2.20}$$

This means that for the quantities of interest, the Fourier transform of the lattice and the Fourier transform of the density can be done separately. The product only needs to be formed at the end. The description of the unit cell will be the topic of the following section. It can be shown that the Fourier transform of the lattice function is another lattice. It has its own three basis vectors in reciprocal space: $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$.

What this means in practice is that the structure factor of a single unit cell can be probed at the lattice points of this *reciprocal lattice* only. The benefit is that the measured intensity from all the individual cells is added up by constructive interference, thereby amplifying the signal to something, which is available even with X-ray sources of common intensity. In kinematic approximation the structure factor can only be non-zero at points in reciprocal space fulfilling the condition:

$$\mathbf{S} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \tag{2.21}$$

where h, k, and l are the Miller indices and can only have integer values. As such our reciprocal space can be mapped in these indices using the vector:

$$\mathbf{h} = \begin{pmatrix} h \\ k \\ l \end{pmatrix} \tag{2.22}$$

2.4. The average density at finite temperature

This section aims to start the description of the electron density of the unit cell. It will introduce why the unit cell is constructed from atomic contributions, while the approaches for constructing the atomic densities will follow in the next section. Important for now is the fact, that the atomic density $\rho_{\text{atom}}(\mathbf{r})$ is constructed with the respective atom being located on the origin of the coordinate system. It can now be moved onto its position in the unit cell by a convolution with a δ function, which is one at the atomic position and zero everywhere else. If there are N_{atom} atoms in our unit cell the density in the unit cell can be described as:

$$\rho_{\text{cell}}(\mathbf{r}) = \sum_{j=1}^{N_{\text{atom}}} \rho_{\text{atom},j} * \delta(\mathbf{r} - \mathbf{r_j}) \tag{2.23}$$

With the knowledge that the Fourier transformation is a linear transformation and the definition of the convolution, the structure factor can then be calculated:

$$F_{\text{cell}}(\mathbf{S}) = \mathcal{F}\left[\sum_{j=1}^{N_{\text{atom}}} \rho_{\text{atom},j} * \delta(\mathbf{r} - \mathbf{r_j})\right] \tag{2.24}$$

$$= \sum_{j=1}^{N_{\text{atom}}} \mathcal{F}[\rho_{\text{atom},j}(\mathbf{r})] \cdot \mathcal{F}[\delta(\mathbf{r} - \mathbf{r_j})] \tag{2.25}$$

$$= \sum_{j=1}^{N_{\text{atom}}} f_j(\mathbf{S}) \cdot e^{2\pi i \mathbf{S}\cdot\mathbf{r_j}} \tag{2.26}$$

The density as a sum of atomic densities has now been successfully constructed. However, in the investigated physical system, atoms are not strictly located in a single position. Atoms are subject to thermal motion even at low temperatures, as they are subject to zero-point vibration. This effect is included by another convolution. The atom can be displaced from its equilibrium position in dependence on the displacement vector $\mathbf{u}$ with the probability $P(\mathbf{u})$. The thermally averaged density is now a convolution of our atomic densities at their position with this probability function:

$$\langle \rho_{\text{cell}}(\mathbf{r}) \rangle_T = \rho_{\text{cell}}(\mathbf{r}) = \sum_{j=1}^{N_{\text{atom}}} \rho_{\text{atom},j} * P(\mathbf{u}) * \delta(\mathbf{r} - \mathbf{r_j}) \tag{2.27}$$

The structure factor becomes:

$$F(\mathbf{S}) = \sum_{j=1}^{N_{\text{atom}}} \mathcal{F}[\rho_{\text{atom},j}(\mathbf{r})] \cdot \mathcal{F}[P(\mathbf{u})] \cdot \mathcal{F}[\delta(\mathbf{r} - \mathbf{r_j})] \tag{2.28}$$

$$= \sum_{j=1}^{N_{\text{atom}}} f_{0j}(\mathbf{S}) \cdot T(\mathbf{S}) \cdot e^{2\pi i \mathbf{S}\cdot\mathbf{r_j}} \tag{2.29}$$

Here, f_{0j} is the atomic form factor, $T(\mathbf{S})$ is the temperature factor and $e^{2\pi i \mathbf{S}\cdot\mathbf{r_j}}$ is the so-called phase factor. Using this description the temperature can be included elegantly with a single additional term in our product. However, the underlying approximation is,

that the atomic density is invariant with displacement. The lower the displacement in the system is, the more this approximation is fulfilled.

The temperature factor is the Fourier transform of the probability density. The simplest assumption is that the probability is a Gaussian function which is only dependent on the distance and not on the direction of the displacement. In this case, the functions take the following values (Coppens, 2010):

$$P(u) = \frac{e^{-|u|^2/2\langle u^2\rangle}}{(2\pi\,\langle u^2\rangle)^{3/2}} \tag{2.30}$$

$$T(|\mathbf{S}|) = \exp\left(-2\pi^2 U_{\text{iso}}\,|\mathbf{S}|^2\right) \tag{2.31}$$

In crystallography, this approach is called an *isotropic* refinement of the atomic displacement, with the isotropic displacement parameter U_{iso}. However, atoms are not vibrating independently of each other, but displacements in bond directions are less likely than displacements perpendicular to bonds. Vibrations of atoms are coupled to each other. To account for these effects the probability is assumed to be dependent on the direction of the displacement with an inversion symmetric trivariate Gaussian function. As such the probability density and temperature factor can be calculated (Coppens, 2010):

$$P(\mathbf{u}) = \frac{\left|\boldsymbol{\sigma}^{-1}\right|^{1/2}}{(2\pi)^{3/2}} \exp\left\{-\frac{1}{2}\mathbf{u}^T\boldsymbol{\sigma}^{-1}\mathbf{u}\right\} \tag{2.32}$$

$$T(\mathbf{S}) = \exp\left\{-2\pi^2\mathbf{S}^T\boldsymbol{\sigma}\mathbf{S}\right\} \tag{2.33}$$

Modelling the thermal motion in this way is called *anisotropic* refinement of the atomic displacement. The variance-covariance matrix $\boldsymbol{\sigma}$ is symmetric, which means that the displacement can be described with six parameters. In small molecule crystallography, a slightly different convention for this temperature factor has been established. First, the matrix $\mathbf{N}$ is defined as (Grosse-Kunstleve and Adams, 2002):

$$\mathbf{N} = \begin{pmatrix} |\mathbf{a}| & 0 & 0 \\ 0 & |\mathbf{b}| & 0 \\ 0 & 0 & |\mathbf{c}| \end{pmatrix} \tag{2.34}$$

with $\mathbf{U}_{\text{cif}}$ again being a symmetric matrix and $\mathbf{h}$ being a vector with the miller indices $\mathbf{h} = (h\,k\,l)^T$, the temperature factor can be defined as:

$$T(\mathbf{S}) = \exp\left\{-2\pi^2\mathbf{h}^T\mathbf{N}\mathbf{U}_{\text{cif}}\mathbf{N}^T\mathbf{h}\right\} \tag{2.35}$$

This is also the convention which will be used throughout this work. This is due to consistency with the main storage file format in crystallography: the 'cif' file and easy comparison to other results.

It is possible to calculate the equivalent isotropic displacement parameter from the anisotropic refinement. It is calculated by the trace of the transformed atomic displacement matrix using $\mathbf{N}$ and the $\mathbf{M}$ from Equation 2.15.

$$U_{\text{equiv}} = \text{tr}\left(\mathbf{MNU}_{\text{cif}}\mathbf{N}^T\mathbf{M}^T\right)/3 \tag{2.36}$$

So far, only harmonic displacement has been investigated. Dependent on the actual system, atoms might be displaced according to probability functions no longer well described by a multivariate Gaussian. In this case, the Gram-Charlier expansion proposed by Johnson and Levy (1974) has proven successful. It is an expansion of the normal distribution in terms of its derivatives (Coppens, 2010):

$$P(\mathbf{u}) = \left[1 - \sum_j D_j + \frac{1}{2!}\sum_j\sum_k c_{jk}D_jD_k - \frac{1}{3!}\sum_j\sum_k\sum_l c_{ijk}D_jD_kD_l + ...\right]P_0(\mathbf{u}) \tag{2.37}$$

where $i, j, k \in 1, 2, 3$ denominate the cartesian coordinate of the derivative operator $\partial/\partial u_i$. Higher derivatives are achieved by applying the operator multiple times. $P_0(\mathbf{u})$ is the unmodified normal distribution displacement function. With the use of Hermite polynomials (see Coppens (2010)). This can be transformed to:

$$P(\mathbf{u}) = \left[1 + \frac{1}{3!}\sum_j\sum_k\sum_l c_{ijk}H_{jkl}(\mathbf{u}) + \frac{1}{4!}\sum_j\sum_k\sum_l\sum_m c_{jklm}H_{jklm}(\mathbf{u}) + ...\right]P_0(\mathbf{u}) \tag{2.38}$$

The reciprocal space representation of this anharmonic displacement description is given as:

$$T(\mathbf{S}) = \left[1 - \frac{4}{3}\pi^3 i\sum_j\sum_k\sum_l c_{jkl}S_jS_kS_l + \frac{2}{3}\pi^4\sum_j\sum_k\sum_l\sum_m c_{jklm}S_jS_kS_lS_m + ...\right]T_0(\mathbf{S}) \tag{2.39}$$

Here T_0 is the harmonic temperature factor and the individual S_i are the components of the reciprocal coordinate of the vector $\mathbf{S}$.

As such, a flexible description of the atomic displacement is available. However, a model for the electron density of the individual atoms is still necessary. Some available options will be given in the next section.

2.5. Atomic density models

As mentioned this section will deal with descriptions of the atomic electron density. The focus will be on the three approximations used in this work (Independent Atom Model, Multipolar Model and Hirshfeld Atom refinement). The explanations will be following some of the historical perspective from the excellent review from Genoni and Macchi (2020).

Electron density can only be suitably described by quantum mechanics and as such the foundation for the electron density has been derived from quantum mechanical methodologies. An early model for atomic form factors based on classical Thomson theory adapted to the Bohr atomic model (Hartree, 1925) did show poor agreement to experimental measurements (Bragg *et al.*, 1922). Subsequently, Waller and Hartree (1929) were able to provide more accurate atomic form factors based on atomic quantum electron densities, which were adopted successfully. The idea to use spherical atomic densities as an approximation for the atomic contribution to the scattering electron density is still in use today in the form of the *Independent Atom Model*, where the atomic form factors in use were derived by Doyle and Turner (1968) and subsequently published in a parametric form in the International Tables for Crystallography, Volume C (Prince *et al.*, 2006).

However, as soon as atoms start to interact with each other the assumption of spherically symmetric densities is no longer valid. Additionally, electron densities will be redistributed from one atom to another. There are two different approaches in crystallography how to include this fact in our density description. One approach is to modify an atomic description in suitable ways. The other approach is to start with a molecular density description, which can be divided into atomic contributions.

2.5.1. The atomic basis for aspherical atoms

Initially, the derivation of improved atomic descriptions has been pushed by Steward. Building on the concept of the generalised atomic form factor by Dawson (1967), Stewart (1969) had the idea to project the calculated molecular densities onto atomic bases and to derive the atomic form factors for X-ray refinement from these projections (Stewart, 1973). As the direct refinement of the wavefunction coefficients proved to be too complicated at this point in time, the model was applied differently, by directly refining atomic parameters against experimental data. As such Stewart (1976) proposed an atom-centred multipolar extension and the derived model became known as the *multipole model*. This work uses the formalism proposed by Hansen and Coppens (1978), which introduces

local coordinate systems, thereby enabling a larger degree of transferability between atomic densities.

The atomic density in this multipolar approach is described by (Coppens, 2010):

$$\rho_{\text{atomic}}(\mathbf{r}) = P_{\text{c}}\rho_{\text{core}}(r) + P_{\text{v}}\kappa^3\rho_{\text{valence}}(\kappa r) + \sum_{l=0}^{l_{\max}} \kappa'^3 R_l(\kappa' r) \sum_{m=0}^{l} \sum_{p} P_{lmp} d_{lmp}(\mathbf{r}/r) \quad (2.40)$$

where $\rho_{\text{core}}(r)$ is the atomic core density, with the coefficient P_{c}, $\kappa^3\rho_{\text{valence}}(\kappa r)$ is the spherical atomic valence density, with the coefficient P_{v}, which is often called the monopole parameter. The parameter κ can be used to vary the distance dependency within the refinement (Coppens *et al.*, 1979). The last part of this sum is the aspherical valence density and is described as a sum of normalised spherical harmonic functions d_{lmp} with their coefficients P_{lmp}, which are multiplied by functions for the distance dependency $\kappa'^3 R_l(\kappa' r)$. This dependency can be modified by multiplying the absolute of the distance from the atomic position by an additional parameter κ'.

The expression for the distance dependency is usually chosen from either a Gaussian or a Slater function and is provided in the form of precalculated databases of atomic density functions derived from quantum chemical calculations. The first of these has been provided by Clementi and Raimondi (1963). This work uses the functions from Su and Coppens (1998), which were extended by Macchi and Coppens (2001).

The Fourier Transform is a linear transform and as such the parts of the atomic density can be transformed individually to get our atomic form factors in reciprocal space. The atomic form factor expression is given by:

$$f(\mathbf{S}) = P_{\text{core}} f_{\text{core}}(S) + P_{\text{valence}} f_{\text{valence}}(S/\kappa) + \sum_{l} \sum_{m} \sum_{p} 4\pi i^l \langle j_l \rangle (S/\kappa') d_{lmp}(\mathbf{S}/S) \quad (2.41)$$

The exact expression for the core atomic form factor $f_{\text{core}}(S)$ and the valence atomic form factor $f_{\text{valence}}(S/\kappa)$ depends on the function chosen for the description of the distance dependence of the density. $\langle j_l \rangle (S/\kappa')$ is the Fourier-Bessel transform and is defined as:

$$\langle j_l \rangle (S/\kappa') = \int_0^\infty j_l(2\pi r S/\kappa') R_l(r) r^2 dr \quad (2.42)$$

where j_l is the lth other Bessel function. Finally, the $d_{lmp}(\mathbf{S}/S)$ function consists of the corresponding spherical harmonic. As the Fourier transform of spherical harmonic functions are the spherical harmonics themselves, the transformation is straightforward.

However, the coordinate system used consists of directions in reciprocal space instead of cartesian coordinates.

2.5.2. Molecular basis for the atomic form factor

In 2008, a new procedure to calculate the atomic form factors was proposed (Jayatilaka and Dittrich, 2008). It does not calculate them starting from atomic densities of atoms or ions but instead partitions a calculated density of a multiatom system into atomic contributions.

First, a representative fragment from our molecular structure within the unit cell is defined and the density of that fragment is calculated *via* computational chemistry methods. In the second step, this density is partitioned using the Hirshfeld stockholder partitioning (Hirshfeld, 1971) scheme. Here, atomic densities are used as weighting functions in order to determine, how much of each point in space should be attributed to which atom.

For each atom, an atomic density of the neutral atom is placed onto the atomic position. The sum of these atomic densities is then called the *promolecular density.*

$$\rho_{\mathrm{promol}}(\mathbf{r}) = \sum_{j=1}^{N_{\mathrm{atoms}}} \rho_{j,\mathrm{neutral\,atom}}(|\mathbf{r} - \mathbf{r}_j|) \tag{2.43}$$

The share of the density at a certain point in space which belongs to the atom j to be evaluated can now be calculated by a simple division.

$$p_{j,\mathrm{Hirshfeld}}(\mathbf{r}) = \frac{\rho_{j,\mathrm{neutral\,atom}}(|\mathbf{r} - \mathbf{r}_j|)}{\rho_{\mathrm{promol}}(\mathbf{r})} \tag{2.44}$$

For a quantum mechanically calculated density $\rho_{\mathrm{system}}(\mathbf{r})$, the individual atomic contributions for each point in space can now be calculated according to the Hirshfeld scheme:

$$\rho_{\mathrm{j,atom}}(\mathbf{r}) = p_{j,\mathrm{Hirshfeld}}(\mathbf{r}) \cdot \rho_{\mathrm{system}}(\mathbf{r}) \tag{2.45}$$

The individual atomic densities are then Fourier transformed to get the atomic form factors needed for the crystallographic refinement, which is used to determine the atomic positions and displacement parameters. For an application and further history of this method, see Chapter 6.

As mentioned in the introduction, methods such as X-ray restrained wavefunction extend this approach. Starting from molecular densities, they fit the wavefunction coefficients to be as low as possible in energy and fit to the X-ray diffraction data at the same time

(Jayatilaka and Grimwood, 2001). However, as this method will not be used as the basis of any of the investigations in this thesis, I will just redirect the interested reader to the aforementioned review (Genoni and Macchi, 2020) for further information.

2.5.3. Non-density contributions to the atomic form factor: Resonance scattering

The electrons can be approximated as bound to a specific nucleus with a strength, that is dependent on the atomic field strength and its quantum state. Therefore they can be considered atomic oscillators with natural frequencies (Giacovazzo *et al.*, 2011). However, this means an incident beam with a wavelength near these frequencies can have resonance with the atomic oscillators. This effect is called resonance scattering or anomalous dispersion. In good approximation, this can be described by modifying the atomic form factor with two additional terms for the resonance scattering:

$$f(\mathbf{S}) = f_\rho(\mathbf{S}) + f'(\lambda) + if''(\lambda) \tag{2.46}$$

To a good approximation, these two factors are not dependent on the diffraction vector. They are, however, highly dependent on the employed wavelength. Usually, the corrections are applied for a given combination of wavelength and element (Creagh, 2010). However recently this approximation has been called into question by Meurer *et al.* (2022), especially if the resonance is strong. The final scope and implications of these issues are an active field of research and discussion at the point of writing.

According to Friedel's law, the intensity of reflections connected by the inversion centre in the origin of reciprocal space should be equal. However, for non-centrosymmetric structures, this is no longer the case when resonance scattering is considered. This effect can be used to determine the absolute structure of the content of the unit cell. In this case, an occupation parameter can be refined between the structure and its inverted counterpart. The intensity is described as:

$$|F(\mathbf{S}, x)| = (1 - x)\,|F(\mathbf{S})| + x\,|F(-\mathbf{S})| \tag{2.47}$$

where x is called the Flack parameter, which should be zero for correct non-centrosymmetric structures. In general, the effect of resonance scattering increases with higher wavelength. As such hard radiation in combination with light elements can often lead to large uncertainty in the Flack parameter. At the same time, the density-dependent part of the atomic form factor decreases in magnitude with higher scattering angles, while the resonant scattering contributions remain similar. As such a large resolution can also increase the certainty in the Flack parameter.

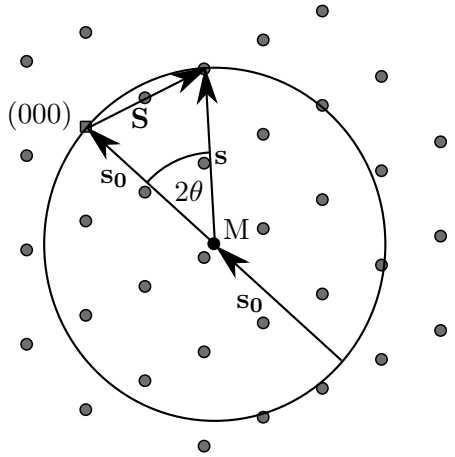


Figure 2.4.: Ewald construction in two dimensions.

2.6. The reciprocal lattice on the detector: The Ewald construction

Section 2.3 demonstrated that the Fourier transform of an ideal crystal is only nonzero if the following equation is fulfilled, *i.e.* if the point in reciprocal space is a point of the reciprocal lattice:

$$\mathbf{S} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \tag{2.21}$$

The question is now how to find out which point in reciprocal space is visible in which direction and in which crystal orientation. Recall that the wave vector of the incident wave is denoted $\mathbf{s}_0$ and the wave vector of the diffracted wave is denoted $\mathbf{s}$. The position in reciprocal space can now be constructed as the difference between these two vectors, which is called the diffraction vector $\mathbf{S}$:

$$\mathbf{S} = \mathbf{s} - \mathbf{s}_0 \tag{2.48}$$

A point on the detector is the point in reciprocal space corresponding to the diffraction vector for a given crystal orientation and diffraction direction. The most useful tool to answer which points are available at a given orientation is the Ewald construction (Figure 2.4).

If the points in the reciprocal lattice correspond to directions from the crystal, the wave vector of the incident wave $\mathbf{s}_0$ can be drawn in a relative orientation pointing to the point (0 0 0) of that lattice. The origin of $\mathbf{s}_0$ is called M. Another way to look at this, is that $-\mathbf{s}_0$ is drawn from (0 0 0). The investigated direction of the diffracted beam $\mathbf{s}$ can be chosen freely. However, the length of the vector is always $1/\lambda$. If $\mathbf{s}$ is now chosen to

start at M in order to fulfil Equation 2.48, all possible diffraction vectors $\mathbf{S}$ are located on a sphere with the radius $1/\lambda$, the length of the two wave vectors.

All the positions where a point of the reciprocal lattice lies on this sphere are directions where visible diffraction is detectable. By turning the crystal relative to the incident beam the reciprocal lattice can be rotated to different positions in relation to the Ewald sphere, thereby mapping different parts of the reciprocal lattice onto the detector.

A final transformation shall be done at this point. It can be shown by simple geometry, that the magnitude of the diffraction vector has to be:

$$|\mathbf{S}| = \frac{2\sin\theta}{\lambda} \tag{2.49}$$

The magnitude of the reciprocal lattice vector is n/d, where d is the spacing between the lattice planes. With this information Bragg's law can be derived by using the magnitudes and the diffraction condition.

$$|\mathbf{S}| = |\mathbf{r}^*| \tag{2.50}$$
$$2\sin\theta/\lambda = n/d \tag{2.51}$$
$$2d\sin\theta = n\lambda \tag{2.52}$$

2.7. From signal to the absolute squared of the structure factor

So far, all influences on the intensities could be neatly partitioned into atomic contributions. Therefore, the structure factor $F(\mathbf{S})$ calculated from these contributions was all that was needed to evaluate the intensity. However, a more complete equation for the intensity at a specific point in reciprocal space is given by (Giacovazzo *et al.*, 2011):

$$I(\mathbf{S}) = k_1\, k_2\, I_0\, L\, P\, T\, E\, |F(\mathbf{S})|^2 \tag{2.53}$$

This represents the number of effects that are acting on the intensity directly. Here I_0 is the intensity of the incident beam, $k_1 = e^4/(m_e c^4)$ is a factor dependent on the physical constants of electron charge e, the mass of the electron m_e and the speed of light c. The rest of the factors will be tackled in the following sections.

Influence of the illuminated crystal volume

k_2 collects the influence of the crystal size. For a crystal of size Ω with a unit cell size V, which is fully enveloped in the X-ray beam, this is a constant calculated as:

$$k_2 = \lambda^3 \Omega / V^2 \tag{2.54}$$

However, for modern microfocus or MetalJet sources, the assumption of full envelopment is no longer valid. Instead, SADABS is used for correction and the calculated intensity is scaled using redundant data to determine the frame-to-frame scale factor $S(n)$, where n is the frame number of the centre of the reflection within the integration.

SADABS then refines this scaling factor together with an absorption correction (see the later section) by minimising:

$$M = \sum \left[w(\langle I_{\mathbf{S}} \rangle - I_{\mathbf{S}})^2 \right] + \sum \left[e_r^{-2} \left\{ S(n) - S(n+1) \right\}^2 \right] \tag{2.55}$$

As such the intensity of nearby frames is assumed to be similar using a restraint parameter e_r, which can be adapted to the system. It is immediately obvious that for the procedure to work reliably the multiplicity (*i. e.* the number of times a set of symmetry equivalent reflections is measured) needs to be suitably high.

The Lorentz correction

During data collection, crystals were rotated at a constant speed and the intensity on the detector during a specific time frame is summed up. Taking Figure 2.4 into account, a problem can be seen. For a given angular speed of the crystal, the travel speed of the individual reflections is dependent on the distance to the origin of the reciprocal lattice. The further away a reflection is located from (0 0 0), the faster its speed. This means, that the reflections also spend different times in the reflection condition of the Ewald sphere and their recorded intensity has a systematic dependence on the angular frequency. This effect is corrected by the Lorentz correction. In general, the Lorentz correction is dependent on the scattering angle θ by (Giacovazzo *et al.*, 2011):

$$L = \frac{1}{\sin 2\theta} \tag{2.56}$$

Polarisation correction

For a non-polarised beam, the diffracted intensity is affected by an additional factor, which is also resolution dependent it is given by:

$$P = \frac{1 + \cos^2 2\theta}{2} \tag{2.57}$$

Transmission / Absorption

If electromagnetic radiation travels through the crystal, part of it will be absorbed. As such the intensity after absorption I_{abs} is smaller than the diffracted intensity without absorption I_{nonabs}. If linear absorption with the coefficient μ and a travel distance through the crystal r is assumed, the intensity of radiation is modified(Giacovazzo *et al.*, 2011):

$$T = \frac{I_{\text{abs}}}{I_{\text{nonabs}}} = \mathrm{e}^{-\mu r} \tag{2.58}$$

The calculation of μ is usually done by summing up contributions for each atom from tabulated mass attenuation coefficients for a given element and wavelength. It is calculated as:

$$\mu = \rho_m \sum_{j=1}^{N_{\text{element}}} g_j \mu_{j,m} \tag{2.59}$$

where ρ_m is the mass density of the crystal, g_j is the mass fraction of the element with index j and $\mu_{j,m}$ is the mass absorption coefficient.

However, for every wave vector of the incident wave $\mathbf{s}_0$ and exiting wave $\mathbf{s}$, *i. e.* for every scattering vector $\mathbf{S}$, every point within the crystal needs to be taken into account with the incident beam length travelled p and the diffracted beam length q (see Figure 2.5).

As such the transmission can theoretically be calculated as:

$$T(\mathbf{S}) = \frac{1}{\Omega} \int_{\Omega} \mathrm{e}^{-\mu[p(\mathbf{S},\mathbf{R})+q(\mathbf{S},\mathbf{R})]} d\mathbf{R} \tag{2.60}$$

where $\mathbf{R}$ is the location within the crystal in cartesian coordinates and the integration is over the crystal volume Ω.

There are no analytical solutions for this integral for arbitrary crystal shapes. Approximations such as the tetrahedron method proposed by De Meulenaer and Tompa (1965)

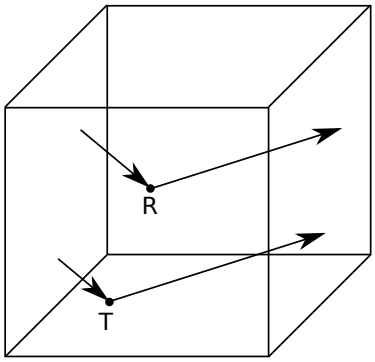


Figure 2.5.: Incident and diffracted beam lengths through the crystal for two arbitrary points P and Q.

are available. This work uses a semi-empirical method implemented in SADABS (Krause *et al.*, 2015). As mentioned in Section 2.7, SADABS uses a refinement against redundant data collection to refine a model for the crystal size. Additionally, it also applies a spherical absorption correction $Q(\mu r, 2\theta)$ and a diffracted beam factor $P(u, v, w)$, which is dependent on the direction cosines of the diffracted beam and is fitted with spherical harmonics. As such the absorption/transmission is treated as:

$$T = P(u, v, q) \cdot Q(\mu r, 2\theta) \tag{2.61}$$

Extinction

The final effect that will be discussed in the context of intensity modification is extinction with the extinction coefficient E. Here two types of extinctions can be distinguished, which both occur with significant crystal size of materials with high scattering cross-section. Extinction mainly affects strong reflections, which often occur at low scattering angles.

The less common effect is primary extinction. It occurs when the single diffracted beam is joined by beams diffracted multiple times. The multiple reflections cause a phase change and as such yield a weaker resulting intensity.

More common is secondary extinction. It occurs when the intensity throughout the crystal is no longer approximately identical. Because of scattering or absorption in the earlier planes of the path of the X-ray beam through the crystal, the intensity is diminished at the scattering planes at the later path of the beam. This reduces the intensity of strong reflections.

In contrast to other effects on the intensity, which are usually treated before refinement, extinction is tackled by the introduction of one or more additional extinction

parameters. A simple approach to describe secondary extinction is by modifying the calculated/modelled intensity by using Giacovazzo *et al.* (2011):

$$(I_{\mathrm{calc}})_{\mathrm{corr}} = \frac{I_{\mathrm{calc}}}{1 + gI_{\mathrm{calc}}} \tag{2.62}$$

where g is the refined extinction parameter. Alternatively, the program SHELXL by Sheldrick (2015*a*) uses an empirical formula which is:

$$(I_{\mathrm{calc}})_{\mathrm{corr}} = \frac{I_{\mathrm{calc}}}{\sqrt{1 + 0.001 I_{\mathrm{calc}} \lambda^3 g / \sin(2\theta)}} \tag{2.63}$$

If possible the best approach is to limit extinction by using small crystals and shorter wavelengths where appropriate.

The treatment of extinction concludes this part on how the mathematical model for the X-ray intensity is constructed.

2.8. Fitting the model to the measured intensity

So far, this chapter has tackled how to model the different effects that determine the intensity of our individual reflections in mathematical terms. Using this model, this section will now deal with how to improve the underlying parameters by minimising the difference between the measured reflection and the modelled intensity. This is certainly not an exhaustive description (as none of the sections in this theory part are) but I hope to provide a feel for the underlying assumptions.

2.8.1. Statistical errors and the method of least squares

Equation 2.53 indicates the intensity in a certain direction given all our assumptions and the intensity of our incoming beam I_0. The equation is written as a continuous number. However, as X-ray radiation consists of photons, which are discrete particles, this cannot be the case. In reality, the given intensity can be used to calculate the probability that a certain number of photons is emitted into the investigated direction. The probability density stems from a Poisson distribution, which has the probability density function (Bronstein *et al.*, 1993):

$$\mathcal{P}(k|\lambda_{\mathrm{P}}) = p(X = k) = \frac{\lambda_{\mathrm{P}}^k \mathrm{e}^{-\lambda_{\mathrm{P}}}}{k!} \tag{2.64}$$

where λ_{P} is the expectation value, *i. e.* the intensity. However, this means that our individual intensity measurement will always have a deviation from this expectation

value, as the variance is also equal to λ_P. Accordingly, for an individual observation, our estimated standard deviation cannot be lower than $\sqrt{\lambda_\mathrm{P}}$. For low expectation values, the Poisson distribution is skewed. However, the normal distribution becomes a good approximation for higher expectation values. This is particularly true as there are additional influences, such as errors in measurement angles, rotation rate, emission rate of the source and detection on the detector. Additionally, there are statistical errors introduced by our data processing, including integration with background correction, absorption correction and frame-to-frame scaling. As such the certainty of our integrated intensities is usually assumed to be normal instead of Poissonian, while the influence of the counting statistics is considered by the integration software.

Whereas the Poisson distribution is discrete, the normal distribution is continuous. Its probability density has the well-known form (Bronstein *et al.*, 1993):

$$\mathcal{N}(x|\mu,\sigma) = p(X=x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.65}$$

with all the derived equations from the previous section and a given vector of parameters $\boldsymbol{\theta}_\mathrm{P}$, it is possible to calculate the intensities of our model $I_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h})$. On the other hand, the observed intensity values $I_\mathrm{obs}(\mathbf{h})$ each with an estimated standard deviation $\sigma_\mathrm{obs}(\mathbf{h})$ have been recorded and corrected.

Assuming that these standard deviations and intensities follow a normal distribution, the probability that a given measured value occurs, given its estimated standard deviation and the calculated value is (Bishop, 2009):

$$\mathcal{N}\left(I_\mathrm{obs}(\mathbf{h})|I_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h}), \sigma_\mathrm{obs}(\mathbf{h})\right) = \frac{1}{\sigma_\mathrm{obs}(\mathbf{h})\sqrt{2\pi}} e^{-\frac{(I_\mathrm{obs}(\mathbf{h}) - I_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h}))^2}{2\sigma^2_\mathrm{obs}(\mathbf{h})}} \tag{2.66}$$

The aim is to find the model, where the collected data has the highest likelihood to appear. First, the joint probability is calculated for N recorded reflections:

$$p(\mathbf{I}_\mathrm{obs}(\mathbf{h})|\mathbf{I}_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h}), \boldsymbol{\sigma}_\mathrm{obs}(\mathbf{h})) = \prod_{i=1}^{N} \mathcal{N}(I_\mathrm{obs}(\mathbf{h}_i)|I_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h}_i), \sigma_\mathrm{obs}(\mathbf{h}_i)) \tag{2.67}$$

The next step is to apply the natural logarithm to this function, which does not change the maximum and yields:

$$\ln p = -\frac{1}{2}\sum_{i=1}^{N} \frac{(I_\mathrm{obs}(\mathbf{h}_i) - I^2_\mathrm{calc}(\boldsymbol{\theta}_\mathrm{P}, \mathbf{h}_i))}{\sigma^2_\mathrm{obs}(\mathbf{h}_i)} + \frac{1}{2}\sum_{i=1}^{N} \ln\left[\sigma_\mathrm{obs}(\mathbf{h}_i)\right] + \frac{N}{2}\ln(2\pi) \tag{2.68}$$

By removing constant terms and factors, it can be seen that maximising this $\ln p$ is equivalent to calculating the sum of least squares, which is the method of choice in crystallographic applications:

$$M = \sum_{i=1}^{N} \frac{(I_{\mathrm{obs}}(\mathbf{h}_i) - I_{\mathrm{calc}}(\boldsymbol{\theta}_{\mathrm{P}}, \mathbf{h}_i))^2}{\sigma^2_{\mathrm{obs}}(\mathbf{h}_i)} \tag{2.69}$$

By varying $\boldsymbol{\theta}_{\mathrm{P}}$, the function M is minimised. Our solution is the set of parameters $\boldsymbol{\theta}_{\mathrm{P}}$, which gives the lowest value M. Usually, this is achieved by using the approximated non-linear least-squares method derived from the closed form for linear least squares. The XHARPy library in Chapter 6 takes a slightly different but related approach by using the BFGS method (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970).

2.8.2. Restraints and Constraints

In the last section, the sum of least squares has been introduced. In the same way that observed intensities have been taken into account, one can also consider a set probability for a target value t to occur given a set probability for the deviation σ_t and the calculated value for the target, which is of course dependent on the parameter vector $t_{\mathrm{calc}}(\boldsymbol{\theta}_{\mathrm{P}})$. The probability is again assumed to be normally distributed:

$$\mathcal{N}\left(t | t_{\mathrm{calc}}(\boldsymbol{\theta}_{\mathrm{P}}), \sigma_t\right) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(t - t_{\mathrm{calc}}(\boldsymbol{\theta}_{\mathrm{P}}))^2}{2\sigma_t^2}} \tag{2.70}$$

These additional assumptions are called restraints in crystallography. Reasonable restraints can stabilise a model and assuming that bond lengths within a group are similar to a disordered second orientation of the same chemical group is one of the most ubiquitous reasonable restraints. Other restraints include similar atomic displacement parameters of atoms located at close distances, atoms following a rigid body motion, certain bond lengths being refined to a certain value or a group of atoms being restrained to a common plane.

The probabilities of all restraints are multiplied to the joint probability to be maximised (Equation 2.66). By again taking the natural logarithm and excluding constant terms, a new function to minimise with N points of data and $N_{\mathrm{restraints}}$ restraints can be obtained:

$$M = \sum_{i=1}^{N} \frac{(I_{\mathrm{obs}}(\mathbf{h}_i) - I_{\mathrm{calc}}(\boldsymbol{\theta}_{\mathrm{P}}, \mathbf{h}_i))^2}{\sigma^2_{\mathrm{obs}}(\mathbf{h}_i)} + \sum_{j=1}^{N_{\mathrm{restraints}}} \frac{(t_j - t^2_{\mathrm{calc},j}(\boldsymbol{\theta}_{\mathrm{P}}))}{\sigma^2_{t,j}} \tag{2.71}$$

On the other hand, the term constrained is used for two values which could be different parameters but are related by a strict mathematical relationship. Examples would

include the assumption that atomic displacement parameters of two atoms at the same site are identical (not just similar as assumed in a restraint) or that the sum of all occupancies of one site is one. There are multiple ways to reduce the redundant parameters. The XHARPy library for example uses the same underlying parameters at model specification.

2.9. Quality indicators and data visualisation

The previous sections introduced the concepts of how to construct a model for the various influences on our observed X-ray intensity and how to find the set of parameters, which fit our observed X-ray intensities most closely. After having achieved this goal, the quality of the resulting model needs to be assessed. In several cases, different models from the same or different data are to be compared in this thesis. The quality indicators used in subsequent chapters will be explained in this section.

2.9.1. Quality indicators for the precision of X-ray data collection

The first class of quality indicators is independent of the model and describes the precision of a set of collected and corrected reflections. It stands to reason that high precision is worthwhile as it indicates how many reflections need to be collected to find a reasonably accurate value for the mean intensity of that set of reflections. As such, several indicators for precision have been proposed. The most commonly used descriptor is the R_{merge} (Weiss, 2001; Giacovazzo *et al.*, 2011) and is calculated as:

$$R_{\text{merge}} = \frac{\sum_{\mathbf{h}} \sum_{i} \left| I_i(\mathbf{h}) - \overline{I(\mathbf{h})} \right|}{\sum_{\mathbf{h}} \sum_{i} I_i(\mathbf{h})} \tag{2.72}$$

where the vector $\mathbf{h}$ denotes the Miller index of a given unique reflection and the index i runs over all measurements of that individual reflection. However, the R_{merge} suffers from the drawback that a very low redundancy/multiplicity might lead to a lower value, thereby leading to incorrect conclusions. Weiss (2001) has proposed the redundancy independent R-value, $R_{\text{r.i.m}}$, which has been corrected for this drawback using the multiplicity N of the individual reflections. It is defined as:

$$R_{\text{r.i.m.}} = \frac{\sum_{\mathbf{h}} \left[N/(N-1)\right]^{1/2} \sum_{i} \left| I_i(\mathbf{h}) - \overline{I(\mathbf{h})} \right|}{\sum_{\mathbf{h}} \sum_{i} I_i(\mathbf{h})} \tag{2.73}$$

Additionally, he also proposed a term indicating the precision of the averaged intensity $R_{\text{p.i.m}}$ defined as:

$$R_{\text{p.i.m.}} = \frac{\sum_{\mathbf{h}} [1/(N-1)]^{1/2} \sum_{i} \left| I_i(\mathbf{h}) - \overline{I(\mathbf{h})} \right|}{\sum_{\mathbf{h}} \sum_{i} I_i(\mathbf{h})} \tag{2.74}$$

Due to their respective advantages but different focus, $R_{\text{r.i.m.}}$ and $R_{\text{p.i.m.}}$ are used for discussion in the Chapter 5 instead of the more commonly used R_{merge}.

2.9.2. Quality indicators for the agreement of the model to the data

The set of quality indicators discussed in this section is concerned with the agreement between a given model and a given set of experimental data. The measured intensities are corrected for all the influences acting on the intensity (Section 2.7) and divided by the refined overall scaling factor. The resulting corrected measured intensity is denoted with F^2_{obs}. The intensity modelled according to the kinematical model is named F^2_{calc}

If the agreement between the model and the calculated data is to be evaluated, the first option is to use the $wR_2(F^2)$, which is defined as:

$$wR_2(F^2) = \sqrt{\frac{\sum_{k=1}^{N} w_k \left(F^2_{\text{obs},k} - F^2_{\text{calc},k} \right)^2}{\sum_{k=1}^{N} w_k \left(F^2_{\text{obs},k} \right)^2}} \tag{2.75}$$

Comparing the minimisation criterion against the data (Equation 2.69), it is obvious, that this is the square root of a scaled M if the weights are chosen as $1/\sigma^2$ and no extinction is refined.

If a weighting scheme is refined and therefore the weights are no longer the inverse variance of the reflection intensities, the $wR_2(F^2)$ cannot be compared between individual refinements. In this case, the $R(F)$ is usually employed.

$$R(F) = \frac{\sum_{k=1}^{N} ||F_{\text{obs},k}| - |F_{\text{calc},k}||}{\sum_{k=1}^{N} |F_{\text{obs},k}|} \tag{2.76}$$

The goodness of fit (GooF) tests whether the differences between calculated and observed intensities match the values of a normal distribution given the number of reflections N and the number of refined parameters $N_{\text{parameters}}$. It is defined as (Clegg and Harrington, 2021):

$$S = \sqrt{\frac{\sum_{k=1}^{N} w_k \left(F^2_{\mathrm{obs},k} - F^2_{\mathrm{calc},k}\right)^2}{N - N_{\mathrm{parameters}}}} \tag{2.77}$$

where a value close to unity indicates good agreement to the expected deviations, a higher value can indicate an under-parametrisation and a lower value over-parametrisation. However, it is dependent on choosing appropriate weights, *i. e.* on the correct determination of the experimental uncertainties.

While the phase is missing from the observed structure factor, it can be assumed that the phases of the calculated structure factor are close to correct and therefore be used to calculate a difference electron density map:

$$\Delta\rho(\mathbf{x}) = \frac{1}{V} \sum_{k=1}^{N} \left(|F_{\mathrm{obs},k}| - |F_{\mathrm{obs},k}|\right) e^{i\phi_{\mathrm{calc},k}} e^{-2\pi i \mathbf{h}_k^T \mathbf{x}} \tag{2.78}$$

Under the assumption that the phase is close enough, this map now reveals features where the calculated and observed reflections result in a difference in density within our unit cell. Therefore, it provides a real-space measurement of our agreement and its analysis is a useful tool during refinement. The simplest quality indicators are the maximum $\Delta\rho_{\mathrm{max}}$ and the minimum $\Delta\rho_{\mathrm{min}}$ values of this difference electron density map.

Additionally, the absolute value of the difference electron density is also a quality indicator. Accordingly, Meindl and Henn (2008) defined the e_{gross}:

$$e_{\mathrm{gross}} = \frac{1}{2} \int_V |\Delta\rho(\mathbf{x})| \, d^3r \tag{2.79}$$

2.9.3. Quality indicators for the agreement of X-ray results to results from neutron diffraction

Secondly, the quality of the derived parameters is to be evaluated. One of the most established applications of HAR is the determination of atomic parameters of hydrogen. X–H distances can be directly compared to the values from neutron refinement in the form of the difference Δr. Obviously, the deviation should be as small as possible. It is desirable to have a lower value which indicates a performance improvement. Therefore the absolute value of the difference $|\Delta r|$ is also employed heavily in the investigations in Chapter 6. The two indicators are calculated as:

$$\Delta r = r_{\mathrm{X}} - r_{\mathrm{n}} \tag{2.80}$$
$$|\Delta r| = |r_{\mathrm{X}} - r_{\mathrm{n}}| \tag{2.81}$$

The comparison of atomic displacement parameters is less straightforward. They can be affected by influences such as absorption or extinction as well as small deviations in temperature. This problem has already been tackled by Blessing (1995). To compare displacement parameters from different models the values from neutron diffraction were scaled using the formula:

$$U^{ij}_{\mathrm{c,n}} = q_{\mathrm{ADP}} \cdot U^{ij}_{\mathrm{c,n}} + \Delta U^{ij} \tag{2.82}$$

Here q_{ADP} represents a scaling due to differences in measurement temperature. ΔU^{ij} can correct for differences in extinction, but some of the individual values can be zero, depending on the crystal system. Both are parameters to be determined by a least-squares refinement to the non-hydrogen atoms. This was done using a custom implementation in PYTHON.

Using the scaled atomic displacement parameters, the difference between the refinements from neutron and X-ray HAR can now be evaluated as a simple difference ΔU_{ij}. Additionally, the absolute of this difference $|\Delta U_{ij}|$ will be used. As such it can be defined:

$$\Delta U_{ij} = U^{ij}_{\mathrm{X}} - U^{ij}_{\mathrm{n,c}} \tag{2.83}$$
$$|\Delta U_{ij}| = \left| U^{ij}_{\mathrm{X}} - U^{ij}_{\mathrm{n,c}} \right| \tag{2.84}$$

Finally, a measure for the overlap of the two probability functions, which are represented by the atomic displacement parameters of the hydrogen atoms, needs to be calculated. Whitten and Spackman (2006) proposed the S_{12} value for this purpose. It is defined as:

$$S_{12} = 1 - \int \left[p_1(\mathbf{x}) p_2(\mathbf{x})\right]^{\frac{1}{2}} d^3\mathbf{x} = 1 - \frac{2^{\frac{3}{2}} \left(\det \mathbf{U}_1^{-1}\mathbf{U}_2^{-1}\right)^{\frac{1}{4}}}{\left[\det\left(\mathbf{U}_1^{-1} + \mathbf{U}_2^{-1}\right)\right]^{\frac{1}{2}}} \tag{2.85}$$

Usually, an additional factor of 100 is contained within the expression. All our values are listed in percent. Subsequently, the numerical values will be the same.

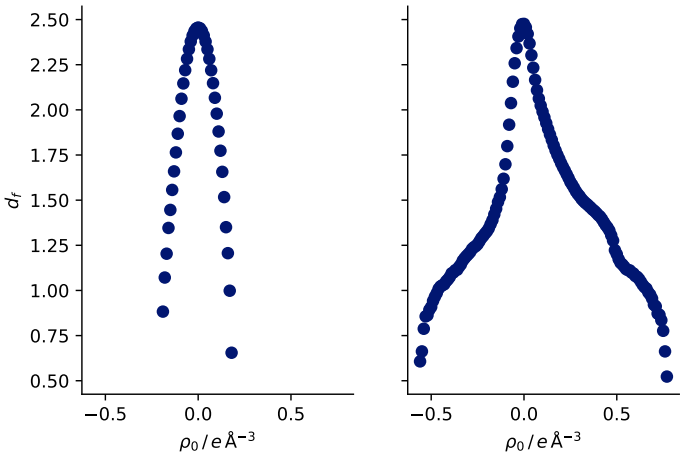


Figure 2.6.: Two examples of Henn-Meindl plots. While the left shows the correct parabola for a normally distributed difference electron density, the right figure shows significant systematic deviations for both positive and negative residual density.

2.9.4. Selected data visualisations

DRKPlot: The DRKPlot (Stash, 2007) program offers several different plots that compare modelled and refined intensities. The one often called DRKPlot is created by binning observed and modelled intensities within set resolution ranges and calculating the quotient $\sum F^2_{\text{obs}} / \sum F^2_{\text{calc}}$.

For a multipole refinement, the usually accepted range is between 1.05 and 0.95 for all bins.

Henn-Meindl-plot: This analytic plot was proposed by Meindl and Henn (2008) to test for the normalcy of the difference electron density of a refined structure. It is less useful for independent atom model refinements, where the bonding density represents a non-statistical contribution to the difference electron density, but is very useful for refinements involving aspheric atom descriptions such as the multipole or the Hirshfeld atom model.

Two examples are depicted in Figure 2.6. The expected behaviour is a parabola. Additionally, the distribution should be as narrow as possible as this indicates the difference electron density is flat and featureless.

Box-Whisker-Plot: In contrast to the first two plots, box-whisker plots are not well established in crystallography but come from the field of statistical visualisation

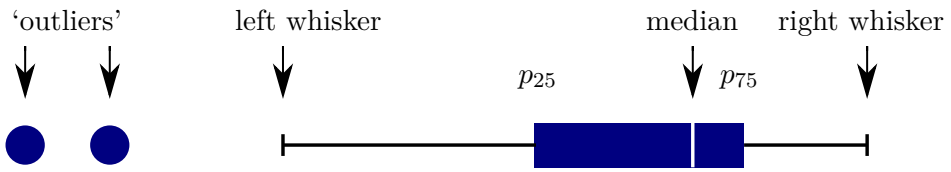


Figure 2.7.: The basic setup of a box-whisker plot. p_{25} and p_{75} denote the 25^{th} and 75^{th} percentiles respectively. The left whisker is the lowest point between p_{25} and $p_{25} - 1.5 \cdot (p_{75} - p_{25})$. The right whisker is the highest point between p_{75} and $p_{75} + 1.5 \cdot (p_{75} - p_{25})$. Values outside this range are denoted separately ('outliers').

(Krzywinski and Altman, 2014). They give more information on the distribution of a given set of data points compared to the usual combination of mean and standard deviation. As some of the differences discussed in Chapter 6 do not seem to follow the normal distribution, they are extensively used there.

The general explanation of how these plots are constructed using an example of a significantly skewed distribution is given in Figure 2.7.

3. Evaluation of structural properties of anthracene derivatives and correlation to their luminescence

This chapter will contribute to the discussion of substituted anthracene structures, which were investigated in collaboration with other members of the Stalke group. Here, intermolecular interactions in the solid state have long been stated as potential candidates for the rationalisation of the presence of the luminescence of one conformer over the other. This chapter will investigate these ideas in some detail. For the second class of compounds, different co-crystals of one molecule exhibited different fluorescence behaviour. This chapter will describe approaches to reproduce this behaviour using methods from computational chemistry and try to rationalise why they were ultimately not successful.

The general structure can be found in Figure 3.1. A useful shorthand naming has been introduced within the group for the two organic groups on the substituent. As the atom directly connected to the anthracene has two phenyl groups by definition, these are omitted from the name. If an oxidant is present this is named first, followed by the atom connected to the anthracene. This is followed by 'An' for the anthracene itself. The last part can either consist of an organic substituent (*e. g.* Me for methyl) or if the substitution has been done with another substituent as depicted in the centre of Figure 3.1, the shorthand notation is used in reversed order. For example, $Ph_2(S{=})P{-}An{-}P({=}S)Ph_2$ is shortened to SPAnPS.

The investigation of this class of structures has a long history within the Stalke group. The initial investigation was published by Fei *et al.* (2003) and included disubstituted anthracene derivatives, where the anthracene held two $Ph_2(S{=})P{-}$) substituents in 9 and 10 positions (SPAnPS). The compound was found to be non-emissive in solution. However, the solid-state structure, which included co-crystallised toluene molecules, showed a strong yellow-green emission upon radiation with UV light. Fei *et al.* reported,

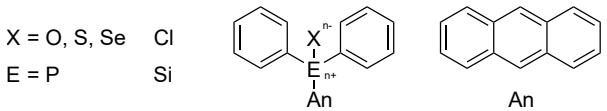


Figure 3.1.: Compounds included in the discussion. If E is a phosphorous atom, n is equal to one. If E is a silicon atom $n = 0$.

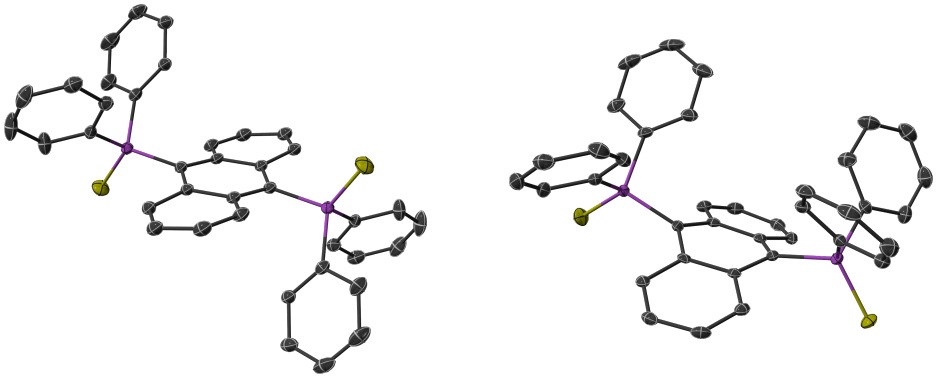


Figure 3.2.: Conformers of $Ph_2(S=)P-An-P(=S)Ph_2$ (SPAnPS). Left: *transoid* conformation. Right: *cisoid* conformation. Hydrogen atoms have been omitted for clarity. Atomic displacement parameters are displayed at 50 % probability.

that this emission could be quenched by applying vacuum to the powdered crystal, thereby removing the toluene molecule from the crystal structure. The fluorescence could not be recovered using benzene, prompting the conclusion that the fluorescence was specific for the presence of toluene. Therefore, Fei *et al.* concluded that the combined structures were responsible for the luminescence properties, making it an exciplex.

This conclusion was challenged by the research of Schwab (2004, 2008) who was able to characterise a number of additional fluorescent solid-state structures of SPAnPS. This included the molecule with other aromatic solvents exhibiting strong luminescence, as well as structures with negligible fluorescence at room temperature incorporating aromatic and non-aromatic solvent molecules. The main difference Schwab observed was the conformation that the SPAnPS molecule adopts within the solid-state structures. While the strongly fluorescent structures all included the molecule in a *transoid* conformation, where the sulphur atoms are located at opposite sites of the anthracene moiety, the structures with negligible fluorescence all included the molecule in a *cisoid* conformation, where the sulphur is located at the same side (Figure 3.2). The additional *transoid* structures also included a luminescent structure with benzene as the guest molecule. However, the exhibited space group was $P\bar{1}$ instead of $P\,2_1/n$ as observed in the toluene structure as well as the majority of *transoid* conformations in the investigated crystals.

Additionally, Schwab also investigated the luminescence properties at a reduced temperature of -135 °C. At this temperature *cisoid* as well as *transoid* solid-state structures did exhibit fluorescence, including the *transoid* structure with removed solvent. As such, the quenching of the emission in some compounds could be attributed to a temperature-dependent process, such as a non-radiative decay *via* inter-system crossing into a non-radiative triplet state. Additionally, Schwab also modified the substituent at the phosphorous group by exchanging the phenyl group for cyclohexyl, o-tolyl and

isopropyl groups. None of these modified compounds showed a non-negligible solid-state fluorescence.

Further characterisations and analysis of the phenyl systems were supplied by Finkelmeier (2013). He pinpointed the emission maxima of all solid-state compounds including the faintly fluorescent species to a range of 515 – 530 nm. He was also able to obtain the solvent-free structure of SPAnPS, which exhibits a *cisoid* confirmation. Building upon initial observations by Schwab, that the *cisoid* conformers show strong deformation he was able to develop quantifiers for the deformations observed in this structure class (see twist and bend angle in Section 3.2). Finkelmeier attributed the quenching to the bending of the anthracene molecule and postulated an enhancement of the fluorescence with the formation of the twist angle, but noted, that both of these features also corresponded to the two different conformers.

Finkelmeier also extended the exciplex theory of Fei by investigating π–H distances of guest molecules to the anthracene moiety for the *transoid* structures, but could not find an overall influence. This approach was later used for further investigations by Bukala. However, further systematic results could be obtained as a procedure for reproducible quantum efficiencies could not be reached, thereby prohibiting detailed discussions of quantum efficiencies. Subsequently, the question to rationalise strong or negligible luminescence at room temperature for 9,10-SPAnPS remained open.

Krause (2017) investigated with additional experimental (powder X-ray diffraction, neutron diffraction, solid-state NMR spectroscopy) and theoretical methods and added further structures and structure conversions to the available pool of information. He revisited the original experiment of Fei *et al.* and subjected vacuum-dried powder of the toluene species to benzene. There, he could reaffirm the observation, that no immediate recovery could be found. However, after a period of about 60 minutes, small fluorescent crystals could be found in the flask, which upon inspection turned out to be the known benzene *transoid* structure in $P\bar{1}$. Further investigations included the application of toluene vapour to dried crystals of the benzene structure. Here a new polymorph containing a fluorescent *transoid* structure with toluene in $P\bar{1}$ was formed. Krause concluded that the dried co-crystals can act as a matrix and this effect can predefine the structure.

For rationalising the difference between the two polymorphs, Krause turned to intermolecular interactions. He noted that the *cisoid* structures exhibited larger vibrational parameters in the X-ray structure refinement, as well as a larger probability of disorder. He explained that this hints towards a restriction of intramolecular vibration as introduced by Li *et al.* (2016). Thereby he concluded that the effect of the guest molecule might not be in the effect on the electronic structure but in the rigidity of the overall system. Therefore, he investigated energy network interaction energies, which however did not yield a conclusive answer. Krause reaffirmed the problem of comparing two very similar classes (*cisoid* and *transoid* SPAnPS) with each other, where the differences are present between the two structure motifs, but the generalisability of explanations cannot be founded on these two points alone.

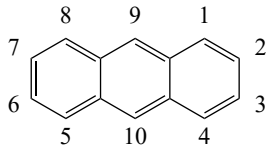


Figure 3.3.: Official numbering for the enumeration of possible substituent positions on the anthracene molecule.

This drawback was partly tackled by Stern (2006, 2009) with the synthesis of the asymmetric compounds $Ph_2(S=)P-An-H$ (SPAnH) and $Ph_2(S=)P-An-Me$ (SPAnMe). The scope of available structures was furthered significantly in the PhD thesis of Schillmöller (2022), who not only supplied additional 9,10 substituted compounds, namely $Ph_2(S=)P-An-Br$ (SPAnBr), but also $Ph_2(S=)P-An-Et$ (SPAnEt), $Ph_2(S=)P-An-SiMe_3$ (SPAnSiMe) and $Ph_2(S=)P-An-Ph$ (SPAnPh). In addition, he also provided synthetic routes to compounds with single and double substitutions at other positions of the anthracene molecule. As such, he enabled a more global analysis beyond the so far dominant SPAnPS compound. The tools described in this chapter were developed in cooperation with him and this cooperation resulted in two publications (Schillmöller *et al.*, 2020*a*) and (Schillmöller *et al.*, 2020*b*).

With respect to the majority of the investigated SPAnR systems, Schillmöller argued convincingly that the absence of a large bathochromic shift for distinct 'host-guest' complexes of the same compound with different solvent molecules makes the presence of an exciplex unlikely. As such he stressed the importance of the substitution pattern for the wavelength shift and agreed with Krause on the importance of rigidity in the system for the quantum yield, in accordance with the principles of the restriction of internal motion (Mei *et al.*, 2014) or the restricted access to conical intersections (Peng *et al.*, 2016; Crespo-Otero *et al.*, 2019).

3.1. The general approach for analysis

To facilitate the fast analysis under the inclusion of new compounds supplied by synthetic chemists, the analysis was automated in the form of a PYTHON script. The required information can therefore be limited to a crystallographic information file 'cif' and the information on how to reconstruct the anthracene moiety from the asymmetric unit. The algorithm for the analysis works using the following steps.

Reading the 'cif' file: The crystallographic information file contains the necessary information about symmetry elements, atomic positions and displacement parameters. In addition, it contains the estimated standard deviations of the refined parameters, which are necessary for error propagation. The first step is to extract this information for further use in PYTHON.

Reconstructing the anthracene: For some structures, the anthracene moiety is generated by symmetry, as only half of the molecule or molecules is or are located within the asymmetric unit. Therefore, atomic positions and displacement parameters need to be projected to their respective equivalent positions with the supplied symmetry element or elements.

Finding the anthracene: First, the molecular graph needs to be constructed. Distances between atomic positions are calculated. Two atoms count as bound to each other if they are closer than 1.2 times the sum of their covalent radii. This enables the construction of a molecular graph in the PYTHON package NETWORKX, which can be analysed for the necessary features. The algorithm searches for sp^2 atoms. In a graph sense, these are carbon atoms, which only have three connecting atoms, as radical or ionic species are not contained within the dataset under investigation. Subsequently, aromatic rings are identified as six-membered rings of sp^2 carbon atoms. This enables the search for anthracene moieties which are sets of three connected aromatic rings, where connecting atoms have only one other connecting atom as a direct neighbour, thus enforcing that the arrangement of the three aromatic rings has to be linear.

Enumerating the positions and connected atoms: The correct numbering for the anthracene molecule is displayed in Figure 3.3. However, due to implementation reasons, the internal numbering is different. This enables the use of pure integer numbers for all carbon atoms within the anthracene moiety instead of using the combination of numbering and letters the IUPAC uses for the connecting atoms. The context of this work will use the official numbering. However, for further use of the script in future projects, the internal numbering shall be documented here.

First, the central carbon atom with the highest IUPAC priority is identified (in official numbering C^9 or C^{10}). This atom gets the index 1. Secondly, the script picks the atom with the highest IUPAC priority within the set of atoms C^1, C^4, C^5 or C^8. This atom will then be set to the internal index 3. The carbon atom between internal indices 1 and 3 gets index 2. As this sets the start and the direction of the numbering around the ring, the remaining carbon atoms are just indexed consecutively up to index 14.

All substituents get assigned the position where they are connected to the anthracene molecule, as well as a distance in atoms to the anthracene.

Evaluate the features: Now, that everything is enumerated, the features can be evaluated. A detailed description of all features is given in the next section and Section B.1.1. However, at the moment the script assumes a phosphorous or silicon atom connected to the carbon atom at position 1 for calculation of the features. In combination with heavier elements than phosphorous at the opposite position, this can lead to a situation where the calculation of certain properties is skipped. It was checked manually, that no structure in the evaluated sets suffered from this problem.

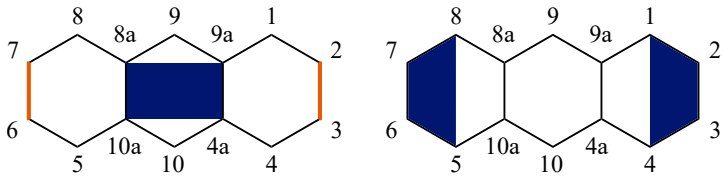


Figure 3.4.: Definition of the twist angle (left) and the bend angle (right). The twist angle is defined as the angle between the bonds of C^6-C^7 or C^2-C^3 (orange) to the mean plane of the atoms C^{4a}, C^{8a}, C^{9a} and C^{10a} (blue). The bend angle is defined as the angle between mean planes through atoms C^1 to C^4 and C^5 to C^8 (both blue).

Automated error propagation: Finally, where possible, the PYTHON package AUTOGRAD is used for automatic error propagation using the estimated standard deviations of the atomic parameters. This is an approximation as a full error propagation would use the covariances as well as the variances for the calculation. However, the covariances are not available from the 'cif' file and therefore were not easily obtainable within the scope of the investigation.

3.2. Investigated features implemented in the automated analysis

As per the discussion between geometric influences and influences from vibration, there are two distinct groups of features which are investigated within this work. Features, which are derived from the atomic positions as determined by X-ray diffractometry and features which are derived from the mean displacements and which are therefore connected to the dynamics of the system. Additional features have been implemented, but as they turned out to not be relevant for the discussion, their description has been moved to the appendix in Section B.1.1

Twist angle: The twist angle is defined as the angle between the mean plane of the four connecting carbon atoms of the central aromatic ring in the anthracene moiety to the bond between the carbon atom pairs C^2, C^3 and C^6, C^7 (see Figure 3.4)

Bend angle: Finkelmeier developed a straightforward approach to calculating the bend angle. Mean planes are calculated through the outer four carbon atoms on both sides of the anthracene molecule (group 1: C^1–C^4, group 2: C^5–C^8, see Figure 3.4). The intersecting angle between these two carbon atom planes is the bend angle of the system.

Overlap and relative anthracene positions: A second anthracene molecule in direct vicinity can significantly influence the emission wavelength of the crystalline material. In order to be able to quantify such influences, a systematic quantification of the overlap, as well as the relative positions of two anthracene molecules in a dimer was necessary.

As such, the calculation of the overlap, as well as the calculation of the relative anthracene position was implemented in the following way:

1. Calculate a mean plane through the carbon atoms of the investigated anthracene molecule.

2. Project all carbon atoms of the anthracene onto this plane.

3. Calculate the rotation matrix, which orients the normal vector of the plane in z direction and the vector from the mean carbon position to C^9 in x direction. This has the consequence that the long side of the anthracene molecule will be oriented in y direction.

4. Calculate the relative position of additional anthracene molecules given the rotation matrix and the symmetry of the unit cell.

5. Find the closest anthracene molecule, which has a mean plane with an interplanar angle to the first plane which is lower than 45°.

6. Calculate the relative position of the centroids of the two anthracene molecules in the rotated coordinate system.

7. Project the carbon atoms of the second anthracene onto the mean plane of the first anthracene, which, due to rotation, just means setting z equal to 0.

8. Construct two-dimensional polygons for the two anthracene molecules.

9. Calculate the area of the intersection polygon and divide it by the area of the first anthracene. This value is the overlap in %.

This provides a systematic way to determine the overlap.

Mean U(equiv) values The displacement parameters should give an idea of the rigidity of the system. (For the details see Section 2.4). In practice, atomic displacement parameters can fit a variety of other effects and therefore the determined values can be skewed by inadequacies in the data treatment, but more relevantly can be also skewed by the crystal quality. With correct data treatment and comparable crystal quality, the mean U_{equiv} value (Equation 2.36) of a part of the molecule can be used as a measure of the flexibility that group enjoys within the crystalline environment.

3.3. Challenging explanations for the 9,10-SPAnR luminescence

As written in the introduction to this chapter, several explanations for the difference in the room temperature luminescence have been proposed during the long period of time these systems have been investigated in the Stalke group. However, almost all of these possibilities have been proposed to explain the behaviour of structures formed by 9,10-SPAnPS with various solvent molecules or without the co-crystallisation of a solvent. The following sections should deal with the individual explanations one by one and add to the discussion with new experiments and data analysis.

3.3.1. Testing the localisation of C−H···π interactions

Finkelmeier (2013) discussed C−H···π interactions as a possible influence on the room temperature luminescence properties. To this end, he determined the distances of hydrogen atoms placed in the riding model to the anthracene plane and investigated a possible connection. In addition to the fact, that hydrogen atoms placed in a riding model provide at best a coarse approximation of position, Schwab (2004) already demonstrated a disconnect between the discussed X-ray structure, which was determined at 100 K and the room temperature measurement of fluorescence intensities.

To bridge this gap, X-ray diffraction data was collected at different temperatures. The corresponding structures were refined using NoSpherA2 (Kleemiss *et al.*, 2021) in Olex2 (Dolomanov *et al.*, 2009), as the XHARPy package discussed in Chapter 6 cannot refine disordered structures. As the size of the atomic displacement parameters increased significantly with rising temperature and libration effects were to be avoided, all data was refined against atomic form factors determined by HAR of the lowest temperature structure (100 K). The DFT calculation used no approximation of the crystal environment and the partitioned molecular density was calculated on the R2SCAN//def2-TZVP level.

Refinement of the room temperature data revealed a disorder of the toluene within the structure, which could subsequently be located in the 100 K structure with a significantly lower occupation. HAR ensured that the bonding density of the major fraction does not introduce a systematic overestimation of the occupation of the less probable fraction. This is a possibility with an independent atom model refinement, where the undescribed aspheric bonding density of the major position might skew the occupancy and position of the minor fraction. The resulting structures for 100 K and 298 K are depicted in Figure 3.5.

The resulting occupancies are listed in Table 3.1. The question is, whether there is a dynamic or static disorder between the two positions. During the experiment, an equilibration time of 5 minutes was used after each temperature change. From the simple fact that the occupancy is temperature dependent, it is clear that the barrier between two potential states can be crossed at 125 K. Alternatively, the pattern could also be the

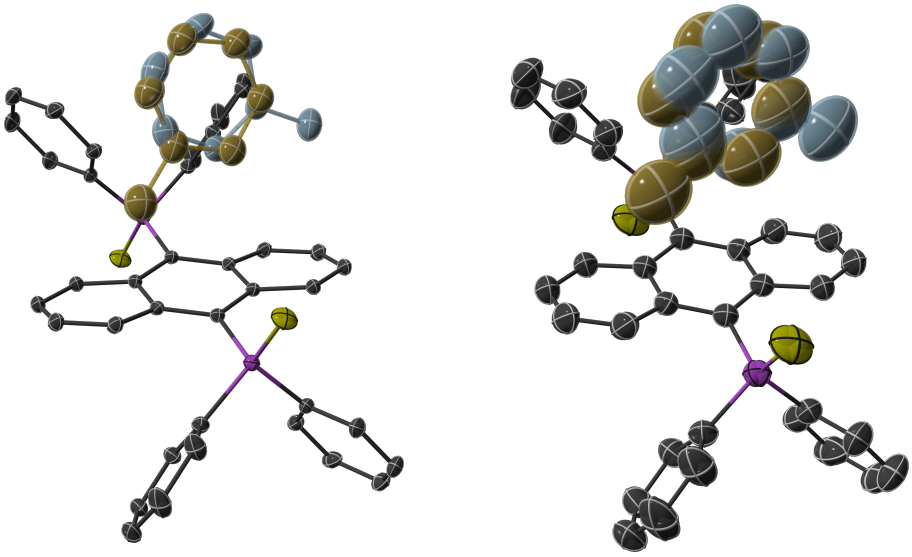


Figure 3.5.: Depiction of disorder (major: grey, minor: orange) in the structure of SPAnPS/Toluene for the diffraction experiment at 100 K (left) and 298 K (right). Atomic displacement parameters are depicted at 50 % probability. Hydrogen atoms are omitted for clarity.

result of a vibrational mode, even though the necessary amplitude of the displacement seems rather pronounced for such a scenario.

Further evidence for the presence of two distinct positions instead of a simple artefact can be found in the development of the unit cell length a with measurement temperature. As depicted in Figure 3.6, the unit cell length decreases between a temperature of 150 K and 250 K. This can be explained by the relative positions of the methyl groups of the disordered toluene molecule in the respective fractions. The major disorder position has the methyl group oriented towards the a-axis, while the minor disorder fraction has the methyl group oriented towards the b-axis. As the occupation of the major disorder is

Table 3.1.: Determined lattice constants a and major occupancies n_{major} for the structures of SPAnPS at different temperatures. Occupancies were obtained with HAR.

T / K	n_{major} / %	a / Å
100(2)	96.85(17)	10.766(2)
125(2)	94.6(2)	10.775(2)
150(2)	91.7(2)	10.777(2)
200(2)	88.9(2)	10.759(2)
250(2)	81.4(5)	10.744(2)
298(2)	78.6(6)	10.752(2)

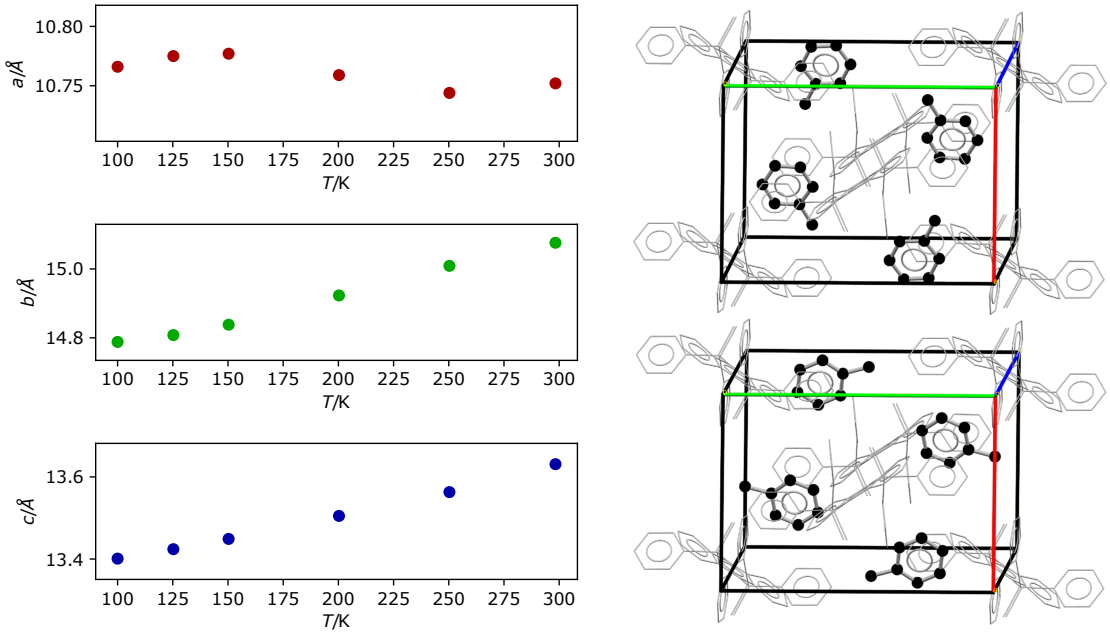


Figure 3.6.: Dependency on the unit cell lengths on the temperature of the diffraction experiment (left). Orientation of the major toluene disorder position (upper right) and the minor toluene disorder position (lower right) to the unit cell constants (a-direction: red, b-direction: green, c-direction: blue).

lowered with increased temperature, fewer methyl groups are oriented in a-direction, thereby reducing the average needed space of the toluene molecule in that direction.

If a static disorder between two distinct minima is assumed, the relative energy of the two states can be calculated by assuming a two-state Boltzmann distribution, where the occupancy of the major disorder fraction would be given by:

$$n_{\text{major}} = \frac{\exp\left(\frac{\Delta E}{RT}\right)}{1 + \exp\left(\frac{\Delta E}{RT}\right)} \tag{3.1}$$

The practical regression was performed using PyMC which is the newest version of the PyMC3 library by Salvatier *et al.* (2016) using a probabilistic model determined using the No U-Turn Sampler by Hoffman *et al.* (2014). The posterior predictive for the energy is given in Figure 3.7. The determined energy is $3.16(3)\,\mathrm{kJ\,mol^{-1}}$.

This determination is under the assumption that the system is well described with two positions. In the datasets of 250 K and 298 K difference electron density can be observed in the region occupied by the toluene molecule (see Section B.2). As such, a third position might be possible. However, due to the needed increase in the number

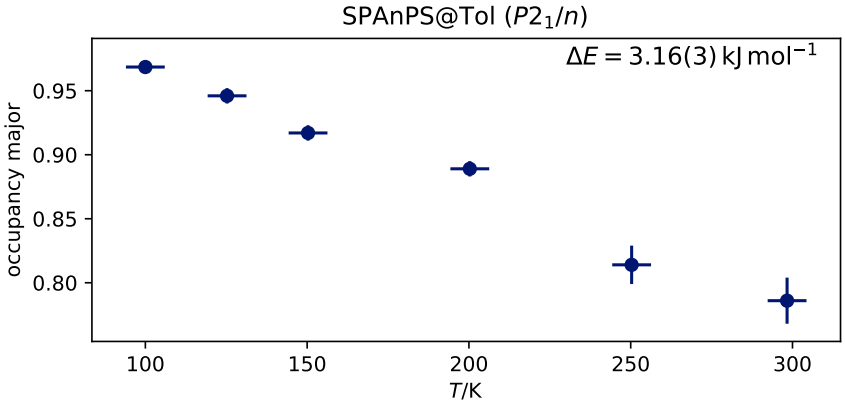


Figure 3.7.: Refined occupancies in dependence of the temperature and 300 draws from the posterior predictive distribution for the evaluated model.

of parameters, this was neglected. This also explains the increasingly large confidence interval for the occupancy parameter in the high-temperature measurements.

In comparison to plane wave DFT-calculations in Quantum Espresso (Giannozzi *et al.*, 2017) conducted to reproduce this energy difference, it can be observed that the difference from experiment is lower than calculated energy differences, especially if dispersion is included with one of the many possible approximations for van-der-Waals interactions (see Figure 3.8). The calculated energies were obtained using the PBE functional with various dispersion corrections by Grimme (Grimme, 2006; Grimme *et al.*, 2010, 2011; Smith *et al.*, 2016), as well as three functionals which included dispersion interaction, namely rvv10 (Sabatini *et al.*, 2013), vdW-DF (Dion *et al.*, 2004; Thonhauser *et al.*, 2015) and vdW-DF2 (Lee *et al.*, 2010). All calculations used Ultrasoft pseudopotentials with a kinetic energy cutoff of 80 Ry and a charge density cutoff of 800 Ry respectively. Occupancy of the respective disorder positions was either completely assigned to the major or the minor position for the energy calculations of the two cells. Cell constants were allowed to relax in accordance with the monoclinic symmetry of the unit cell.

The calculated energy differences are systematically too large. As such, the PBE calculation without any dispersion gave the closest agreement. As the physical system should include van-der-Waals interactions, it remains a question where the source of the systematic difference is located. Possible candidates include the fact that the calculated unit cells only contained either conformation but not the mixture of the two or the complete neglect of entropic contributions.

However, there is significant flexibility in the toluene orientation at room temperature and this basically excludes the possibility of directed and strong interactions of the guest molecule to the anthracene group. This makes an influence of these interactions on the room-temperature luminescence unlikely. Even if the change in occupancy is a change of one minimum into another, the interconversion already happens between 100 and 125 K,

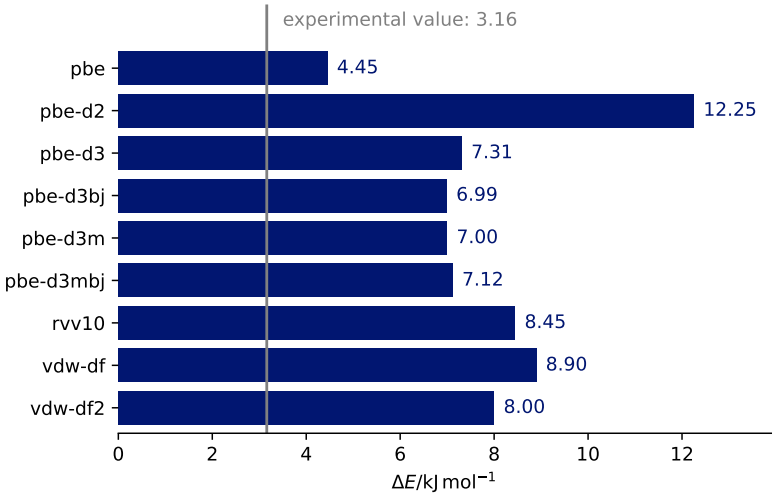


Figure 3.8.: Calculated energy difference for the two orientations of toluene in the unit cell for different schemes for including dispersion. Part of the calculations was conducted with the PBE functional and dispersion correction by D2 (Grimme, 2006), D3 (Grimme *et al.*, 2010), D3BJ (Grimme *et al.*, 2011), D3M and D3MBJ (Smith *et al.*, 2016) schemes. The remaining calculations used the van-der-Waals functionals rvv10 (Sabatini *et al.*, 2013), vdw-df (Dion *et al.*, 2004; Thonhauser *et al.*, 2015) and vdw-df2 (Lee *et al.*, 2010).

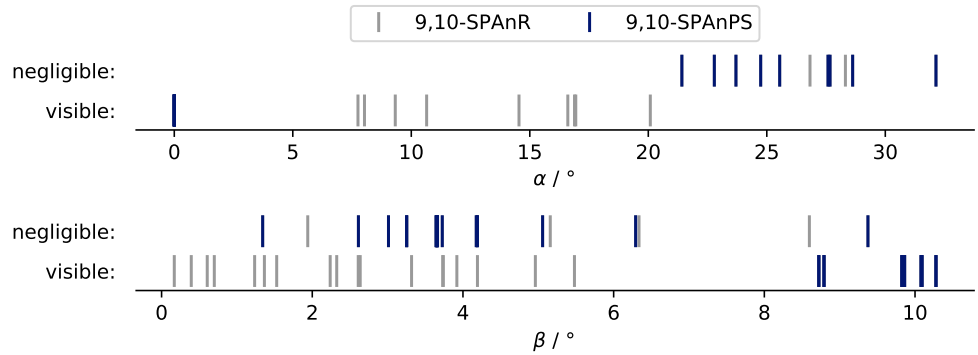


Figure 3.9.: Bend angle α and twist angle β for 9,10-SPAnR structures. Angles determined from 9,10-SPAnPS are marked in blue. Rows in the individual plots represent negligible fluorescence or visible fluorescence at room temperature.

which means a rather low barrier. In the more unlikely case that the observed disorder is purely dynamic, the amplitude of the corresponding phonon mode is exceptionally large and therefore a directed interaction is even less plausible.

3.3.2. Twist and Bend in global investigations

Using the additional structures provided by Schillmöller (2022), as well as the automatic evaluation, the geometric criteria by Finkelmeier (2013) can be investigated in more detail.

The determined bend and twist angles for 9,10-SPAnR structures are depicted in Figure 3.9. The molecules have been grouped by either negligible luminescence at room temperature or visible luminescence at room temperature to avoid any ambiguity connected to the determination of quantum efficiencies. The two angles seem to hold very differently with the additional data.

For the bend angle α, the distance in the separation seen in the 9,10-SPAnPS data alone is not retained. However, there still is a border at about 20.5°, where structures with a lower bend angle show visible luminescence and structures with a higher bend angle show negligible luminescence within the investigated group of structures. Additional data or structures might provide a more gradual change, but the general connection between a large bend angle and a quenched luminescence is valid.

At the same time, the twist angle does not seem to be connected to the presence of luminescence, where the additional structures include low twist angle structures with visible luminescence, as well as high twist angle structures with negligible luminescence. The solvent-free 9,10-SPAnPS does also yield a large twist angle by the definition of Finkelmeier, while structures with visible fluorescence and low twist angles only appear with the additional structures within the SPAnR set.

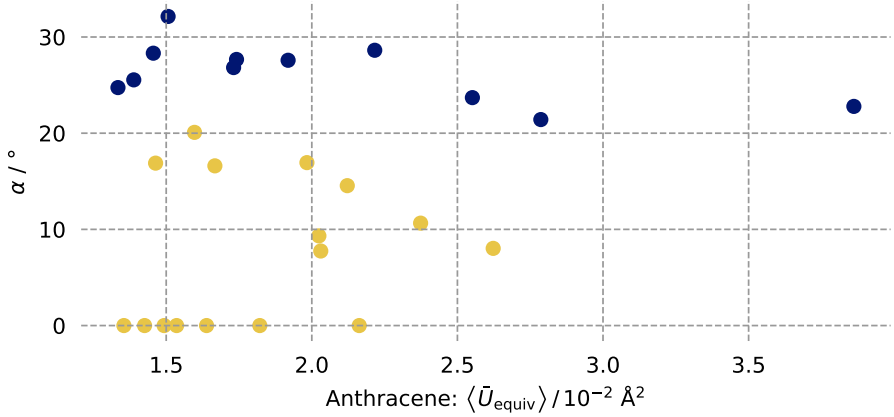


Figure 3.10.: Bend angle α and mean equivalent displacement parameter of the anthracene carbon atoms $\left\langle \bar{U}_{\text{equiv}} \right\rangle$ for structures exhibiting negligible (blue) or visible (yellow) fluorescence.

As such the additional data has sharpened the look upon the discussed structures. It is clearly visible, that the 9,10-SPAnPS samples discussed in isolation before the works of Schillmöller are not a representative sample of the compound class with respect to the twist angle.

3.3.3. Completing the picture with displacement

Krause and Schillmöller both favoured an explanation including the restricted access to conical intersections (Crespo-Otero *et al.*, 2019). This approach tries to explain, why some compounds show fluorescence in the solid state, but not in solution. The theory states that there is a radiationless relaxation pathway including a conical intersection between two states, which can be reached by the relaxation of the molecular structure upon optical excitation. In the solid state, the relaxation pathway is blocked due to the reduced degree of free movement the molecule can exhibit. Therefore, the conical intersection is no longer accessible and no radiationless decay can happen. This means that a decay involving radiation occurs instead.

This approach can be transferred onto the different polymorphs and solvent structures of 9,10-SPAnR by assuming that the radiationless conical intersection has a geometry exhibiting a bend angle. The amount of reorientation needed to reach the conical intersection is therefore much smaller in the compounds exhibiting negligible fluorescence, while the amount of reorientation needed in the absence of a preformed bend angle is higher.

A second influence can be in the flexibility of the overall system. Even if the bend angle is not quite near the ideal value, given a large degree of freedom for the anthracene molecule to reorient itself, the barrier to the conical intersection might still be low enough for a radiationless decay. As such, the dependency on both the mean displacement parameter of the anthracene moiety and the bend angle has been investigated (Figure 3.10). In agreement with the presented theory, the three structures closest to the dividing line in the bend angle, located at about $\alpha = 20.5°$ all show larger displacement parameters, hinting that they might reach the intersection by relaxation of the positions in the excited state.

3.4. Rationalising the wavelength shift in 1-SPAnH structures

As written in the introduction to this chapter, Schillmöller (2022) was able to provide synthetic routes to different regioisomers. This included the synthesis of 1-Ph_2(S)P−An (1-SPAnH). In the solid state, this compound can adopt several different structures, of which most contain a dimer of the molecule. In dependence on the presence or absence of co-crystallised solvent molecules, the dimer structures adopt different overlaps, which can be correlated to the emission wavelength. In total there is one solvent-free structure without any dimer formation, two solvent-free polymorphs which form dimers (free1, free2) and four structures which include a solvent molecule and dimers in the solid-state structure (quinoline, benzene, aniline and pyridine).

The contribution of my work in this project consisted of two parts. The first task was to find a systematic way to calculate the overlap of the dimer structures. The developed approach is described in Section 3.2.

As depicted in Figure 3.11, the overlap between dimer structures can be correlated with the wavelength shift for these structures. As the overlap is correlated with the relative position of the two anthracene molecules a correlation of the luminescence with Δx and Δy can be observed as well. The distance between the anthracenes (Δz) is not correlated with the shift in wavelength. The second step within this section is to try to reproduce this behaviour with the Gaussian program (Frisch *et al.*, 2016).

3.4.1. Anthracene dimers without substituents

The investigation started with the evaluation of anthracene dimers without any substitution. This assumes, that the influence of the substituent of the anthracene on the relative energies of the fluorescence is small. The overlap of the anthracene molecules themselves would then be the central influence. To test this hypothesis, dimer structures of anthracene were generated using the geometries from the crystal structure. Before the calculation of the excited state, the substituent was replaced by a hydrogen atom and all hydrogen atoms were relaxed using either the CAM-B3LYP or M062X functionals and

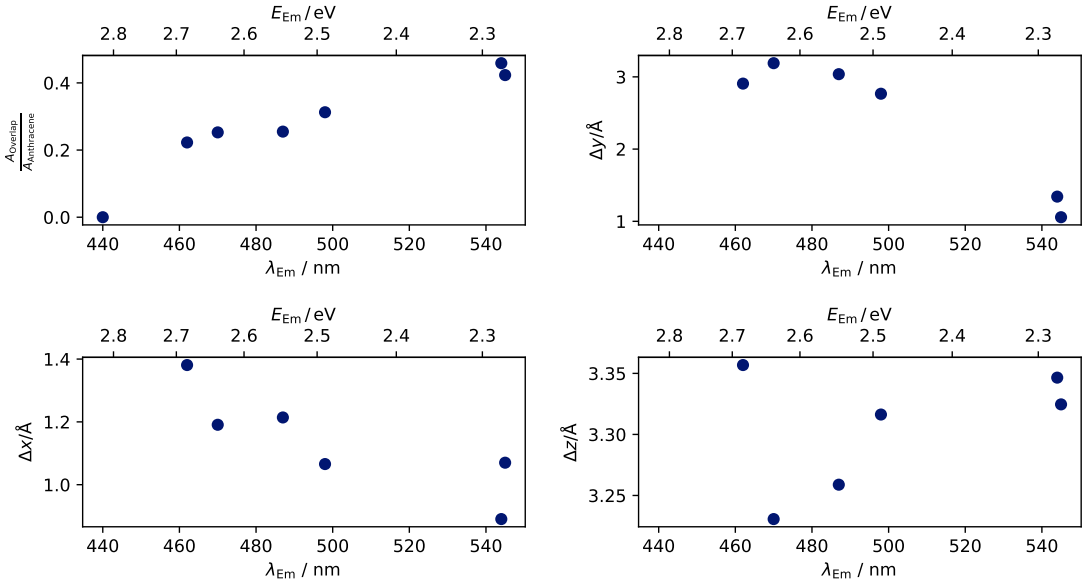


Figure 3.11.: Maximum emission wavelength and energy in dependence on the determined overlap for the different crystals of the 1-SPAnH structure.

a 6-31G(d,p) basis set. The energies of the excited state were obtained with TD-DFT calculations with the same basis set and functional and can be found in Table 3.2.

This simplistic approach proved to be unsuccessful. Not only is there a comparatively large difference between the two solvent-free structures, but the relative position of the other structures is also not reproduced. However, there is a minimal set of requirements which can be drawn from the difficulties of this very simple model. First, the two solvent-free structures should be as close as possible. Secondly, the dimer geometry derived from the quinoline co-crystal should have a higher transition energy than both of these structures. Thirdly, the geometry derived from the solid state structure formed as a co-crystal with aniline should lie between the free structures and the quinoline structure.

3.4.2. TD-DFT calculations for the substituted dimer structures

To investigate whether the missing effect is the result of a too-small system within the calculation, dimers and dimers of dimers of 1-SPAnH were investigated. As can be seen in Table 3.3 the differences are small, both between the dimers themselves and the dimers of dimers. Additionally, no systematic development for the transition between dimers and dimers of dimers can be seen. Solvent molecules were not included in the calculations.

Table 3.2.: Transition energy $\Delta E_{S_0 \to S_2}$ / eV for the different dimer structures of anthracene that match the geometry of the 1-SPAnH dimers. All theoretical calculations were carried out using a 6-31G(d,p) basis set.

Structure	CAM-B3LYP	M062X	Experimental in 1-SPAnH
free anthracene	4.1585	4.1770	3.35
quinoline	3.6932	3.7174	2.68
benzene	3.6549	3.6760	2.64
aniline	3.6821	3.7118	2.55
pyridine	3.7159	3.7437	2.49
free1	3.6786	3.6935	2.28
free2	3.6083	3.6056	2.28

Table 3.3.: Energy of first visible excitation of the dimers or dimer of dimers of 1-SPAnH in the CAM-B3LYP/6-31G(d,p) level of theory.

	Dimer	Dimer of Dimers
Structure	$\Delta E_{\text{vis},1}$ / eV	$\Delta E_{\text{vis},1}$ / eV
quinoline	3.5864	3.5901
aniline	3.6065	3.6036
free1	3.5831	3.5894
free2	3.5167	3.5016

The general problems persist, with the relative excitation energies following a similar pattern to the free anthracene calculations. The largest difference in the investigation is still located between the two free isomers, which have been observed to be the same in experiment.

To rule out that there is a basis set effect, that is specific to the dimer structures, the properties of the first solvent-free structure (free1) were evaluated using the CAM-B3LYP functional. As can be seen in Table 3.4 the overall differences between the basis sets that did converge are rather small. Unfortunately, some augmented variants within the evaluation did not show convergence, neither with DIIS nor with Fermi annealing.

As all structures contain the same two molecules, an optimisation of the excited state without any constraints should lead to the same global minimum. As such, an optimisation with fixed anthracene carbon atoms was performed. As the overlap still seems to be the determining influence, fixing the overlapping atoms seems reasonable. In practice, this means that the substituents of the anthracene, as well as the hydrogen atoms, were relaxed in the excited S_2-state. As the calculation is computationally expensive a preliminary investigation using 1-SPAnH@free1 and 1-SPAnH@quinoline was performed. With an experimental difference in emission energy of 0.4 eV they represent the largest difference observed in the measurement of the powdered compounds. The calculation was done using the CAM-B3LYP functional with a 6-31G(d,p) basis set. However, both

Table 3.4.: Results for the free1 polymorph-derived dimer of 1-SPAnH obtained with different basis sets. All calculations were done with the CAM-B3LYP functional.

Basis set	ΔE_{S_1}/eV	ΔE_{S_2}/eV	$\Delta(\Delta E)$/eV	$\mu_{tr,S_2,x}$	$\mu_{tr,S_2,y}$	$\mu_{tr,S_2,z}$
6-31G(d,p)	3.4064	3.5831	0.1767	1.4017	-0.4241	0.6865
6-31+G(d,p)	3.3317	3.5193	0.1876	-1.4205	0.4499	-0.6754
6-31++G(d,p)	3.3311	3.5188	0.1877	1.4202	-0.4509	0.6747
6-311+G(d,p)	3.3093	3.4905	0.1812	-1.4062	0.4480	-0.6585
cc-pVDZ	3.3655	3.5401	0.1746	1.3877	-0.4299	0.6728
aug-cc-pVDZ	3.2946	3.4691	0.1745	-1.3749	0.4428	-0.6463
pcseg1	3.3452	3.5306	0.1854	-1.4122	0.4273	-0.6744
pcseg2	3.2985	3.4762	0.1777	1.3983	-0.4483	0.6524
def2-TZVP	3.2956	3.4698	0.1742	1.3838	-0.4477	0.6458
aug-pcseg1	SCF did not converge					
aug-pcseg2	SCF did not converge					

structures yielded a practically indistinguishable transition energy (free1: 3.5079 eV, quinoline: 3.5084 eV).

3.4.3. Coupling Calculations

As is obvious from the discussion so far, the two molecules which can be excited and are located in close proximity can interact with each other. A direct way to quantify the strength of this interaction is the coupling J_{ij} (Davydov, 1964).

Dommet (2022) and Crespo-Otero supplied a GitHub repository, giving different possibilities for calculating the coupling with PYTHON from log files from the GAUSSIAN program. Four different methods were supplied and these were evaluated using the CAM-B3LYP functional with both the 6-31G(d,p) and 6-31+G(d,p) basis sets. The resulting coupling plotted against the emission wavelength maxima are depicted in Figure 3.12.

Energy Splitting: The interaction of the two individual excited states generates two new states, where one is raised in energy and one is lowered. Therefore, the coupling can be calculated as one half of the difference between the energies of the first E_{S_1} and second E_{S_2} excited state in the dimer calculation:

$$J_{ij} = \frac{1}{2}(E_{S_2} - E_{S_1}) \tag{3.2}$$

As the discussion on the relative S_2 excitation energies was unsuccessful, it is perhaps expected that this alternative analysis does not yield a reasonable coupling value. As

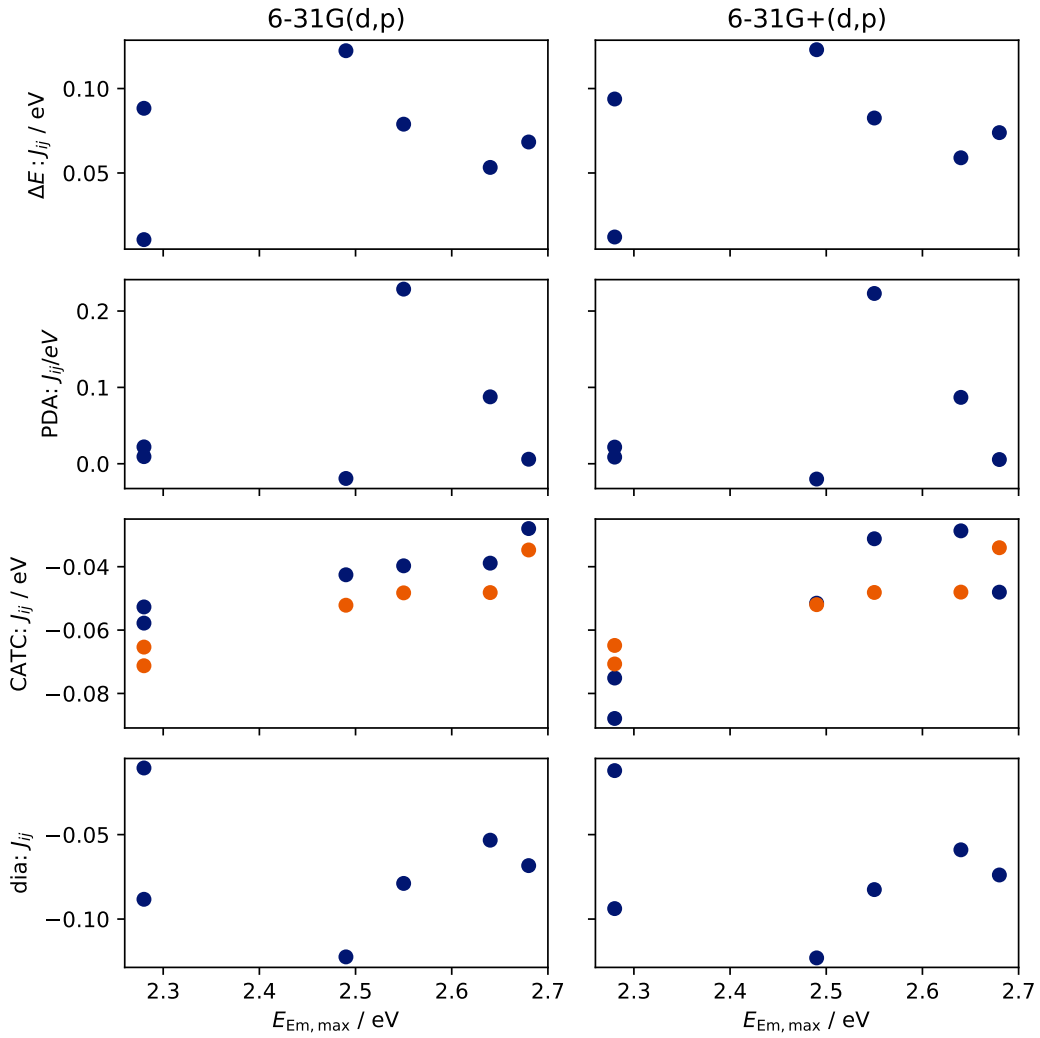


Figure 3.12.: Results of the four methods for the basis sets 6-31G(d,p) (left) and 6-31+G(d,p) (right) calculated with CAM-B3LYP. For both basis sets, the different colours in CATC represent the different methods to calculate the charges. Mulliken is depicted in blue and Hirshfeld in orange.

depicted in the top row in Figure 3.12, no systematic connection to the emission energy can be observed.

Point Dipole Approximation: The interaction can also be roughly approximated by the Coulomb interaction of the transition dipole moments of the excitation of the two individual molecules.

$$J_{ij} = \frac{\boldsymbol{\mu}_i \boldsymbol{\mu}_j}{R^3} - \frac{3\left(\boldsymbol{\mu}_i \boldsymbol{R}_{ij}\right)\left(\boldsymbol{R}_{ji} \boldsymbol{\mu}_j\right)}{R^5} \tag{3.3}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are the transition dipole moments of the individual molecules. $\boldsymbol{R}_{ij}$ and $\boldsymbol{R}_{ji}$ are the vectors from the centre of mass of molecule i to the centre of mass of molecule j and the other ways around which both have the length R.

While the relative positions of the two solvent-free molecules at 2.28 eV at least show a very similar coupling, the relative positions of all other dimers do not correlate with the observed emission maxima (second row in Figure 3.12).

Coulomb Atomic Transition Charges: A more sophisticated version of the point dipole approximation can be reached by summing the interaction energies of the transition charges.

$$J_{ij} = \sum_{a}^{N_i} \sum_{b}^{N_j} \frac{q_a q_b}{|\boldsymbol{R}_{i,a} - \boldsymbol{R}_{j,b}|} \tag{3.4}$$

This scheme is dependent on the way the transition charges are calculated. In addition to the Mulliken partitioning scheme, the Hirshfeld partitioning scheme was also employed. As can be seen in the third row of Figure 3.12, both schemes produce a reasonable correlation with the emission maxima for the CAM-B3LYP//6-31G(d,p) calculations. For the basis set including diffuse functions, the Mulliken scheme starts to show a lower correlation. As the Mulliken scheme simply assigns all orbitals centred on an atom to this atom, this is reasonable. Due to their diffuse nature, the additional orbitals do not only include contributions from the atom they are centred on. The Hirshfeld partitioning scheme on the other hand retains its correlation.

The good performance is consistent with the fact, that the overlap also has a strong correlation with the emission maximum. As a large overlap means a large number of close contacts, the calculated point interaction energies grow larger at the same time.

Approximate Diabatisation Method: This method aims to retrieve the coupling constant as part of the adiabatic Hamiltonian:

$$\mathbf{H}^{\mathrm{D}} = \mathbf{C}\mathbf{H}^{\mathrm{A}}\mathbf{C}^{T} \tag{3.5}$$

$$\begin{bmatrix} E_i^{\mathrm{D}} & J_{ij} \\ J_{ij} & E_j^{\mathrm{D}} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} E_i^{\mathrm{A}} & 0 \\ 0 & E_j^{\mathrm{A}} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} \tag{3.6}$$

In order to yield $\mathbf{C}$ the methods of Aragó and Troisi (2015) are implemented in the aforementioned Python package by Dommet and Crespo-Otero. The unitary matrix $\mathbf{C}$ is calculated by minimising the transition dipole moments from the ground state of the diabatic states and the transition dipole moments from the ground state of the isolated states.

The resulting couplings are connected to the couplings from the energy splitting by a factor of -1. As such they do not show correlation to the emission wavelengths.

3.5. Conclusion

The original aim of this project was to provide a dataset for benchmarking excited state structures. However, for different reasons the evaluated systems proved to be unsuitable for this purpose. The set of SPAnPS structures proved to be very similar for the properties that could be obtained with a suitable degree of experimental certainty. Reevaluation of the original data of the quantum efficiencies revealed a lack of reproducibility, prompting the omission of this physical property. This is quite possibly connected to the very fact prompting the initial study of this system by Fei *et al.*: the quenching of fluorescence upon solvent loss. At the same time, the calculation of quantum efficiencies is certainly non-trivial and the investigation in the solid state would have further complicated the theoretical description, even if there had been an established protocol for their measurement. The possible influence of flexibility and vibration within the system shows that the involved geometries are probably not limited to the ground state.

The energy difference between the two orientations was also evaluated as a potential target. However, host-guest complexes of larger molecules showed a less clear second position, while the crystallisation of SPAnPS with singly halogenated benzene proved to be unsuccessful.

The study of the 1-SPAnH dimer structures looked to be another promising candidate for a set of connected structures to be a potential target for a benchmarking run. However, the employed quantum chemical methods failed to reproduce the observed behaviour. The calculation of emission wavelengths is non-trivial in the solid state as the relaxed structure in the excited state is not easily available. A comparison to the absorption spectra is not possible, as the investigated powders do not easily allow

for the experimental observation of absorption. The initial hypothesis that the dimer structures without their crystal environment are sufficient to reproduce the observed shift in wavelength did not hold. Using the CATC method, the behaviour was reproduced from the atomic transition charges. However, this is unsurprising, as the distances and therefore the calculated Coulombic interactions are directly correlated with the overlap. As such a large overlap means a larger number of shorter distances and therefore a stronger shift. In summary, a relaxation of the excited state in the solid-state environment seems to be needed, which is beyond the scope of this work.

While the application to benchmarking could ultimately not be provided, the investigations delivered valuable insight into the systems under investigation. The study of 1-SPAnH did result in the development of a reliable overlap indicator and the large number of elucidated other structural parameters did raise confidence in the unique importance of the relative position. The first section could successfully evaluate previous hypotheses used for the explanation of the difference in luminescence properties of *cisoid* and *transoid* structures of SPAnPS by additional experiments conducted by myself as well as additional structures provided by Schillmöller.

Of the three possibilities determining whether the luminescence is visible, the first can be excluded with some certainty. The strong disorder, as well as the change in occupancy of the guest molecule observed for toluene, should mean that only a small fraction of solvent molecules are present in the exact relative orientation used for the discussion. As will be shown in the next chapter, changes in the behaviour due to temperature dependencies should not be overlooked.

While the twist angle β does not hold for the larger set of 9-SPAnR systems, the bend angle α seems to be connected to the quenching of the fluorescence. Additionally, the discussion of the rigidity within the system, as pursued by Krause and Schillmöller can also play a role. This is underpinned by the fact that the combined discussion of displacement and bend angle strengthens the separation between structures along their luminescence properties.

Despite the new valuable insight won, it was clear that benchmarking would ultimately need another (quantum) crystallographic approach as well as a different set of systems to study.

4. Evaluating the low-temperature phase transition in decamethylsilicocene

This chapter will illustrate the importance of looking at the low-temperature ground state when discussing interactions, while at the same time investigating a system where dispersion interactions play a decisive role in the assumed geometry.

4.1. Background

The structure belongs to the group of cyclopentadienyl compounds of group 14. Synthetic access to this type of structure was first made available by Fischer and Grubert (1956*b*) in the form of cyclopentadienyl lead. The same group also published the synthetic route to cyclopentadienyl tin in the same year (Fischer and Grubert, 1956*a*). After this quick succession, it took considerably longer to provide access to the next structure of this class in the form of cyclopentadienyl germanium reported by Scibelli and Curtis (1973). The solid-state structure of this compound was measured and reported by Grenz *et al.* (1984), while the structure of monomeric Cp_2Sn was reported by Atwood *et al.* (1981).

Changing to a heavier and more protective ligand Atwood *et al.* (1981) also reported a synthetic route and the structure of the heaviest member of the decamethylcyclopentadienyl family. At this point, the synthesis and structures of Cp^*_2Sn and Cp^*_2Ge had already been reported by Jutzi *et al.* (1980), in addition to the corresponding cations $CpSn^+$ and $CpGe^+$. The currently available set was completed by following their previously published route (Jutzi and Hielscher, 1985), which Jutzi *et al.* (1986) successfully used for providing synthetic access and the crystal structure of decamethylsilicocene, which also corresponded to the first reported structure where silicon adopts the formal oxidation number of II. Jutzi *et al.* (1989) also provided additional characterisation shortly afterwards.

The solid-state structure as initially reported is very comparable to the structure depicted in Figure 4.1 with two distinct conformers of the molecule within the same unit cell. In agreement with the germanium and tin structures (Jutzi *et al.*, 1980), the first molecule exhibits a bent formation. In contrast, the second molecule adopts an inversion symmetric staggered conformation of the two ring planes, where these planes are parallel to each other. The inversion centre is not only present in the molecule, but also in the crystal symmetry, which results in only half of this molecule's atoms being in the asymmetric

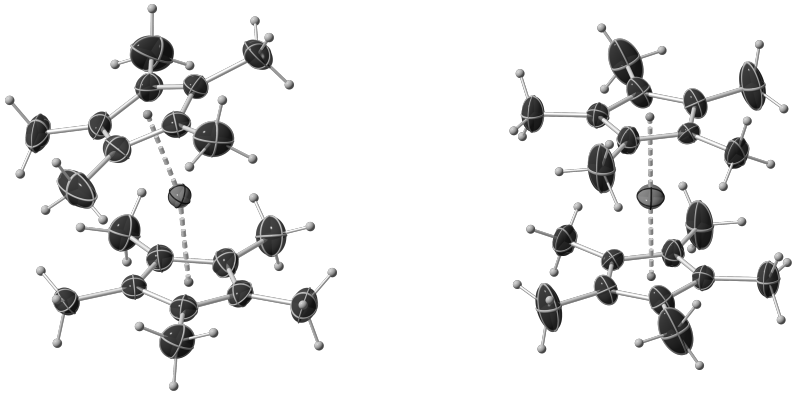


Figure 4.1.: Structure of the two different confirmations present in the high-temperature phase of Cp^*_2Si. The relative ratio of the two conformers bend/left to symmetric/right is 2:1. All ADPs are depicted at the 50 % probability level.

unit, while the asymmetric unit contains one complete bent molecule. Subsequently, the relative population of bend to symmetric molecules in the unit cell is two to one.

In the original publication by Jutzi *et al.* (1986), the presence of two distinct conformers was explained by 'intermolecular interactions and packing effects' which were presented to be responsible for the occurrence of the bent conformer. Later Jutzi *et al.* (1989) showed that the bend conformer was the most stable in the gas phase. However, the idea of packing effects kept being used in subsequent publications (Minyaev and Gribanova, 2005; Hutton, 1986) and used to rationalise differences between the theoretical structure of Cp_2Si and the experimental structure of Cp^*_2Si (Smith and Hanusa, 2001).

4.2. A phase change

The apparent conundrum of why a straight conformer is present in addition to the more stable bend conformer can be resolved by the observation of a non-destructive phase change. The low-temperature phase of decamethylsilicocene is best described in the space group $P\,2_1/c$ and it contains three molecules in bend conformation within the asymmetric unit (Figure 4.2). Due to a lowering of symmetry the asymmetric unit is twice the proportion of the total unit cell volume compared to the high-temperature phase, which is best described in $C\,2/c$.

To track this phase change, sets of single crystal diffraction data were collected between 80 K and 160 K. The phase change will now be tracked in three different descriptions, two of which are geometric and one is dependent on systematic absence differences between the space groups.

Additionally, quantum chemical calculations were employed for comparison and analysis of the phase change. Periodic plane wave calculations were conducted with QUANTUM

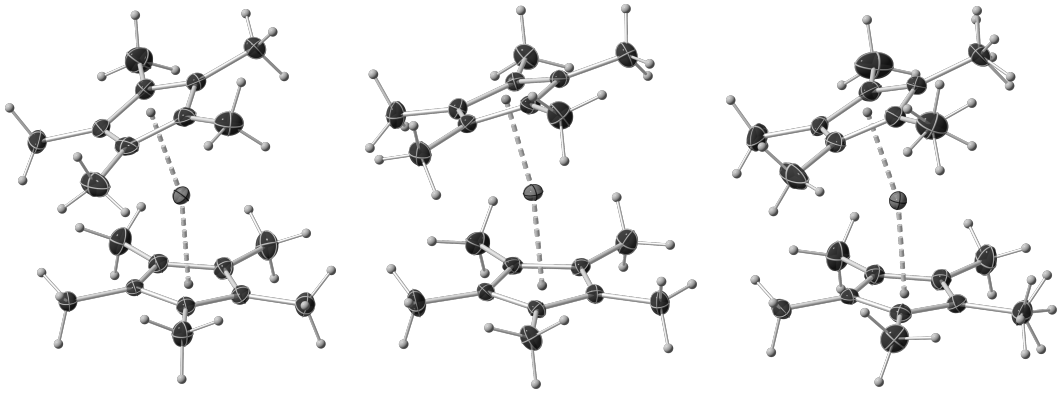


Figure 4.2.: Conformers for the three molecules of Cp^*_2Si present in the asymmetric unit of the low-temperature phase. Outer molecules are symmetry equivalent and bend at high temperature, while the structure in the centre becomes centrosymmetric. All ADPs are depicted at 50 % probability level. Additional hydrogen atoms indicate two groups refined with different occupancies summing to a total of 1.

Espresso using the PBE functionals with cut-off energies of 80 Ry for the wave function and 800 Ry for the density. Non-periodic calculations were conducted with Orca using the PBE0 functional and a def2-TZVP basis set. For all calculations, Grimme's D3BJ method was used for the dispersive energy correction.

The simplest way to track the difference in molecular geometry is the interplane angle α. It is defined as the angle between the mean planes of the carbon atoms of the two cyclopentadienyl rings of the structure. The development of the interplanar angle with temperature is depicted in Figure 4.3. The comparison geometries were obtained from the periodic plane-wave calculations, where the respective structures were relaxed under enforcement of the space group symmetry starting from the experimental geometries.

The two bend molecules, which are symmetrically equivalent at high temperature, exhibit a slight deviation of angles at low temperature. At high temperature, the angles start at 27.59(18)° and 27.15(18)° respectively (calc 26.50°). Upon cooling to 80 K, the first angle widens by 2° to 29.31(11)° (calc. 27.43°). Meanwhile, the second angle shows a much smaller deviation to 27.96(11)° (calc. 26.97°). The remaining molecule changes its conformation significantly with the interplanar angle starting from a nominal value of 0.8(2)° at 160 K (set to 0° in the calculation) when the high-temperature phase is refined in $P\,2_1/c$. The interplanar angle increases upon cooling, reaching a value of 20.40(10)° (calc 20.40°) at 80 K. As such, the resulting angle is still smaller than the angle observed in the other two molecules in the asymmetric unit.

The second possibility is to track how much atom positions break the symmetry of the high-temperature phase. To quantify this difference, the atom positions were mapped onto their symmetry equivalents using the additional inversion centre of $C\,2/c$. Then

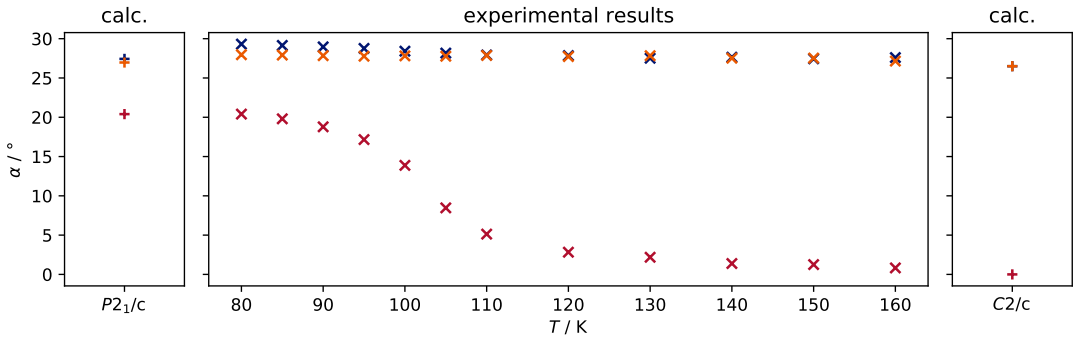


Figure 4.3.: Interplane angle α in dependency of the temperature measured for the three conformers of the $P\,2_1/c$ asymmetric unit. Calculated values were derived from plane-wave calculations.

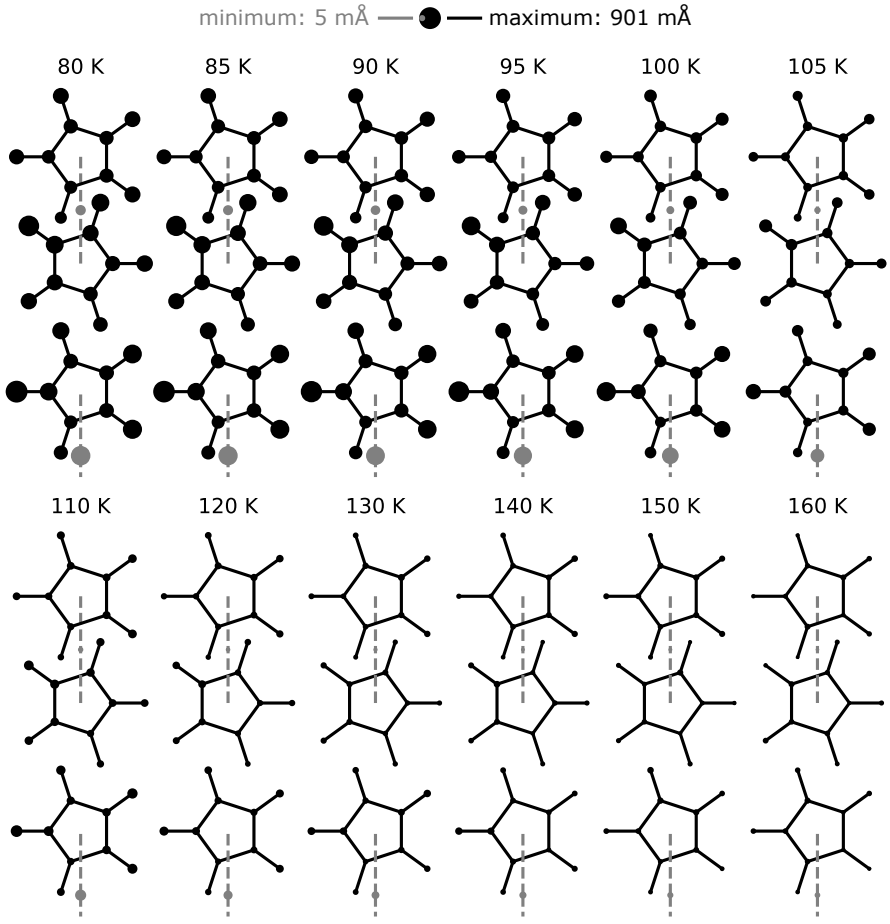


Figure 4.4.: Distance of atom positions from their symmetry equivalents in $C\,2/c$ mapped onto the atomic position by the additional inversion centre of the high-temperature space group.

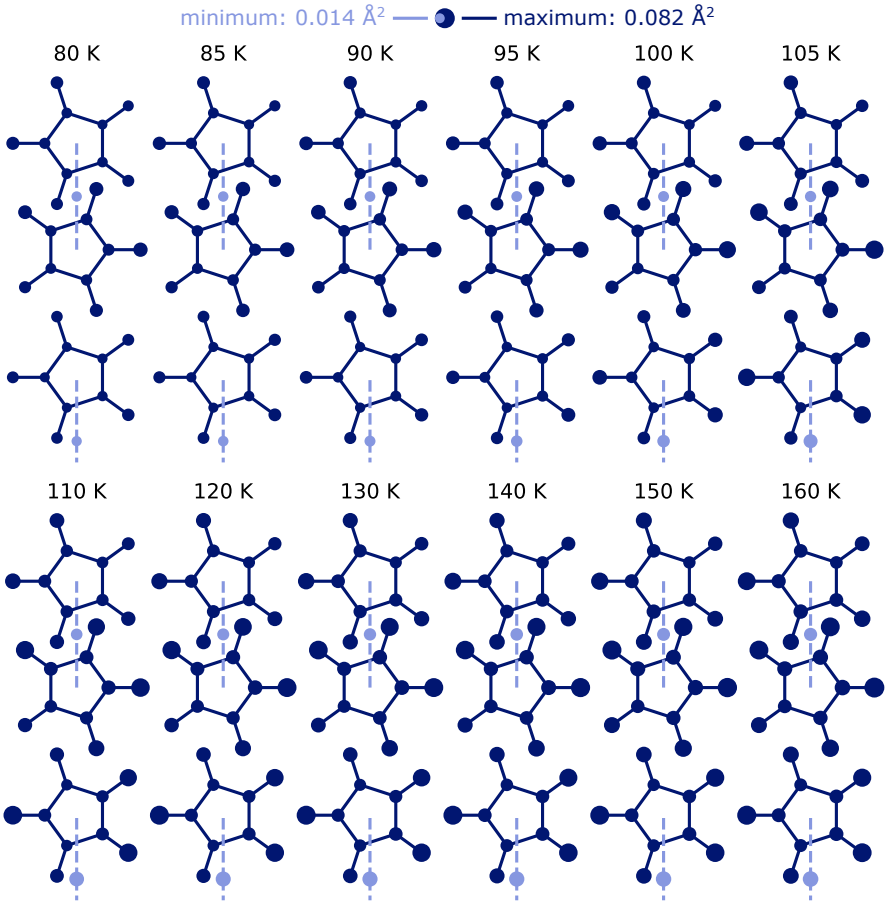


Figure 4.5.: Averaged U_{equiv} values of the two symmetry equivalent atoms in $P\,2_1/c$.

the distance between the symmetry-generated atoms and the atoms with the refined second position can be measured. The resulting values are depicted in Figure 4.4. It is clearly visible what kind of geometry change leads to the breaking of symmetry for the two respective molecules. For the molecule, which is completely in the asymmetric unit in both phases, one of the two rings (here depicted at the centre) opens up, with the largest point of movement being on a single atom. This atom is not pointed directly to the opening but to the side, which indicates that the molecule itself is already close to an optimal opening angle. In contrast, the second molecule shows the highest deviation on three of the methyl carbon atoms, which form one pair of neighbours, while the other atom lies on the opposite side of the cyclopentadienyl ring. This indicates a pitching geometry change, which results in the inversion centre present at high temperature being broken upon cooling.

Interestingly, the potential room for this movement is filled by vibration at high temperature. Figure 4.5 shows mean U_{equiv} values for the atomic positions. Comparison to

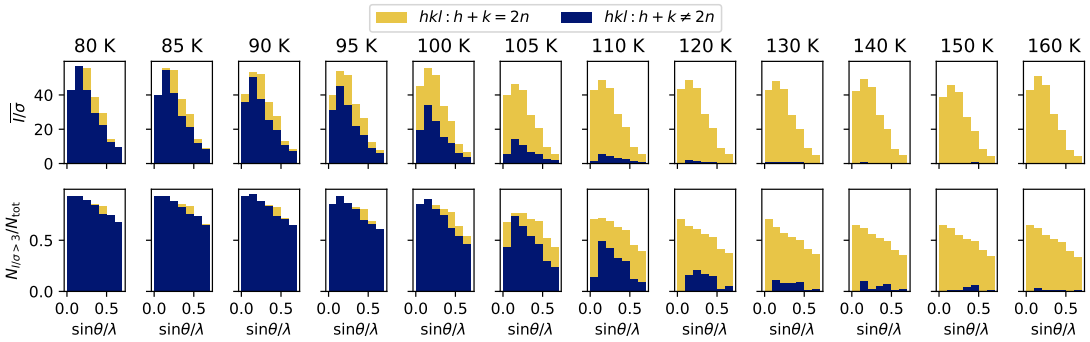


Figure 4.6.: Development of reflection intensity over estimated standard deviation (top) and the fraction of reflection fulfilling the significance criterion (bottom) in dependence of temperature. Colouring indicates whether reflections are systematically absent in $P\,2_1/c$ (blue) or present in both space groups (yellow).

Figure 4.4 shows that the same atoms**,** which break the symmetry at low temperature, exhibit larger thermal motions at high temperature.

The third possibility to observe the phase change is to track the difference in systematic absences between the two space groups $C\,2/c$ (high temperature) and $P\,2_1/c$ (low temperature). While reflections with $hkl : h + k \neq 2n$ are systematically absent in the space group $C\,2/c$, they are present in the low-temperature phase. The development of the reflections is depicted in Figure 4.6 and the change can be assigned to the temperature region between 90 K and 120 K. The mean intensity over estimated standard deviation for the reflections fulfilling this criterion changes from 20.3 to 0.9 within this region. At the same time, the reflections present in both space groups change only from 20.9 to 15.9.

4.3. The importance of dispersion for the minimum structure

Both structures are present in the solid state at high temperature, while the low temperature favours the bend formation, even if the angle of the third molecule does not converge to the larger bend angle of the two remaining molecules. A calculation will now be employed to answer two questions. Firstly, what is the theoretical energy difference between the two structures in isolation and between the two phases in the solid state? Secondly, how important are dispersive interactions for the observed difference?

The overall energy difference in the ultrasoft pseudopotential calculations with optimised geometries of full unit cells, and optimised unit cell parameters under the constraint that resulting cells still need to be monoclinic, yields an energy difference of 0.191 eV between the two complete unit cells with the lower energy structure being the low-temperature

structure in $P\,2_1/c$. This corresponds to an energy difference of only -1.54 kJ/mol Cp^*_2Si with 12 molecules per unit cell.

With both conformers present in the solid state, the discussion of how large the energy difference between the two conformers is, has already had significant contributions. Schoeller *et al.* (1999) provided a very in-depth discussion and used CCSD(t) calculations to show that the analogue without methyl groups (Cp_2Si) is bent. Using B3LYP/6-31G(d) calculations, they also concluded that the influence of the methyl groups is small. A more global approach by Sapunov *et al.* (2001) included Cp^*_2Si. The approach investigated through space coupling and attributed the resulting structure to a compromise between van-der-Waals interaction and repulsion.

Consequently, calculations including dispersion interactions were performed and the structure of Cp^*_2Si was optimised using the PBE0 functional and Grimme's B3BJ scheme in a non-periodic calculation. The straight conformation was obtained by enforcing a C_i symmetry, while the bend conformation was optimised without symmetry constraints (C_1 symmetry). The non-periodic calculation yielded an energy difference of about $4.5\,\mathrm{kJ\,mol^{-1}}$. A large part of this difference ($3.5\,\mathrm{kJ\,mol^{-1}}$) can be attributed to dispersion, which corresponds to a large part of the stabilisation of the bend conformation. In fact, the straight conformation does not constitute a local minimum in these calculations, shown by two Cp* ring slippage vibration modes yielding imaginary frequencies in the calculation.

In the final step, the interaction will be localised. While it seems most plausible, that dispersive interactions occur between the methyl groups, this fact will be verified using NCIPlot (Contreras-García *et al.*, 2011). The resulting depictions can be found in Figure 4.7. They clearly show that the dispersive interactions are located between the outer methyl groups. This further underpins the reasoning of dispersive interactions between the methyl groups being a factor in the overall stabilisation. Arguably, the centrosymmetric structure also shows dispersive interactions. However, the larger distance between methyl groups reduces the strength of these interactions as van-der-Waals interactions have a distance dependency of r^{-6}.

4.4. Conclusion

What was so far known as a structure behaving uniquely compared to structures from other elements from the same group, was rationalised as a high-temperature phase under entropic driving force for the observed conformation. At low temperature, all structures could be shown to adopt a bend position, albeit with different angles. A large part of the stabilisation could be attributed to dispersive interactions, even though the overall difference was calculated to be low, which is reasonable as both conformers are present under ambient conditions.

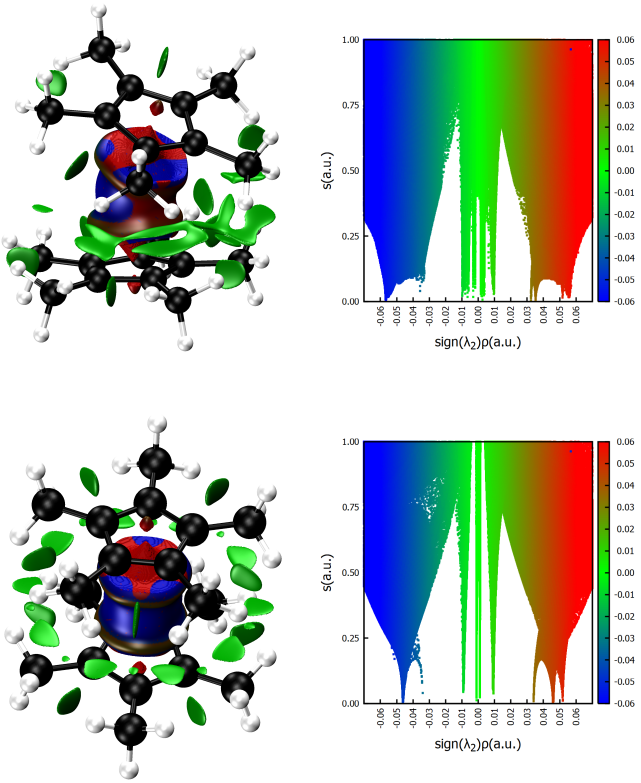


Figure 4.7.: NCI Analysis (Contreras-García *et al.*, 2011) plots for the structures optimised in ORCA. The depictions in the top row are calculated for the bend conformer, while the depictions in the bottom row are calculated for the straight conformer. On the left, the isosurfaces represent a reduced density gradient at a level of 0.5 atomic units. The green colour indicates a low-density accumulation corresponding to dispersive interactions. The right column depicts the reduced density gradient plotted against the sign of λ_2 multiplied by the electron density ρ.

The investigation shows the problem with the 'explanation' of differences due to not further specified packing effects. Profiting from over 35 years of technological advancement since the original publication of the structure, this chapter reveals that the difference explained with the packing effect term is not to be localised in energy but entropy effects. The equilibrium at room temperature is therefore not at the minimum of the electronic energy surface. Keeping this in mind, it should be a warning that the explanation with packing effects was still successful, ultimately revealing the fuzzy nature of the concept and its limited explanatory and predictive power, when the interactions comprising these packing effects are not further specified.

At the same time, this investigation is one of many demonstrations, that routine crystal structures should be performed at low temperature whenever possible, especially if statements with a high degree of generalisability are to be derived from the measurements.

5. A new metal for in-house use: Evaluating the performance of an Indium MetalJet diffractometer

Chapter 2 has introduced the kinematical model of describing X-ray diffraction experiments and the approximations used to include effects of non-kinematical diffraction (Section 2.5.3 and Section 2.7). These models are highly successful and are a very good description of a lot of structures determined with molybdenum K_α radiation, which can be seen as one of the work-horses of high-resolution charge density refinements, with the other being synchrotron radiation.

However, for heavier elements resonance scattering, as well as extinction can become a problem. As long as the measurement is not close to an absorption edge, a lower resonance scattering can be achieved with a lower wavelength / higher energy radiation (Becker and Coppens, 1974*a*; Hamilton, 1965; Becker and Coppens, 1974*b*).

Therefore, the choice of the correct wavelength for the X-ray diffraction experiment is highly dependent on the content of the crystal which is to be investigated. Softer radiation shows benefits for weakly diffracting samples which need its larger scattering power, as well as the higher intensity within the same class of X-ray source. Additionally, the relative size of the resonant scattering to the elastic scattering increases, which enables a higher degree of certainty of the determined absolute structure in non-centrosymmetric structures.

Hard radiation is particularly desirable for compounds which are bound by ionic interactions and which include no or only small organic ligands. These compounds combine a high scattering density and the strong interactions often exacerbate high diffraction and thereby extinction. While corrections exist, they introduce additional parameters with a potential correlation to the crystallographic quantity of interest. Very strong resonance scattering and extinction can even bring these models to their limits.

So far, employing silver K_α has been the only widely available choice for in-house sources when facing these problems. This chapter aims to explore the potential of an X-ray source with even higher energy, namely using indium K_α radiation produced by an Excillum MetalJet source for the study of compounds with high absorption and extinction potential, as well as the performance for compounds already well established for benchmarking in X-ray crystallography.

5.1. The MetalJet X-ray source

Generally, the MetalJet and all other in-house sources follow well-established principles. The original discovery of X-ray radiation by Röntgen (1895) already used a setup where electrons emitted from a cathode were accelerated onto an anode target, which then produced the X-ray radiation. The specific apparatus was of the type described by Crookes (1879). For in-house X-ray diffraction, this principle for the production of X-rays has been retained until the present day. Improvement in produced intensity was achieved by using tungsten cathodes (Coolidge, 1916, 1925). However, the current of electrons onto the anode was still limited by the thermal load put onto the anode by the impact of accelerated electrons. A successful mitigation strategy was provided in the form of rotating anodes, where the spread of the thermal load enables higher X-ray intensities (Gray, 1930) and the current iterations of this technologies provide some of the highest intensities for in-house use.

The problem of heat dissipation was elegantly solved with the introduction of the liquid MetalJet method (Hemberg *et al.*, 2003). For this technique, the anode is a liquid jet, which is expanded into the vacuum at high pressure. Subsequently, heat dissipation is no longer a problem as the heated fraction of anode material is exchanged for new anode material in an instant. The initial demonstration used an alloy consisting of tin and lead as jet material, but the authors already mentioned the use of other liquid alloys as potential candidates. At the same time, the employed cathode consisted of an LaB_6 crystal instead of the usual tungsten filament. This enables the production of electrons at the sharp edge of the crystal. While traditional X-ray sources do not profit from the gain in focus of this cathode as they are limited by heat dissipation, the presented MetalJet can see a benefit from the higher focus.

With the stability of the source confirmed by the Hertz group (Otendal *et al.*, 2007), the first application with a liquid gallium alloy followed quickly (Otendal *et al.*, 2008), while the first production of 24.2 keV indium radiation was presented three years later (Larsson *et al.*, 2011).

In 2007, the company 'Excillum AB' was founded to make the technology available commercially. They introduced the pumping of the alloy in a closed loop instead of using a reservoir to be refilled, thereby making the continuous operation of the machine viable. Additionally, they introduced several different alloys of gallium, indium and tin, to be chosen depending on the wavelength to be used in the experiments (gallium K_α or indium K_α).

The X-ray diffraction experiments used the high brilliance of the gallium radiation to obtain diffraction patterns refined with the independent atom model (see for example Diop *et al.*, 2016; Rajak *et al.*, 2020), where the high brilliance can save on measurement time and enable the study of weakly diffracting samples. Consequently, the use for the study of metal-organic frameworks has also been reported (Klein *et al.*, 2017). However, gallium K_α has a wavelength of about 1.34 Å. The resolution limit for X-ray diffraction

experiments used for the refinement of charge density is at least 0.5 Å and the desired resolution is usually beyond 0.45 Å. With a maximum theoretical resolution for gallium K_α of only 0.67 Å ($d_{\min} = \lambda/2$), the indium K_α part with a wavelength of about 0.51 Å is the only sensible option for high-resolution X-ray diffraction with a MetalJet source and the alloys commercially available at this moment.

The MetalJet setup in Göttingen was installed in September 2017. Starting from that date it underwent some significant optimisation, where the investigation from the university side was done in cooperation between Nico Graw and me. For the details, see the PhD thesis of Graw (2022), where the first chapter sketches the path to the setup in its current iteration. In addition, this chapter in his thesis also contains a comparison conducted by Graw which evaluates the performance of the MetalJet at 200 W against an Incoatec Ag IμS 2.0 High Brilliance.

In the iteration of the setup used for this work, the MetalJet D2 in Göttingen works at an acceleration voltage of 70 kV and a power of 250 W. The used alloy is Excillum I3 which consists of 75 % gallium and 25 % indium. The machine is housed in a Bruker D8 Venture enclosure with a Bruker goniometer of the same name. The Incoatec Helios Montel optics are the ones, which were provided in the first place as the experiments in the optimisation evaluated these optics favourably.

5.2. Detector technology for reducing low-energy contamination

After the optimisation until this point, two questions remained. The first question was which of two possible detectors, the Bruker Photon III or the Eiger2 R CdTe 1M, provides superior performance with the MetalJet. This includes investigations for high-absorbing compounds, which were evaluated using the independent atom model. The second question was, what performance is to be expected for high-resolution measurements with aspherical refinements, especially with potential applications using the multipolar model. This was evaluated for both detectors, as well as a setup using an Incoatec Ag IμS 3.0 source with a Photon III detector for comparison.

5.2.1. Bruker Photon III

This detector is an improved version of its predecessor, the Photon II detector and likewise uses a rare-earth X-ray scintillator to convert X-rays into visible light. This light is subsequently detected by a CMOS sensor chip, with a large detector area of 110 mm · 140 mm covered with pixels of a size of 135 µm · 135 µm. The detector noise is efficiently reduced by adaptive oversampling, where the fast non-destructive readout rate of 70 Hz is used to distinguish photon detection events from background noise. This signal processing is done in real-time.

The detector offers two measurement modes, dynamic mode and mixed mode. While the dynamic mode offers a higher dynamic range, mixed mode offers a more sensitive detection, scaling to single photon detection events and thereby lower background noise. The scintillator cannot be easily saturated. However, the memory well on the detector in dynamic mode is limited to 163810 counts. To account for this fact, strong reflections are usually measured with a lower frame time / a higher rotation speed of the crystal. Optionally they are determined in a fast scan mode and only reflections which cannot be obtained otherwise are included from these fast scan measurements. Mixed mode usually shows a lower maximum detection rate as the price for its increased sensitivity.

In the iteration present on the machine in Göttingen, the Photon III offers a comparatively low quantum efficiency for indium radiation of only 53 %. According to Bruker, this has been mitigated with the new 'High Energy' or HE variant. However, this was not available at the point of this investigation.

5.2.2. Dectris Eiger2 R CdTe 1M

The second detector is a single photon counting detector using a cadmium telluride layer for direct detection of individual photons *via* the charge separation induced by X-ray photons on the detector. Despite their fairly complex composition individual pixels have a size of only 75 µm · 75 µm with the total active area being 77.1 mm · 79.65 mm.

The direct conversion enables the setting of an energy threshold. Usually, this is set to half the photon energy to exclude the possibility of double counting X-ray photons hitting the detector at the border between two pixels. The Eiger2 family even offers the possibility of setting a second higher energy threshold to exclude artefacts from cosmic radiation.

Using a home source the detector has a high dynamic range with a maximum count rate of $5.5 \cdot 10^6$ photons per pixel and second and basically no background noise from the detector itself. The photon efficiency of the cadmium telluride semiconductor for indium K_α radiation is near unity with 95 %. At the same time, the maximum readout rate is 100 Hz.

5.2.3. Tailored approaches to low energy contamination

The combined alloy of indium and gallium poses a challenge from an instrumental perspective. While the Helios Mirror optics on the instrument focuses the indium radiation more efficiently compared to gallium, four factors lead to a high detected share of gallium radiation. Firstly, there is a higher overall gallium content in the alloy, compared to indium. Secondly, compared to indium, the production of gallium K_α radiation is more efficient. Thirdly, the higher scattering cross section for gallium radiation means a larger share of the gallium radiation arriving on the crystal is scattered. The final consideration concerns the Photon III detector, where the higher quantum

efficiency for gallium radiation also favours the detection of the lower energy. All of these points combined lead to a significant amount of unwanted low-energy contamination in the measurement. This low-energy contamination has to be treated differently depending on the employed detector.

The approach for the Eiger2 CdTe is straightforward. Using the energy cutoff feature of the detector, the gallium radiation can be filtered. As the photon energy of gallium (9.6 keV) is less than half of the indium radiation energy (24.2 keV), the optimal value of the energy cutoff for indium radiation can be chosen.

In contrast, the Photon III detector does not offer energy discrimination to use for this purpose. This means that an attenuator has to be used to eliminate the low-energy gallium radiation. Most filtering materials do absorb gallium radiation more efficiently. The approach in this work was to use 40 µm of palladium. It offers additional filtering with its absorption edge just above the energy of indium K_α. While the theoretical attenuation of gallium with this attenuator is 99.9 %, the indium radiation suffers from an attenuation of 39.4 % as well.

5.3. Structures in the investigation

In total, five structures will be used to assess the performance of the two detectors on the setup in Göttingen. For the determination using the Independent Atom Model, three strongly absorbing inorganic salts were used. The first test crystal was scandium cobalt carbide $[ScCoC_4]_n$ **1**. The solid-state structure of this compound was originally described by Jeitschko *et al.* (1989). Subsequently, it has been the subject of a number of investigations (Zhang *et al.*, 2007; He *et al.*, 2015), including being evaluated using high-resolution X-ray diffraction (Rohrmoser *et al.*, 2007; Eickerling *et al.*, 2013; Langmann *et al.*, 2021). The other two compounds selected for the investigation were scandium platinum silicate [ScPt$_9$Si$_3$]n **2** and sodium tungstate dihydrate $[Na_2WO_4 \cdot H_2O]_n$ **3**. To also include an indication for performance on crystals with low absorption, L-Alanine [$C_3H_7NO_2$] **4** was included, being a common compound for benchmarking investigations.

The MetalJet setup with either detector was also to be compared to a second machine, routinely used for X-ray diffraction with comparatively hard radiation. To this end, data was collected on an YLID crystal [$C_{11}H_{10}O_2S$] **5** and refinements using aspherical form factor descriptions were conducted. While the use of this compound for benchmarking diffractometers is well-established using room-temperature measurements, both polymorphs have also been characterised at low temperature by Guzei *et al.* (2008). Measurement at this temperature provides diffraction well suitable for the purposes of the multipole and Hirshfeld atom refinements used for the comparison presented in Section 5.7. The details of the measurement and the data processing can be found in Section A.3.

Table 5.1.: Crystal properties for the comparison measurements of this chapter.

		Space group	Crystal dim. / mm	μ / mm^{-1}	μr
1	$[ScCoC_4]_n$	*Immm*	0.592/0.063/0.031	3.9	0.06
2	$[ScPt_9Si_3]_n$	*C*2/*c*	0.059/0.049/0.041	64.3	1.29
3	$[Na_2WO_4 \cdot 2\,H_2O]_n$	*Pbca*	0.208/0.157/0.086	8.2	0.35
4	L-Ala	$P2_12_12_1$	0.214/0.155/0.128	0.1	0.00
5	YLID	$P2_12_12_1$	0.395/0.387/0.312	0.1	0.03

Table 5.2.: Measurement details for the comparison measurements in this chapter.

		d_{min} / Å	T / K
1	$[ScCoC_4]_n$	0.39	100
2	$[ScPt_9Si_3]_n$	0.38	100
3	$[Na_2WO_4 \cdot 2\,H_2O]_n$	0.36	100
4	L-Ala	0.45	150
5	YLID	0.45	110

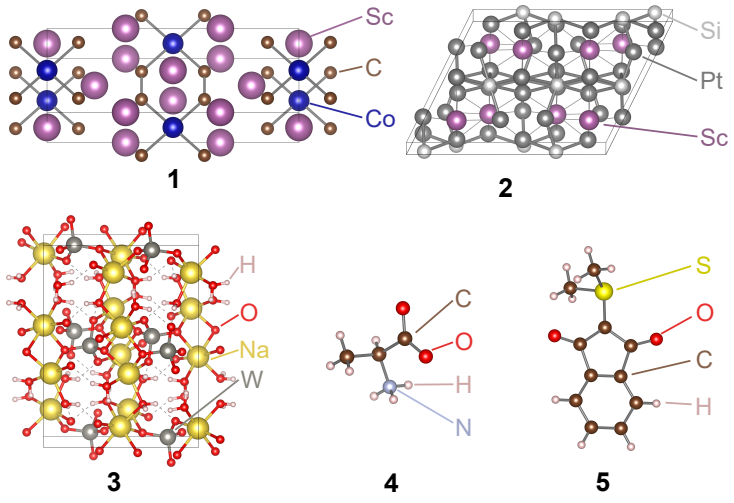


Figure 5.1.: Solid-state structures for the compounds investigated in this chapter.

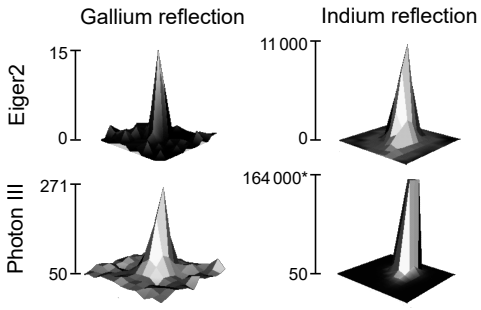


Figure 5.2.: Reflection profiles for (2 1 2) measured with an exposure time of 15 s (dataset **1**). The filtering is incomplete for both cases. The indium reflection on the Photon III is overexposed and the recorded intensity marked with * is the detector limit.

5.4. Evaluating the efficiency of the low energy filtering

The different solutions for filtering the low energy contamination for the two detectors were discussed in Section 5.2.3. The question is now, how well these approaches work in practice. Qualitatively, this question can be answered by looking at the remaining intensity at the gallium spots. As such, four reflection profiles of the reflection (2 1 2) were created from the measurement of **1**. The resulting comparison can be found in Figure 5.2.

Both approaches still show maxima at the position of the reflection corresponding to the peak position for gallium radiation. As such, it is obvious that neither approach offers a full suppression of the gallium intensity. Accordingly, the next step was to investigate the difference in filtering quantitatively using the twinning routine of SAINT. In the first step, the integration was conducted normally for the unit cell corresponding to the indium radiation. The determined cell, instrument parameters and crystal orientation resulting from this integration were used as the basis of the following investigation. The obtained cell was copied and scaled to match the equivalent gallium positions. Using a fixed orientation and fixed cell parameters, the intensities of both wavelengths were then evaluated by a second integration in SAINT. After absorption correction and scaling in TWINABS (Sevvana *et al.*, 2019), the two components were refined as a twin in SHELXL (Sheldrick, 2015*a*). The determined relative percentages carry more information than the absolute ones, as differences in scattering cross sections, absorption and anomalous dispersion have been neglected. The determined twin partitions were 0.226(11) % for the Eiger2 CdTe and 0.44(4) % for the Photon III detector, respectively. This indicates improved filtering of the Eiger2 CdTe compared to the Photon III detector. This could be counteracted with an increase in the thickness of the palladium attenuation, but this would further reduce the indium K_α intensity.

Table 5.3.: Precision indicators for the datasets evaluated with the independent atom model (**1**: $[ScCoC_4]_n$, **2**: $[ScPt_9Si_3]_n$, **3**: $[Na_2WO_4 \cdot 2\,H_2O]_n$, **4**: L-Ala). All indicators were determined for the full measured and the multiplicity equivalent sets of data. Values in brackets denote the standard deviation of the indicator over the 100 multiplicity equivalent sets.

			1	**2**	**3**	**4**
Multiplicity	full	Eiger2	31.22	7.83	15.31	12.56
		Photon III	60.78	32.63	24.38	31.83
	equal	both	25.71	7.76	10.68	7.61
$R_{\text{r.i.m}}$ / %	full	Eiger2	2.35	6.29	3.76	3.75
		Photon III	4.15	9.35	3.97	5.04
	equal	Eiger2	2.303(5)	6.3296(20)	3.502(5)	3.323(8)
		Photon III	3.305(12)	6.98(2)	4.539(10)	5.94(4)
$R_{\text{p.i.m}}$ / %	full	Eiger2	0.39	1.82	0.97	0.90
		Photon III	0.47	1.35	0.69	0.85
	equal	Eiger2	0.42	1.8397(17)	1.04	1.099(5)
		Photon III	0.598(4)	2.135(8)	1.339(3)	1.963(14)
$\langle I/\sigma \rangle$	full	Eiger2	117.24	20.03	41.00	41.90
		Photon III	110.13	28.46	44.37	45.61
	equal	Eiger2	108.063(13)	18.958(4)	36.529(7)	37.73
		Photon III	80.19(3)	14.830(9)	28.899(8)	20.95(2)

Table 5.4.: Precision indicators for the YLIDE (**5**) dataset evaluated with the aspherical atom models. All indicators were determined for the full measured data.

Radiation	Detector	Multiplicity	$R_{\text{r.i.m}}$ / %	$R_{\text{p.i.m}}$ / %	$\langle I/\sigma \rangle$
In K_α	Eiger2	29.04	2.48	0.37	101.99
	Photon III	27.99	3.83	0.59	78.14
Ag K_α	Photon III	23.33	2.46	0.39	93.91

5.5. Comparison of the precision of the measured data

The precision measures how large the statistical spread is between symmetrically equivalent reflections. This was evaluated using the redundancy-independent merge $R_{\text{r.i.m}}$, the precision indicating merge $R_{\text{p.i.m}}$ and the mean intensity over estimated standard deviation $\langle I/\sigma \rangle$. Their definitions can be found in Section 2.9.1. The indicators were calculated for the full datasets as well as the multiplicity equivalent datasets, which were prepared as described in Section A.3.2. The resulting values for datasets **1**–**4** can be found in Table 5.3.

Comparing the multiplicity differences of the individual datasets reveals why a second evaluation using multiplicity equivalent sets of reflections was included. The large discrepancy stems from three factors: Firstly, the Photon III has a larger active area.

Secondly, multiple runs with increasing exposure times were included in the Photon III measurement to ensure a maximum exposure time within the overexposure limit of the detector. Finally, both mixed-mode and dynamic mode measurements were included for the Photon III measurements. Subsequently, the multiplicity of the Eiger2 CdTe datasets is lower as the higher maximum exposure limit reduced the need for such an approach and only one measurement mode needs to be considered. Obviously, the multiplicity equivalent datasets have the same multiplicity.

The difference in multiplicity does affect both the determined values for $\langle I/\sigma \rangle$ and $R_{\text{p.i.m}}$. While the values determined from the full dataset favour the Photon III detector in three of four cases, the indicators determined from the multiplicity equivalent datasets favour the Eiger2 CdTe without exception. Due to its construction, the $R_{\text{r.i.m}}$ value is independent of the multiplicity. Accordingly, it favours the Eiger2 CdTe detector for all datasets for both the complete as well as the redundancy equivalent datasets.

As the comparison for the aspherical atom refinements includes three datasets, the sorting procedure was not employed. Instead, care was taken to keep the multiplicities of the individual measurements within a closer window and this aim was achieved to a reasonable degree. The precision indicators of the Metaljet equipped with an Eiger2 CdTe detector and the silver IμS were basically identical, except for the $\langle I/\sigma \rangle$, which was higher for the MetalJet/Eiger2 CdTe setup. The combination of MetalJet and Photon III detector fares worse. All indicators show inferior values to the other two setups.

5.6. Results of Independent Atom Model refinements

In this evaluation, the direct comparison of the two detectors on the MetalJet in refinements using the independent atom model approximation was evaluated. The evaluations used the dataset preparation as described in Section A.3.2 for preparing the full, as well as the multiplicity equivalent datasets. All refinements were done in SHELXL (Sheldrick, 2015*a*) and employed a weighting scheme. The results are listed in Table 5.5. The definition of the evaluated quality indicators can be found in Section 2.9.2.

The crystallographic agreement factor ($R(F)$) of the full datasets does not favour either detector. The Eiger2 CdTe is clearly superior in two of the four measurements (**1** and **3**), while the Photon III shows a lower value for one dataset (**2**). The difference in agreement factor for the last dataset (**4**) is only 0.05 %. Recalling the approximations of the independent atom model, this can be considered equivalent. The e_{gross} closely follows the relative performance of the crystallographic agreement.

The ambiguity in relative performance is not retained for the multiplicity equivalent evaluations. Here, the Eiger2 CdTe shows an improved performance in all four datasets for both the $R(F)$ and the e_{gross}. Differences are always significantly larger than the uncertainty introduced by the choice of which reflections are to be omitted for assuring

Table 5.5.: Quality indicators of the independent atom model refinement of the four structures: **1**: $[ScCoC_4]_n$, **2**: $[ScPt_9Si_3]_n$, **3**: $[Na_2WO_4 \cdot 2\,H_2O]_n$and **4**: L-Ala. $R(F)$ was calculated using all data. Bond precision $\sigma(d)$ was evaluated for the following bond types: **1**: Co−C, **2**: Pt−Si, **3**: W−O and **4**: C−C.

			1	**2**	**3**	**4**
$R(F)$ / %	full	Eiger2	0.97	3.04	1.92	3.34
		Photon III	1.35	2.72	2.28	3.29
	equal	Eiger2	0.930(5)	3.029(3)	2.079(6)	3.221(2)
		Photon III	1.352(15)	3.85(3)	2.361(6)	4.53(4)
e_{gross}	full	Eiger2	9.8	424.5	143.7	10.3
		Photon III	10.3	395.7	173.2	10.8
	equal	Eiger2	5.70(8)	419(7)	143.0(5)	8.45(3)
		Photon III	9.52(8)	518(4)	149.3(3)	11.20(11)
$\sigma(d)$ / mÅ	full	Eiger2	0.6	0.9(3)	0.6	0.7
		Photon III	0.7	0.81(16)	0.7	0.6
	equal	Eiger2	0.6	0.9(3)	0.6	0.7
		Photon III	0.7	1.0(3)	0.65(5)	0.83(5)

multiplicity equivalence. The standard deviations of the results from the 100 refinements are well below the differences between the two detectors.

The uncertainties in bond lengths are always below 1 mÅ and are not systematically different between the two detectors for either the multiplicity equivalent or the full datasets. As such, the MetalJet is suitable for independent atom model measurements with either detector, which should be somewhat unsurprising.

5.7. Results of aspherical refinements and comparison to an established X-ray source

As explained in detail in Section 2.5, the electron density around an atom is spherically symmetric only in a coarse approximation and newer methods have been developed to describe atoms aspherically. Using the data measured for the YLIDE (**5**) crystal, this section shall evaluate the performance of the two detectors for both the PAW-HAR method, tackled in detail in Chapter 6 and by Ruth *et al.* (2022), as well as the multipolar model.

In contrast to the last section which only compared the two detectors, this evaluation will also include a comparison with data collected on an Incoatec IμS 3.0 silver microsource.

5.7.1. Comparison of quality indicators

The numerical factors for determining the agreement between data and model are very similar for both aspherical models (Figures 5.3 and 5.4). The unweighted crystallographic agreement factor $R(F)$ as well as the e_{gross} value follow the $\langle I/\sigma \rangle$ of the individual measurements discussed in Section 5.5. Accordingly, the Eiger2 CdTe dataset shows improvement over the Photon III datasets. Within the two datasets collected with the Photon III detector, the e_{gross} for the dataset from the Ag IμS is superior to the dataset of the Metaljet independent of the refinement method, the $R(F)$ only shows an improvement in the same direction for the multipole model. The performance measured by $R(F)$ is basically identical in the Hirshfeld atom refinement.

A less conclusive picture is presented by the analysis of the weighted agreement factor $wR_2(F^2$ and the goodness of fit (GooF). While the Metaljet – Photon III setup still performs the worse of the three investigated setups, the Ag IμS shows similar performance in $wR_2(F^2)$ for the Hirshfeld atom model refinement but a small improvement in the multipole refinement compared to the MetalJet - Eiger2 CdTe setup. The GooF shows a superior performance of the MetalJet/Eiger2 CdTe in the Hirshfeld atom refinement and for the Ag IμS/Photon III in the multipole refinement.

Both Photon III measurements show a higher absolute value for the most negative difference electron density. This effect can be connected to an overall shift in the Henn-Meindl plot towards values of more negative difference electron densities. While this seems to be an overall shift for the Ag IμS / Photon III measurement, resulting in an overall less pronounced maximum in the positive difference electron density, the MetalJet / Photon III dataset shows the largest value for this maximum overall.

For the resolution-dependent quotient of the sum of observed and modelled intensities (DRKPlot), the Ag IμS shows a significant underdetermination in the innermost shells of the Hirshfeld atom refinement. The effect is less pronounced in the multipolar refinement but still present. The MetalJet / Photon III data suffers from a similar effect, albeit to a lesser degree. A possible explanation would be the larger number of lower exposure time reflections in the Photon III datasets. While these needed to be included to supplant overexposed reflections on the Photon III dataset, the overall higher dynamic range of the Eiger2 CdTe detector made the inclusion of these reflections unnecessary.

The overall difference in performance can also be illustrated by the difference electron density of the refinements (Figure 5.5). The region of high difference is mostly concentrated around the sulphur region. The difference electron density is low for all refinements. However, the already low difference electron density of the Photon III datasets is significantly improved for the refinements against the data obtained with the Eiger2 CdTe. This is consistent with the difference in the resolution dependence seen in the DRKPlots but is also a very promising example of the achievable performance in charge density refinements of the MetalJet setup.

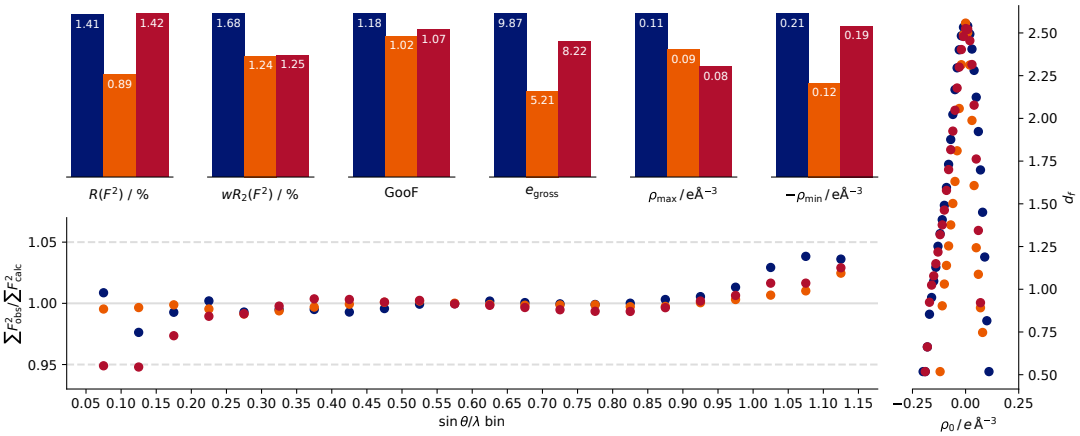


Figure 5.3.: Evaluated quality indicators for the PAW-HAR refinement of **5**. Displayed are numerical quality indicators as defined in Section 2.9.2 (top), the Henn-Meindl plot (right) and the DRKplot (bottom). Evaluated setups are MetalJet D2 – Photon III (blue), MetalJet D2 – Eiger2 CdTe (orange) and IμS 3.0 Ag – Photon III (red).

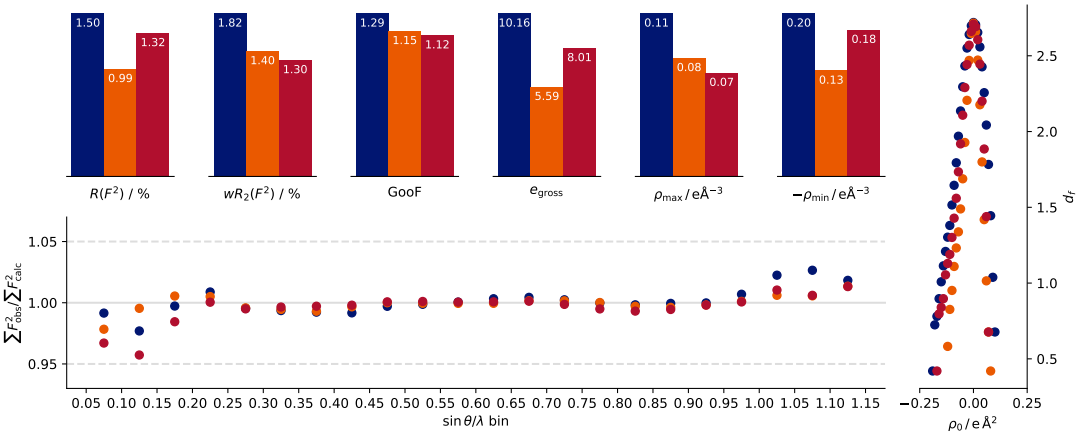


Figure 5.4.: Evaluated quality indicators for the multipolar refinement of **5**. Displayed are numerical quality indicators as defined in Section 2.9.2 (top), the Henn-Meindl plot (right) and the DRKplot (bottom). Evaluated setups are MetalJet D2 – Photon III (blue), MetalJet D2 – Eiger2 CdTe (orange) and IμS 3.0 Ag – Photon III (red).

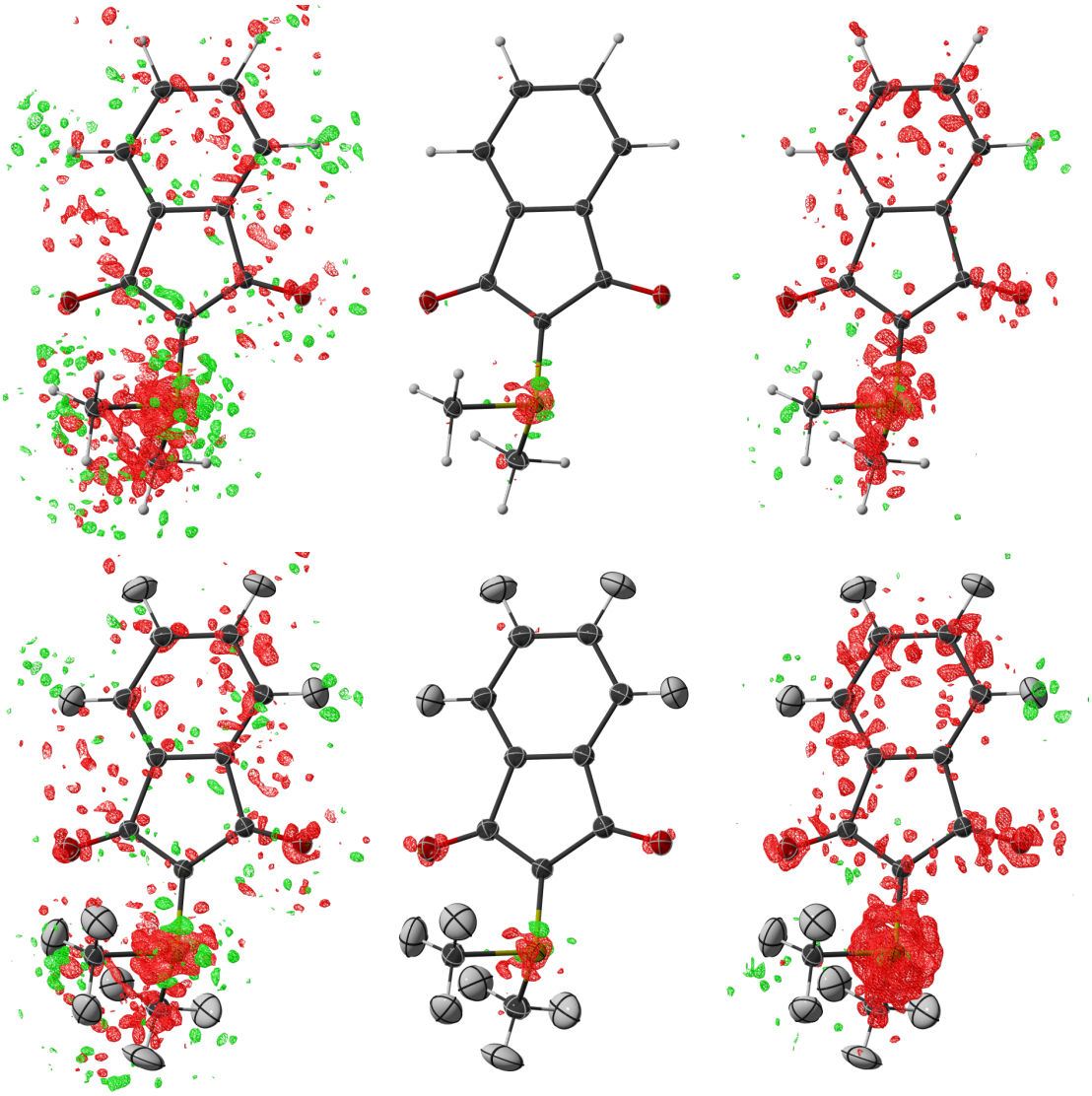


Figure 5.5.: Difference electron density (red: $-0.05\,e\,\text{Å}^{-3}$, green: $0.05\,e\,\text{Å}^{-3}$) of the multipole refinement (top row) and HAR (bottom row) of **5** against the data obtained from the MetalJet D2 – Photon III (left), MetalJet D2 – Eiger2 CdTe (centre) and IμS 3.0 Ag – Photon III (right). Atomic displacement parameters are depicted at the 50 % probability level.

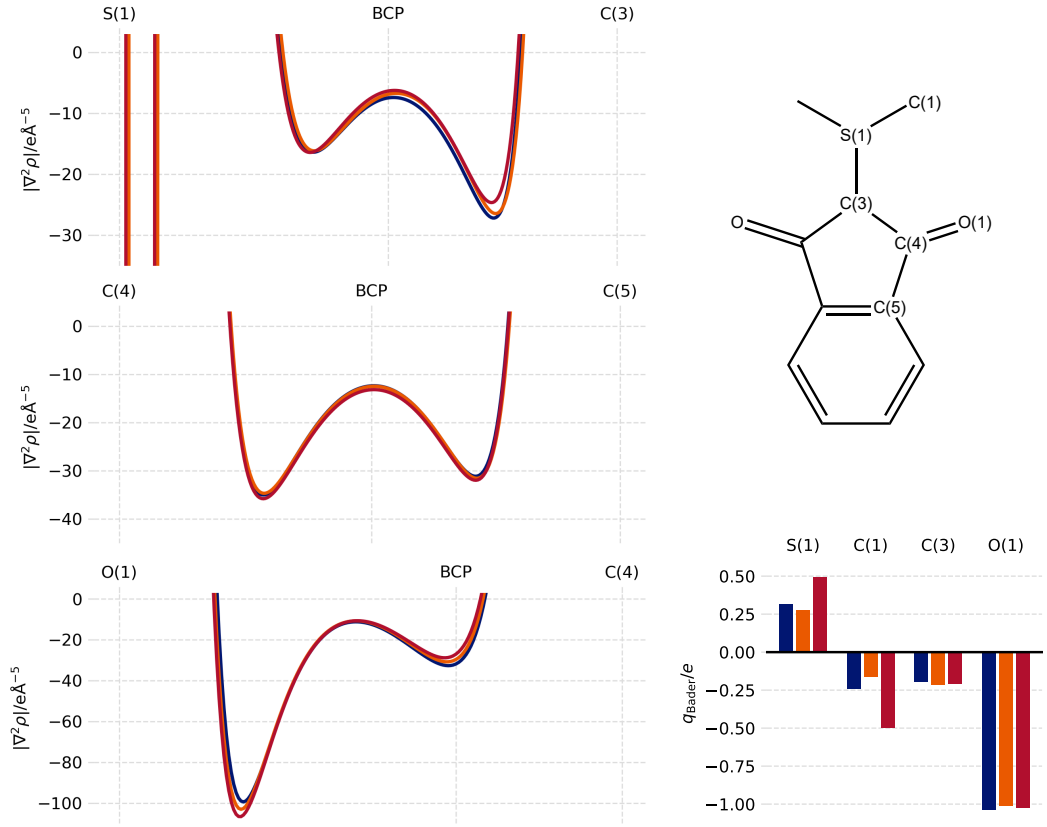


Figure 5.6.: Selected QTAIM indicators determined from the experimental charge density analysis of **5**. Laplacians along the bond critical points for three bonds (left), integrated Bader charges (lower right) and numbering of the displayed atoms (top right). Evaluated setups are MetalJet D2 – Photon III (blue), MetalJet D2 – Eiger2 CdTe (orange) and IμS 3.0 Ag – Photon III (red).

5.7.2. Comparison of QTAIM properties derived from the multipole refinements

The aim of a traditional charge density refinement using the multipolar model is to analyse the derived density to extract meaningful insight. Usually, this is done in the framework of the *Quantum Theory of atoms in molecules* (QTAIM) by Bader (1985). Selected properties are displayed in Figure 5.6.

The Laplacians along the three selected bond paths are well retained irrespective of the employed setup and accordingly, the Laplacians at the bond critical points are basically indistinguishable. There is a difference in the integrated Bader charges, where the sulphur atom and the connected methyl carbon atom show a difference of 0.25 e between the datasets collected using the MetalJet source and the Ag IμS. The charges of the

remaining carbon atoms as well as the oxygen atom are well retained between all three measurements. The difference in the sulphur might be connected to the comparatively poorer performance of this dataset in the DRKPlot, but is surprising when considering that the MetalJet – Eiger2 CdTe showed the lowest levels of difference electron density around the sulphur atom, while the MetalJet – Photon III refinement showed the highest levels at this site.

5.8. Conclusion

The MetalJet setup in Göttingen has been shown to work well with both independent atom refinements and two investigated refinement methods using aspheric atomic form factors.

In direct comparison, the Eiger2 CdTe offers a set of properties uniquely well tailored to the challenges of the specific setup as well as the designated application of the MetalJet in Göttingen. The ability to eliminate the low-energy gallium radiation by energy cutoff makes the use of attenuation unnecessary. Together with the very high photon efficiency of the cadmium telluride in the energy region of indium K_α radiation, this results in a higher effective intensity of the source, both at the sample and by detecting a larger number of the diffracted photons. The only possible downside of the lacking attenuator would affect samples susceptible to radiation damage by gallium K_α radiation specifically. The detector's advantages result in superior performance in both the independent atom model refinement of highly absorbing compounds, as well as the Hirshfeld atom model and multipole model refinements of the YLIDE crystal.

In these aspherical refinements, superior performance of the combination of the MetalJet D2 with the Eiger2 CdTe 1M compared to an Incoatec Ag IμS 3.0 with a Photon III detector was established. Unfortunately, the set of data does not allow the separation of the influence of the source and the detector, which would be of general interest but is irrelevant for the operation of the machine in Göttingen itself. The high dynamic range of the Eiger2 CdTe is advantageous for charge density refinements, where the strong inner data needs to be collected with as high accuracy as possible.

The use of indium K_α enables a higher resolution limit with its lower radiation wavelength, given that the specific sample diffracts up to this high resolution. Inorganic compounds without organic groups are a logical application of the MetalJet technology. Firstly, the unit cells are smaller on average and here the lower wavelength promises a faster experiment as a larger volume of the reciprocal space is mapped within a given 2θ range. Secondly, problems usually associated with these well-diffracting compounds such as absorption and extinction are also mediated, making it a well-tailored machine for the study of strong ionic interactions.

In future the update of the high voltage generator from 70 to 160 kV might further increase the intensity the machine can provide for such experiments.

6. Tapping of PAW-DFT for use in Hirshfeld atom refinement

As introduced in Section 2.5.2, Jayatilaka and Dittrich (2008) proposed the use of Hirshfeld atoms for crystallographic refinement. In a second publication, the method was extended to the iterative approach (Capelli *et al.*, 2014) currently known as HAR: If the atomic positions resulting from the refinement deviate from the atomic positions used for the calculation of the density, the density needs to be recalculated to be truly representative of the structure within the crystal.

The greatest benefit compared to the ubiquitous Independent Atom Model (IAM) was the possibility to obtain X−H bond lengths in much closer agreement with the values derived from neutron diffraction. This was shown in a number of studies for elements hydrogen is commonly bonded to in organic molecules. Comparison studies were done using densities calculated using Hartree-Fock calculations (Fugel *et al.*, 2018) or DFT (Woińska *et al.*, 2014; Sanjuan-Szklarz *et al.*, 2020).

The choice of the theoretical method for the evaluation of the density of the selected fragment affects the agreement factor with measured intensities. More important from a practical point of view is the significantly higher agreement of hydrogen positions and Atomic Displacement Parameters (ADPs) obtained by the HAR to atomic parameters obtained by refinement against neutron diffraction data. Obviously, the quantum chemical method of choice also affects the required computation time. Subsequently, progress can be made in two directions. Either one can go to a method enabling results in shorter computations times, by applying previous knowledge or approximations or one can go to more accurate methods and pay the computational price for more accuracy.

An example of the first direction would be the application of Extremely Localised Molecular Orbitals (ELMOs) in HAR, which were shown to yield accurate results at low computation times (Malaspina *et al.*, 2019). More recently fragmentation approaches have been used for a significant reduction in the computation time of large molecules, for example as reported by Chodkiewicz *et al.* (2022).

Higher accuracy quantum chemical methods have also been investigated with the application of MP2 and CCSD (Wieduwilt *et al.*, 2020). While a benefit for the crystallographic agreement factors could be shown, no systematic improvement of the X−H bond lengths of L-Alanine over DFT methods could be observed. As such the authors have provided a strong indication to search for further improvements by improving the description of the

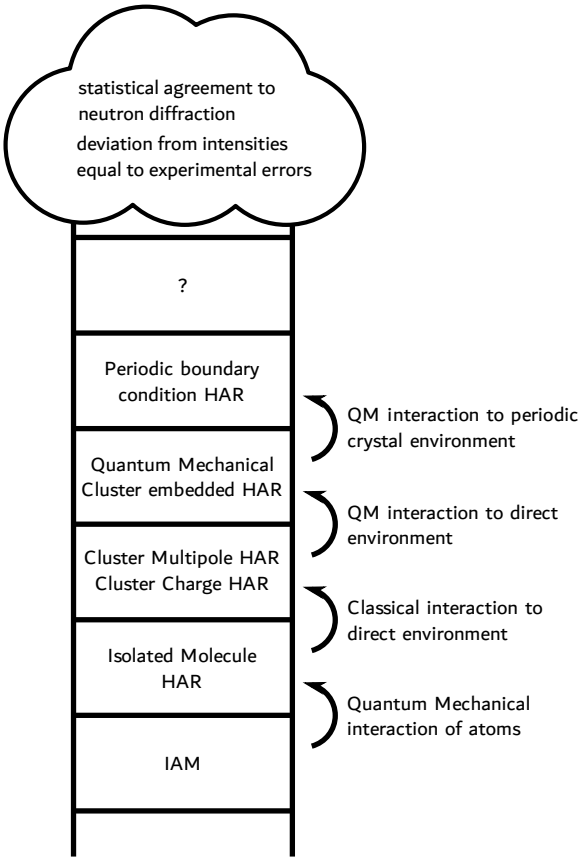


Figure 6.1.: Hierarchy of HAR methods in the form of a Jacob's Ladder. Adapted with modifications from Wieduwilt *et al.* (2021).

interaction of the investigated fragment with its surrounding instead of improving the level of quantum mechanical method used for the description of the fragment itself.

As such, Wieduwilt *et al.* (2021) proposed arranging the different descriptions of the crystal environment in the form of a Jacob's ladder as depicted[1] in Figure 6.1.

[1] In the originally published version the authors added the multipole model to the second rung of the ladder, together with the HAR using densities derived from isolated molecules. I disagree with the placement on this rung or on the ladder at all. The multipole model is a way of describing aspherical density. As such it contains all influences included in the object, which is described by the multipolar model. If the multipolar model is refined against theoretical calculations, as done in several of the databank approaches, the placement would be correct. However, if fitted against experimental X-ray diffraction intensities the multipolar model includes all influences the density of the physical system experiences. So assuming that the procedure is applied correctly, the Multipole model would be at the top. Accordingly, the multipolar model should not be on the Jacob's ladder in the same way that the Linear Combination of Atomic Orbitals (LCAO) or plane-wave basis sets used in the calculation do not belong onto the ladder. With exception of this caveat,the ladder is a highly useful tool to conceptualise the different environment descriptions.

Jumping from the second rung of the ladder (isolated molecule) to the third (cluster charges) was already established in the original publication of (non-iterative) HAR by Jayatilaka and Dittrich (2008) in the form of cluster charges surrounding the investigated fragment. The more involved form of the description of the surrounding in the form of multipoles was first established in the form of dipoles (Dittrich and Jayatilaka, 2012; Dittrich *et al.*, 2012), but later the application of higher order multipoles was also successfully demonstrated (Chodkiewicz *et al.*, 2020).

The next rung is formed by embedding the selected fragment into a cluster of quantum mechanically interacting molecules. Calculating the surrounding cluster at the same level of theory has been applied in the publication of Chodkiewicz *et al.* (2020). A more efficient approach is the embedding into other molecules described with an ELMOs basis set as reported by Wieduwilt *et al.* (2021).

All of the previously presented methods involve the same choice at the beginning: A representative fragment of the crystal needs to be chosen which can be either the asymmetric unit or constructed from the asymmetric unit and the symmetry elements of the unit cell. In a second step, the scientist would then decide if and how the surrounding the fragment experiences in the crystal is approximated. Often different rungs are combined into cluster shells of increasingly simple interactions. In the following sections, a new approach will be presented and contrasted with the methodologies presented so far. Namely, I will introduce the calculation of atomic densities for unit cells in periodic boundary conditions. In addition to removing the ambiguity of defining what a representative fragment for a specific structure actually would be, this method also enables the application of HAR to structures where a sensible definition of a fragment is not possible at all. This approach has only been demonstrated for non-iterative HAR of the urea molecule in a pilot study by Wall (2016). This work will demonstrate the systematic application in iterative HAR. For a discussion of additional differences between the two approaches see Section 6.8.

Please note that the underlying assumption is still that the electronic density of the vibrating system can be represented by atomic densities derived from the ground state density, which are convoluted with a probability function for the change of atomic position (see Section 2.4). It is known for Car-Parinello theory, that the band gap or HOMO-LUMO gap is a measure of how tightly the electrons are bound to the atomic positions (Pastore *et al.*, 1991) and this in turn could affect how well the approximation of vibrational invariance is fulfilled. All of the structure presented in this chapter have a large HOMO-LUMO gap.

6.1. Projector augmented waves

In chemistry, most quantum chemical and therefore also most HAR calculations are performed with the highly successful LCAO approach, where the molecular electronic wave function is represented as a linear combination of atom-centred functions, which are

usually Gaussians for performance reasons. These are provided in the form of well-defined basis sets, such as the correlation-consistent basis set family by Dunning (1989) or the Karlsruhe basis sets (Schäfer *et al.*, 1992).

In contrast, evaluations in the solid-state physics community are largely dominated by non-atom-centred bases for the wave function description, which were developed from the eigenfunctions of a free electron gas. The introduction of pseudo potentials removes the occurrence of rapidly oscillating wave functions near the atomic positions, thereby enabling the use of significantly smaller bases, be it in the form of plane waves, which represent a reciprocal space description of the wave function or the more recently introduced use of real-space grids, which divide the unit cell into voxels with a specific value for each position for each wave-function. The introduction of super soft pseudo potentials has further sped up convergence times, thus enabling the treatment of larger systems or larger bases for a given system.

The pseudopotential-based methods cannot adequately describe properties which require the interaction with core electrons, such as in the calculation of chemical shifts for Nuclear Magnetic Resonance (NMR) spectroscopy. It also prohibits the use for HAR as X-ray radiation is diffracted from all electrons, including the core electrons and therefore an all-electron density is required. Projector methods such as LAPW or the LMTO method (Andersen, 1975) were developed to achieve an all-electron approach while still relying on plane waves for most of the wave function description.

A very efficient solution to this problem was presented by Blöchl (1994) with the introduction of the Projector Augmented Wave (PAW) method. It combines the versatility of the LAPW approach with the simplicity of the use of plane wave pseudopotentials. The method is an all-electron method and as such can provide us with the all-electron density needed for calculating the atomic form factors. The explanation here will be limited to the wave function and density description. I will mainly follow the notation and completely follow the derivation by Blöchl (2022).

With the pseudopotential method, the PAW approach shares the presence of an auxiliary wave function $|\tilde{\psi}_n\rangle$. However, the all-electron wave function can be recovered by use of a transformation operator $\hat{\mathcal{T}}$:

$$|\psi_n\rangle = \hat{\mathcal{T}}|\tilde{\psi}_n\rangle \tag{6.1}$$

In the atomic regions, this operator has to modify the wave function to include the correct nodal structure. As such, the transformation includes the identity of the wave function plus a sum of atomic contributions $\hat{\mathcal{S}}_R$:

$$\hat{\mathcal{T}} = \hat{1} + \sum_R \hat{\mathcal{S}}_R \tag{6.2}$$

These atomic contributions add the difference between the true wave function and $|\psi_n\rangle$ and auxiliary wave function $|\tilde{\psi}_n\rangle$ for every atom. The atomic transformation operators are local, meaning they only affect the wave function near the atomic position. They are defined in terms of the solution of the Schrödinger equation for the isolated atoms $|\phi_i\rangle$. Additionally, core wave functions do not spread out to other atoms within the structure. As such a frozen core approach can be employed, where the core density and energy can be added separately, as long as the transformation $\hat{\mathcal{T}}$ only produces wave functions orthogonal to the core electrons, which it does by construction. Near the nuclei the wave functions are represented as a superposition of the partial waves:

$$\psi(\mathbf{r}) = \sum_{i \in R} c_i \phi_i(\mathbf{r}) \text{ for } |\mathbf{r} - \mathbf{R}_R| < r_{\mathrm{c,R}} \tag{6.3}$$

Here $i \in R$ is used for indicating that the partial wave with index i belongs to the atom with index R. For defining our atomic contribution to the projector function $\hat{\mathcal{S}}_R$ an auxiliary partial wave $|\tilde{\phi}_i\rangle$ needs to be chosen using the identity:

$$|\phi_i\rangle = \left(\hat{1} + \hat{\mathcal{S}}_R\right) |\tilde{\phi}_i\rangle \tag{6.4}$$

$$\hat{\mathcal{S}}_R |\tilde{\phi}_i\rangle = |\phi_i\rangle - |\tilde{\phi}_i\rangle \tag{6.5}$$

The atomic contribution should be different from one only near the atomic position of the respective atom with index R, *i. e.* the partial wave and its auxiliary counterpart should be identical beyond a certain defined radius $r_{\mathrm{c},R}$. To obtain the projector functions $|\tilde{p}_i\rangle$, the auxiliary wave function is expanded locally onto the auxiliary partial waves.

$$\tilde{\psi}(\mathbf{r}) = \sum_{i \in R} c_i \tilde{\phi}_i(\mathbf{r}) = \sum_{i \in R} \tilde{\phi}_i(\mathbf{r}) \langle \tilde{p}_i | \tilde{\psi} \rangle \text{ for } |\mathbf{r} - \mathbf{R}_R| < r_{\mathrm{c,R}} \tag{6.6}$$

Equations 6.5 and 6.6 can now be combined to apply $\hat{\mathcal{S}}_R$ to any auxiliary wave function.

$$\hat{\mathcal{S}}_R |\tilde{\psi}\rangle = \sum_{i \in R} \hat{\mathcal{S}}_R |\tilde{\phi}_i\rangle \langle \tilde{p}_i | \tilde{\psi} \rangle = \sum_{i \in R} \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle \tilde{p}_i | \tilde{\psi} \rangle \tag{6.7}$$

and the transformation operator is defined as:

$$\hat{\mathcal{T}} = \hat{1} + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle \tilde{p}_i | \tag{6.8}$$

where the index i now runs over all partial waves of all atoms at their respective positions. Finally, the complete all-electron wave function can be expressed as:

$$|\psi\rangle = |\tilde{\psi}\rangle + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle\right) \langle \tilde{p}_i|\tilde{\psi}\rangle = |\tilde{\psi}\rangle + \sum_R \left(|\psi^1_R\rangle - |\tilde{\psi}^1_R\rangle\right) \tag{6.9}$$

with the atomic wave functions and auxiliary wave functions given by:

$$|\psi^1_R\rangle = \sum_{i\in R} |\phi_i\rangle\langle\tilde{p}_i|\tilde{\psi}\rangle \tag{6.10}$$

$$|\tilde{\psi}^1_R\rangle = \sum_{i\in R} |\tilde{\phi}_i\rangle\langle\tilde{p}_i|\tilde{\psi}\rangle \tag{6.11}$$

What remains for the purpose of this work, is the expression for the density. Following Equation 6.9 it can be expressed as:

$$\rho(\mathbf{r}) = \tilde{\rho}(\mathbf{r}) + \sum_R \left(\rho^1_R(\mathbf{r}) - \tilde{\rho}^1_R(\mathbf{r})\right) \tag{6.12}$$

where the auxiliary density $\tilde{\rho}(\mathbf{r})$ can be calculated as the sum of all n wave functions with the respective occupation f_n:

$$\tilde{\rho}(\mathbf{r}) = \sum_n f_n \tilde{\psi}^*_n(\mathbf{r})\tilde{\psi}_n(\mathbf{r}) + \tilde{\rho}_\mathrm{c}(\mathbf{r}) \tag{6.13}$$

The atomic density and auxiliary density can be calculated with:

$$\rho^1_R(\mathbf{r}) = \sum_{i,j\in R} D_{i,j}\phi^*_j(\mathbf{r})\phi_i(\mathbf{r}) + \rho_{\mathrm{c},R}(\mathbf{r}) \tag{6.14}$$

$$\tilde{\rho}^1_R(\mathbf{r}) = \sum_{i,j\in R} D_{i,j}\tilde{\phi}^*_j(\mathbf{r})\tilde{\phi}_i(\mathbf{r}) + \tilde{\rho}_{\mathrm{c},R}(\mathbf{r}) \tag{6.15}$$

where $\rho_{\mathrm{c},R}(\mathbf{r})$ is the core density of the corresponding atom, $\tilde{\rho}_{\mathrm{c},R}(\mathbf{r})$ is the auxiliary core density. It is identical to the core density outside of the atom region, but smooth inside. Finally, $D_{i,j}$ are the elements of the one centre density matrix, which is defined as:

$$D_{i,j} = \sum_n \langle\tilde{p}_i|\tilde{\psi}_n\rangle f_n \langle\tilde{\psi}_n|\tilde{p}_j\rangle \tag{6.16}$$

As such, all the elements of the all-electron ground-state density of the periodic calculation are present. This in turn enables the use as the basis for a periodic approach to HAR.

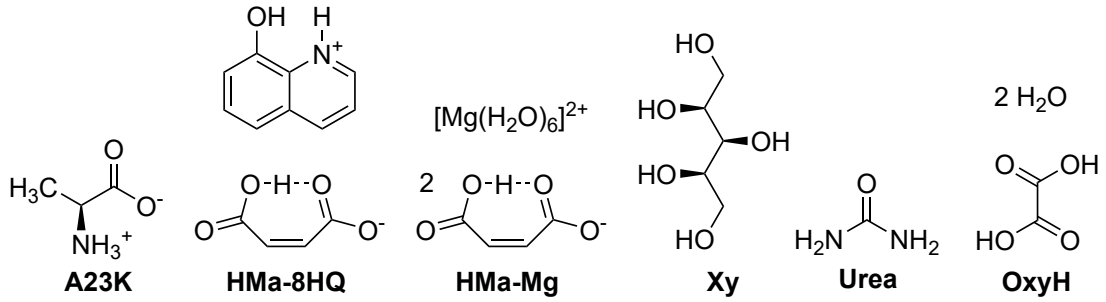


Figure 6.2.: Structures included in the investigation of this chapter.

6.2. Evaluation details

Before application comes validation, which is therefore the main focus of this work in its quest to establish a new method. Accordingly, the methodology should be applied to established high-quality datasets. For most of these, high-quality HARs are already available, which enables the comparison without the possibility to choose options which are beneficial to our own method. However, as these refinements have been done using a plethora of different options, consistent combinations of description of the surrounding, density functional, basis set and program for calculation, self-performed non-periodic HARs will still be presented in addition to the reference descriptions obtained from literature.

6.2.1. Datasets used for evaluation

A variety of datasets was chosen for the validation of the new methodology. While a good crystallographic agreement factor is one factor in the evaluation, error compensation might lead to erroneous conclusions. One factor to reduce this risk is a suitable high resolution. This work included only datasets with a resolution of at least 0.6 Å.

At the same time, the stated benefit of HAR is the closer agreement of X−H bond lengths and the possible description of atomic displacement of hydrogen atoms. The success in this manner can only be evaluated against neutron diffraction experiments which represent the gold standard for the determination of positions and atomic displacement. Therefore the presence of a neutron diffraction measurement is paramount for validation. A good performance in this regard then informs confidence in the application even when neutron data is not available.

The depictions of the structures included in the investigation can be found in Figure 6.2. Information on the individual structures will be provided in the following:

L-Alanine at 23K (A23K)

The number of atoms in the structure L-Alanine is relatively small with one complete molecule with thirteen atoms in the asymmetric unit and four molecules in the unit cell. The heaviest element is oxygen, which also limits the number of electrons, which need to be included in the calculation. As such it is employed for the evaluation of options, as computational times are still affordable even with a larger relative basis per atom. Additionally, there are no atoms on special positions. Therefore the maximum number of atomic parameters for each atom are to be used for this structure. The original high-resolution X-ray data at 23 K up was measured by Destro *et al.* (1988), while the corresponding neutron dataset was measured more recently by Malaspina *et al.* (2019).

Urea at 123K (Urea)

The second dataset suitable for comparatively cheap comparisons stems from a crystal containing urea. In this structure all of the atoms are located on special positions and as such the performance of XHARPy on this problem can be evaluated. There is only half a molecule in the asymmetric and only two molecules with eight atoms each are located in the unit cell. Again the heaviest element is oxygen, thereby limiting the number of needed basis functions. The original high-resolution X-ray dataset was measured by Birkedal *et al.* (2004), while the neutron dataset is older (Swaminathan *et al.*, 1984).

Hydrogen maleate systems (HMa-Mg, HMa-8HQ)

These two structures contain the hydrogen maleate anion paired with different cations. While the original investigation includes more cations, this investigation focuses on the two cations which have been treated with the highest level of HAR in a previous investigation, namely hexaquamagnesium and 8-hydroxyquinolinium. The X-ray data is more recent (Malaspina *et al.*, 2020) than the neutron dataset (Malaspina *et al.*, 2017).

Xylitol (Xy)

This high-resolution dataset shows several hydrogen bonds and is well-established in the charge density community. Additionally, the original authors (Madsen *et al.*, 2004) also refined extinction, while a later X-ray restrained wave function fitting investigation assigned the effects of the intensity-dependent differences to the density instead (Malaspina *et al.*, 2021). Attributing the effect to extinction effects seems more reasonable and as such this system can be used to test the treatment of extinction effects in our implementation. The corresponding neutron dataset was published by Madsen *et al.* (2003).

Oxalic acid dihydrate (OxyH)

All other investigated datasets supply one measurement per structure. In contrast, this dataset by Kaminski *et al.* (2014) provides thirteen datasets measured on thirteen crystals employing three different diffractometers. Therefore this dataset allows testing the robustness of benchmarking with PAW-HAR in dependence on the individual measurements. A neutron measurement for comparison was provided in the same publication.

6.3. Application to Hirshfeld atom refinement

The resulting density of PAW calculations is usually expanded on a rectangular grid. Fourier Transformation in this basis can efficiently be done using the Fast Fourier Transform (FFT) algorithm (Gauß, 1866; Cooley and Tukey, 1965). However, there is the potential for a similar challenge as described in Section 6.1: The expansion of the localised core density requires a much denser grid than the less localised valence density.

As such, it was decided to follow a similar route, namely, to evaluate the core density separately. Most often PAW setups and software use a frozen core approximation. This core is defined on a core-centred spherically symmetric grid which is evaluated on points using a logarithmic scale. This core density was assigned completely to the respective atom and as such excluded from the density to be partitioned in the Hirshfeld partitioning scheme and from the atomic densities used for the Hirshfeld weights.

In general, the generation of the projector functions for the PAW method does include the calculation of atomic densities. As such these atomic densities are also used for calculating the Hirshfeld weights. In an ideal setting, this means, that the atomic and molecular densities are calculated using the same functional. However, the implementations of projector function generation only support density functionals using the generalised gradient approximation. Therefore, the projector functions for meta-GGA functionals or hybrid functionals are generated with PBE. Additionally, the libraries of Quantum Espresso compatible setup files are mainly available for PBE or PBEsol. As such PBE calculated setup files were used for all calculations in Quantum Espresso.

In general, there is a number of different codes for evaluating structures in the PAW formalism such as GPAW (Enkovaara *et al.*, 2010), Quantum Espresso (Giannozzi *et al.*, 2017), CP-PAW and ABINIT (Gonze *et al.*, 2020), of which only the first two were implemented and evaluated.

Starting from loaded X-ray intensities, including estimated standard deviations and atomic parameters (x, y, z, U_{ij}) the following algorithm is employed for the actual refinement:

Algorithm Basic Approach to PAW-HAR

```
for atomtype in structure do
  Load distance grid of PAW setup
  Load distance values of PAW setup
  for h in measured intensities do
    Calculate the frozen core structure factor for the atom by numerical Fourier-Bessel-
    Transform
  end for
end for
while New atomic positions not equal to old positions do
  Do a PAW Calculation of the density
  for atom in asymmetric unit do
    Calculate the Hirshfeld density
    Calculate the atomic form factor by FFT
    for h in measured intensities do
      for symmetry s in symmetry elements do
        Get the corresponding atomic form factor value for R_s h for symmetry
        equivalent atoms
      end for
    end for
  end for
  Do a crystallographic refinement using the calculated form factors
end while
Calculate the variance covariance matrix
```

The basic foundations of the XHARPy[2] package were written in 2020 before the publication of NoSpherA2 by Kleemiss *et al.* (2021). As such, I wrote a complete refinement library in PYTHON, by leveraging the just-in-time compiling and automatic gradient generation provided by the JAX (Bradbury *et al.*, 2018) package in combination with NumPy (Harris *et al.*, 2020) and the implementation of the BFGS algorithm (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970) in SciPy (Virtanen *et al.*, 2020).

The code for the package is currently available under `https://github.com/Niolon/XHARPy`.

In the first step, the dependence of the performance of PAW-HAR on grid-spacings and the functional will be investigated in GPAW, which was implemented first. Subsequently, this work will compare the performance to already established approaches, recapping the investigations presented in Ruth *et al.* (2022). The employed quality indicators are $wR_2(F^2)$ as defined in Equation 2.75, $|\Delta r|$ as defined in Equation 2.81, $|\Delta U_{ij}|$ as defined in Equation 2.84 and S_{12} as defined in Equation 2.85.

6.4. Dependence of the performance on the functional

The first investigation will focus on the performance of different functionals within the PAW-HAR method. Investigations have been conducted using GPAW with a number of functionals available within that software package, which unfortunately excluded the possibility to use hybrid functionals in periodic calculations. The resulting quality indicators are depicted in Figure 6.3 for **HMa**–**Mg**. Tables with numerical values for all datasets can be found in the appendix in Section B.5. An explanation of the employed box-whisker plots can be found in Section 2.9.4.

The difference in the performance of the crystallographic agreement as measured by the $wR_2(F^2)$ is immediately obvious. Unsurprisingly, the PW functional using only the LDA approximation shows a poorer agreement compared to the functionals of the GGA family. Within the GGA family itself, an improvement can be seen for the newer PBE functionals compared to the older BLYP and PW91 functionals. Moving to the group using the mGGA approximations TPSS shows a worse performance compared to the PBE family. The two investigated functionals from the SCAN family are again a further improvement from the GGA functionals. The functionals including van-der-Waals interactions yield disappointing results. They perform poorer compared to functionals from the PBE and SCAN families with vdW-df2 yielding the poorest performance in the complete set.

[2]The name is an acronym for 'X-ray diffraction data Hirshfeld atom refinement in PYTHON.' I agree that it falls into the name-follows-acronym category of naming packages (instead of acronym-follows-name), but hope the reader agrees that the stretch is limited as far as scientific software packages or methodologies are considered.

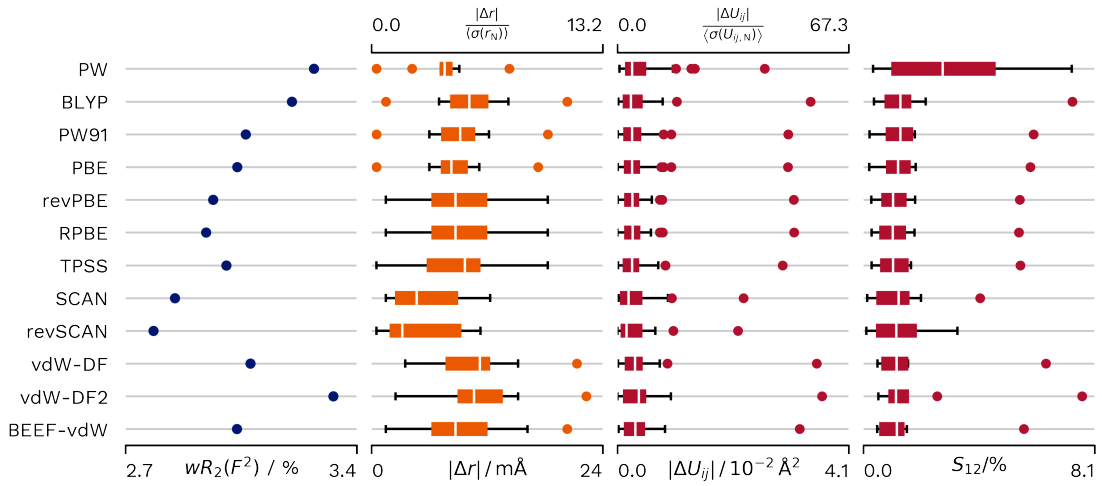


Figure 6.3.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **HMa**–**Mg** in dependence of the chosen functional.

In the agreement of the refined parameters to neutron values, PW does provide comparatively good distances, but at the same time very poor atomic displacement parameters. An improvement in the ADPs can be seen for the GGA functionals, while the distances do not profit for this dataset. This effect is unique for the **HMA**–**Mg** dataset, but BLYP usually performs worse than the PW functional overall. At the meta-GGA level, TPSS is again not an improvement but performs worse than the PBE family using a simpler approximation. However, the functionals from the SCAN family do yield the best displacement parameters and distances. Again, the performance of the van-der-Waals functionals is disappointing, which indicates that the approaches do not reflect the van-der-Waals interactions in the determined density, while they may do so in the determined energy.

An outlier is present in the determination of the atomic displacement parameters for all functionals. This corresponds to the U_{11} value of the hydrogen atom within the maleate molecule between the two oxygen atoms. The disagreement is located along the bond direction. A probable explanation lies in the fact that hydrogen atoms in such configurations are known to tunnel between the two possible locations. The atomic density should change shape completely with the change of position, where the bonding density is now located at the other side of the atomic position. This effect cannot be accurately represented within the chosen approximations.

In light of these results, the SCAN functional was chosen for further investigations and comparisons. The other option would have been the revSCAN functional, which performed similarly in the conducted investigations.

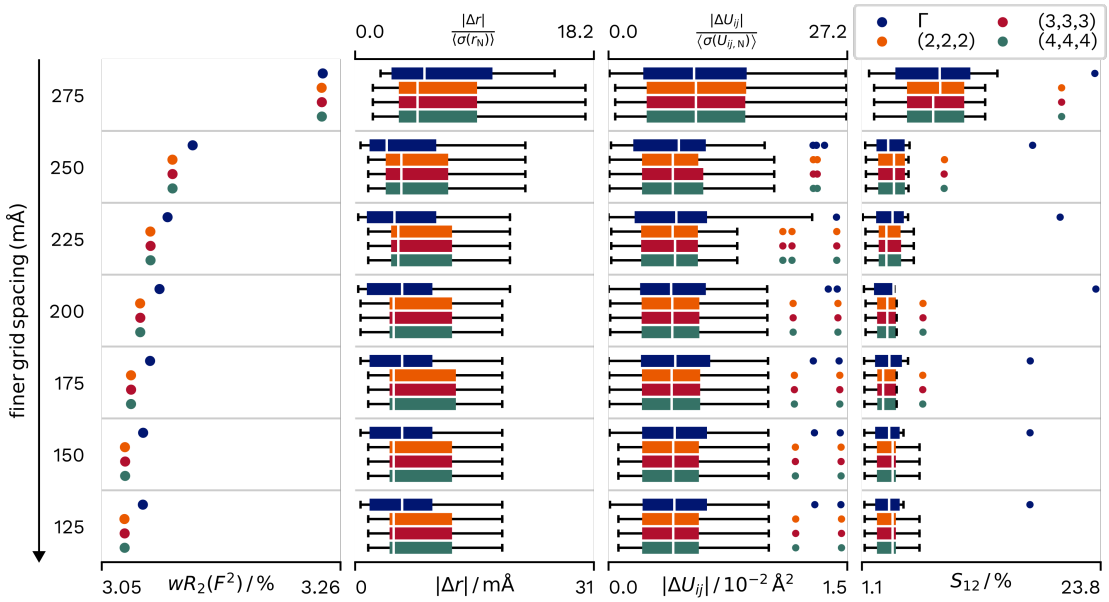


Figure 6.4.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **A23K** in dependence of the choice of the real-space grid for the wave function description and the chosen k-point grid using the SCAN functional. Real-space grids for the density for functional evaluation and density for the atomic form factor calculation are interpolated two and four times on each axis respectively. Colours indicate k-point sampling, with a k-point grid of a single point in the origin denominated with the common symbol Γ.

6.5. Dependence of the performance on real-space grid spacing and k-point grid

In this section, the influence of the two main grids that can be set in a GPAW calculation is to be evaluated. Computing a larger number of grid points improves the achieved results until convergence is reached. This section should now demonstrate the evaluation of convergence for the **A23K** dataset. While the determination of the required real-space grid size should be applicable to other compounds, especially with similar elements, the k-point grid should be evaluated for individual compounds. Therefore, this investigation should be seen as an example of the influence of the k-point grid on crystallographic parameters. The resulting quality indicators are depicted in Figure 6.4.

All quality indicators profit from a finer real-space grid spacing. Convergence for the $wR_2(F^2)$ is reached at 1.50 Å. Convergence in comparison to the neutron-derived distances and atomic displacement parameters is reached at 1.75 Å.

The picture is less obvious for convergence with the k-point grid. Both the $wR_2(F^2)$ and the agreement to neutron atomic displacement parameters profit from raising the number of k-points from the single Γ point to a (2,2,2) Monkhorst-Pack grid, which was centred on the origin. At the same time, the mean absolute difference in X−H bond distances increases from 0.007 to 0.009 Å. There is no additional difference if the k-point grid size is increased even further.

6.6. Comparison to established approaches

The individual reference refinements from other groups represent a climb of the ladder as depicted in Figure 6.1. In order to also compare to Hirshfeld atom refinements using a cohesive set of options, further Hirshfeld atom refinements without any consideration of the crystal environment were performed in Olex2 (Dolomanov *et al.*, 2009) using NoSpherA2 (Kleemiss *et al.*, 2021) and Orca (Neese, 2018). These will be listed in the following figures with the indicator '(None)'. In addition, tonto (Jayatilaka and Grimwood, 2003) in Olex2 was employed to perform HAR with a shell of 4 or 8 Å of cluster charges. The connected indicator is '(4/8 Å cc)'. All these investigations were performed on the B3LYP/def2-TZVPP level to ensure that the performance is not limited by a small basis set.

PAW-HAR calculations in this section were done using GPAW as the calculation program with a real-space grid as the basis for the pseudo-wave functions. The listed grid size for a given real-space grid spacing corresponds to the grid for the description of the wave functions. Density-dependent properties are calculated by GPAW from a once interpolated grid, *i. e.* with half the size of that one listed. The density grid used for atomic partitioning of the unit cell density was interpolated once more, thereby quadrupling the number of points in each dimension compared to the wave function. Calculated k-points were either limited to the Γ-point or selected using the scheme proposed by Monkhorst and Pack (1976).

6.6.1. L-Alanine at 23K

The comparison can be done with a CCSD calculation without consideration of the crystal environment (Wieduwilt *et al.*, 2020). To my knowledge, this published structure can be seen as the limit of what can be reached without consideration of the crystal environment. As such it answers the question, of whether the quality of the density is limited by the quality of the calculation method or the description of the crystal environment. The fragment-based HAR methods conducted for this work used the asymmetric unit as the evaluated fragment. The PAW-HAR calculations used a (2 2 2) k-point grid and real-space grid spacing of 0.15 Å.

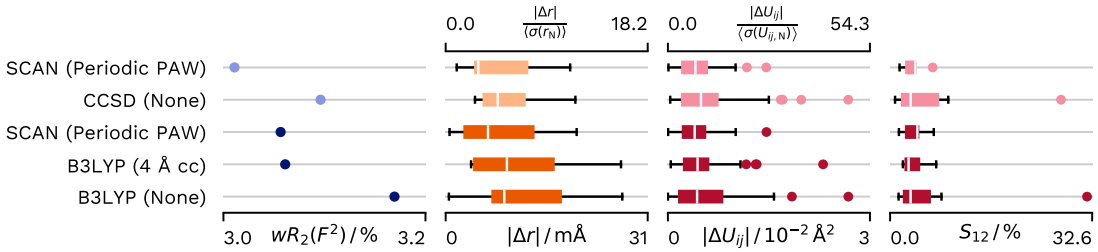


Figure 6.5.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **A23K**. CCSD (None) results are computed from the calculations of Wieduwilt *et al.* (2020). Lighter-coloured quality indicators are computed with an $F/\sigma(F)$ cutoff of 3, as applied in the reference.

The reference calculation used an $F/\sigma(F)$ cut-off. Accordingly, two PAW-HAR refinements were done. One used the employed cut-off while the other used the complete data.

All Hirshfeld atom refinements result in a featureless difference electron density. A summary of the performance can be found in Figure 6.5. The $wR_2(F^2)$ indicates that a consideration of the crystal environment is necessary. Both the refinement using cluster charges in TONTO as well as the calculation using the periodic PAW density basis show a significant improvement over the ORCA calculation without any consideration of the crystal surrounding, while the wR_2 is almost identical for the two methods using a model for the crystal environment. Against intuition, the increase of the cluster radius from 4 Å to 8 Å did lead to worse results, whereas one would normally expect convergence to a minimum with increasing cluster radius. Neither of the TONTO calculations converged in 20 HAR iterations.

The agreement to neutron values increases from cluster charge calculations to periodic PAW. A higher impact can be seen on the agreement in X−H distances, whereas the improvement in atomic displacement parameters is only small. A direct comparison of the reference CCSD calculation to the PAW-HAR calculations clearly shows the need for the crystal environment description for all investigated agreement factors.

6.6.2. 8-Hydroxyquinolinium hydrogen maleate and hexaquamagnesium hydrogen maleate at 15 K

The reference calculation (Malaspina *et al.*, 2021) used 8 Å of cluster charges and the B3PW91 functional for a refinement on the absolute of the structure factors (F) instead of the intensities with an $F/\sigma(F)$ cutoff of 4. For each structure, two PAW-HAR calculations were done. One against the data reduced to the same reflection as in the reference to enable a direct comparison and one additional refinement against

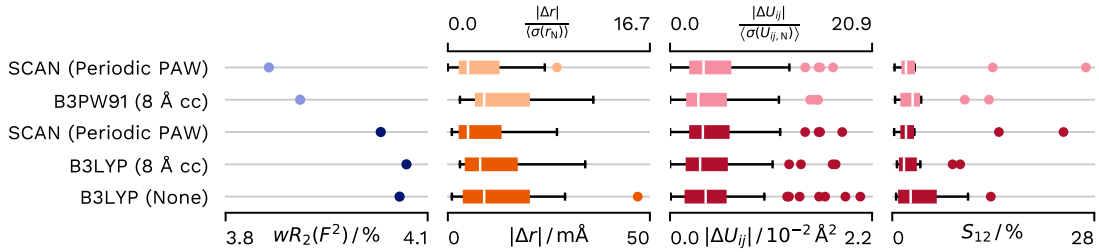


Figure 6.6.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **HMa-8HQ**. B3PW91 results are computed from the calculations of Malaspina *et al.* (2021). Lighter-coloured quality indicators are computed with an $F/\sigma(F)$ cutoff of 4, as applied in the reference.

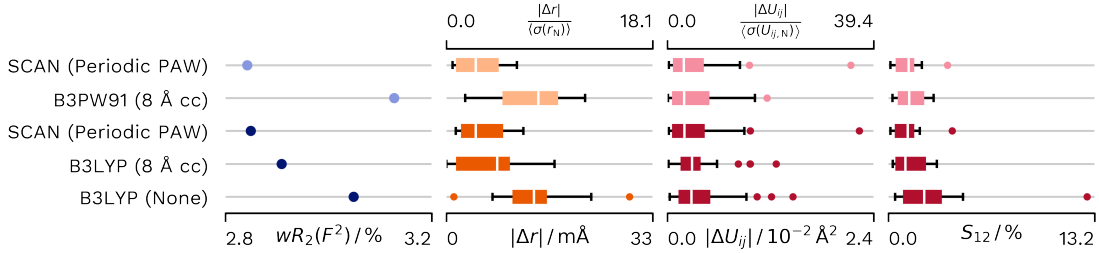


Figure 6.7.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **HMa-Mg**. B3PW91 results are computed from the calculations of Malaspina *et al.* (2021). Lighter-coloured quality indicators are computed with an $F/\sigma(F)$ cutoff of 4, as applied in the reference.

the full data. The fragment-based HAR methods conducted for this work used the asymmetric unit as the fragment for the hydroxyquinolinium structure. The fragment of the magnesium-based structure was generated by applying the inversion centre on the magnesium atom onto all atoms, thereby retaining the inversion symmetry of the structure within the fragment. The PAW-HAR calculations were done at the Γ-point with a real-space grid spacing of 0.15 Å.

The resulting summary of the performance can be found in Figures 6.6 and 6.7. Refinement against a periodic density shows obvious benefits for both the $wR_2(F^2)$ and the agreement in X−H bond distances. However, the reference calculation exhibits smaller outliers for both the ΔU_{ij} and S_{12} values.

This improvement in X−H bond distance agreement is also exhibited by the NoSpherA2/tonto cluster charge calculation compared to the reference, albeit to a smaller degree. This could be an indication that refinement against F^2 can show improved results to a refinement against F.

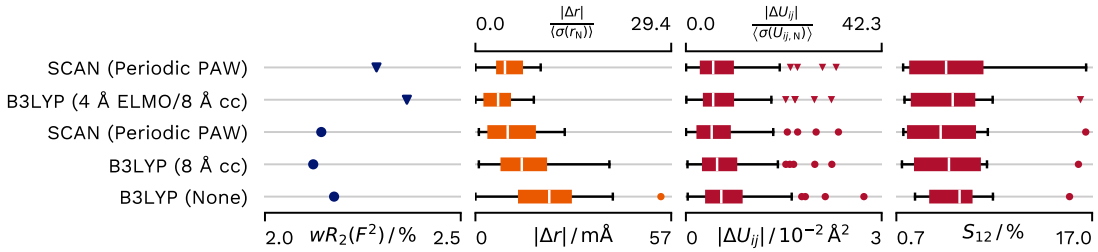


Figure 6.8.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **Xy**. B3LYP (4 Å ELMO/8 Å cc) results are computed from the calculations of Wieduwilt *et al.* (2021). Refinements marked with a triangle do not account for extinction effects.

The residual density is featureless in neither of the two datasets (For a depiction see Figures B.12 and B.13 in the appendix). This could point to an undescribed effect within the data. Nevertheless, PAW refinements still result in a superior agreement to both the data as well as the neutron X−H distances.

6.6.3. Xylitol

The reference description of the **Xy** dataset is located another rung higher on the Jacob's ladder (Figure 6.1) than the previous one. The investigation provided by Wieduwilt *et al.* (2021) embedded the structure calculated with the B3LYP functional into a 4 Å cluster where the molecules were described by ELMOs. This layer was surrounded by another layer with a 4 Å width consisting of cluster charges. The fragment-based HAR methods conducted for this work used the asymmetric unit as the evaluated fragment. The PAW-HAR calculation was conducted in GPAW, using a k-point grid of (3 3 3) and a real-space grid spacing of 0.175 Å.

Initial HAR results showed strong reflections, which were weaker than calculated by the employed model for all descriptions of the crystal environment. The behaviour is consistent with the effects of extinction and the original authors treated it as such (Madsen *et al.*, 2004). I would also interpret the observed behaviour as such.

In contrast, the reference (Wieduwilt *et al.*, 2021) seems to follow the route by Malaspina *et al.* (2021). The treatment of extinction does affect the results, but in order to be consistent with the approach in this chapter, comparison to reference is still necessary. Accordingly, two PAW-HAR calculations were performed. The first calculation includes a refinement of extinction using the scheme as implemented in SHELXL by Sheldrick (2015*a*) and should be considered the correct refinement. For comparison, an additional refinement without extinction has been performed. The results can be found in Figure 6.8.

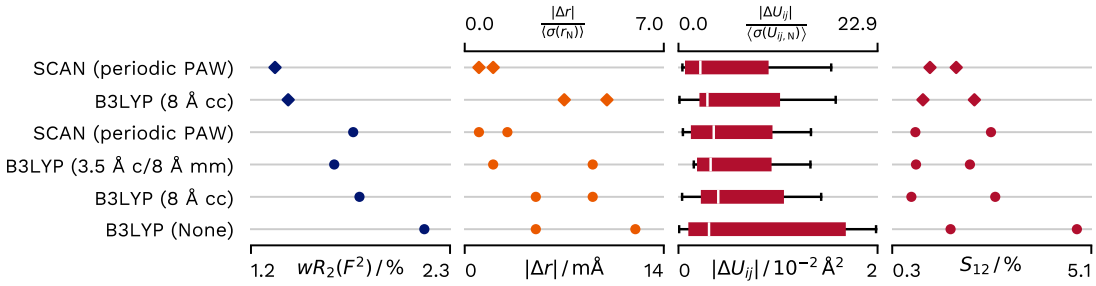


Figure 6.9.: Crystallographic agreement measured with the $wR_2(F^2)$, as well as the agreement to neutron values for **Urea**. B3LYP (3.5 Å c/8 Å mm) results are computed from the calculations of Chodkiewicz *et al.* (2020). Refinements marked with a diamond glyph used models including a Gram-Charlier expansion of the atomic displacement.

In comparison to the reference (and therefore in the calculation without extinction), the calculated wR_2 is slightly lower for the PAW-HAR calculation and the agreement in atomic displacement parameters is improved slightly. However, the agreement to X−H neutron bond lengths is lower compared to the reference. An investigation, into whether the relative performance holds with an extinction refinement still remains to be conducted. However, superior performance for the ELMO cluster on the X−H bond lengths seems likely.

As expected the crystallographic agreement factor profits from the refinement of an extinction parameter. In contrast to all other datasets, the cluster charge calculation shows an improvement in wR_2 over the PAW-HAR calculation. However, the agreement in distances and atomic displacements is still superior for PAW-HAR.

6.6.4. Urea at 123K

The reference refinement calculates the density by placing the molecule at the centre of a 3.5 Å cluster calculated at the same level of theory as the molecule itself, which is a B3LYP//cc-pVTZ description. Additionally, the cluster was surrounded by a cluster with an 8 Å width, which was described by classical point multipoles. The fragment-based HAR methods conducted for this work used a single completed molecule as the evaluated fragment. The PAW-HAR calculation was conducted in GPAW using a (3 3 3) k-point grid and a real-space grid spacing of 0.1 Å.

The resulting quality indicators are depicted in Figure 6.9. As can be seen, the wR_2 and the atomic displacement parameters of the PAW-HAR refinement show a slightly worse performance compared to the reference. However, the N−H bond distances are in agreement with the neutron values, exhibiting differences of only 0.001 and 0.003 Å, with the estimated standard deviation of the neutron value being 0.002 Å.

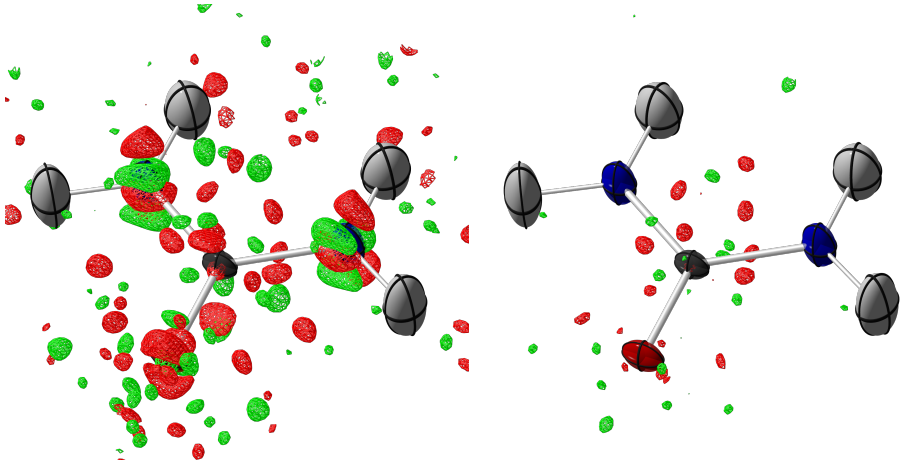


Figure 6.10.: Difference electron density (red: $-0.05\,e\,\text{Å}^{-3}$, green: $0.05\,e\,\text{Å}^{-3}$) for the PAW-HAR refinement of urea without Gram-Charlier parameters on the left and with Gram-Charlier parameters (N, O) on the right.

The resulting difference electron density (Figure 6.10) is on a low level but not entirely featureless. The features were successfully modelled using third and fourth-order Gram-Charlier parameters for urea. While this modelling resulted in a decrease in agreement to neutron values for the description of the surrounding with cluster charges using TONTO/NOSPHERA2, the PAW-HAR refinement resulted in a further convergence to the neutron values. A possible reason would be the error compensation of the missing Gram-Charlier and the less accurate density description in the cluster charge calculation, while the more accurate description using a periodic calculation does not suffer from this drawback.

The description with Gram-Charlier parameters needs to be verified. Three criteria need to be met for the additional refinement parameters to be permitted. The first criterion is the check for negative probability, resulting from the anharmonic refinement. A region of negative probability is impossible in the physical system and as such a refinement that gives this result is not allowed. The second criterion was formulated by Kuhs (1992), who formulated a necessary resolution of the data, depending on the harmonic vibrational parameters of the investigated atoms. Kuhs' rule allowed for the refinement of third-order parameters for the carbon atom and fourth-order refinement of nitrogen and oxygen with the given data with a maximum resolution of $\frac{\sin(\theta)}{\lambda} = 1.405\,\text{Å}^{-1}$. Both criteria were checked using XDPDF from the XD2016 suite (Volkov *et al.*, 2016). The final criterion for refinement is that Gram-Charlier parameters of the given order are significantly different from zero. As the third-order Gram-Charlier parameter refined for the carbon atom did not meet this criterion, Gram-Charlier parameters were only refined for the nitrogen and oxygen atoms of the urea molecule. As such, the procedure is consistent with the treatment of anharmonic vibration in the original publication by Birkedal *et al.* (2004).

6.6.5. Comparison of aggregated values

This section will aggregate the results from all the structures in order to draw more general conclusions. To this end, the hydrogen atoms were allocated into three groups based on whether they are located in hydrogen bonds or not. The applied criterion is the narrow definition of a classical hydrogen bond (X−H···Y) where X and Y are either nitrogen or oxygen atoms in the systems evaluated.

The first group consists of hydrogen atoms in classical hydrogen bonds, which were not cut in the fragment definition for the ORCA/TONTO calculations with NOSPHERA2. This includes the bridging hydrogen atom of the hydrogen maleate structures as well as the hydrogen bonds connecting the maleate anion with the cation. This means four bonds in total (intra X−H···Y).

The second group consists of hydrogen bonds, which were cut during the fragment definition. These include the N−H bonds of alanine. The remaining hydrogen bonds from the hydrogen maleate structures include most of the bonds in the water molecules, except for one, all O−H bonds in the xylitol molecule and the N−H bonds of the urea molecule. In total, this means 16 atoms/bonds are included (inter X−H···Y).

The third group contains hydrogen atoms, which are not in a hydrogen bond. In the case of the evaluated structures, these are only C−H bonds. In total there are 21 atoms/bonds in this group (C−H).

There are two slight differences concerning the comparison to the reference neutron refinements in this section in relation to the last one. Firstly, signed values instead of absolute ones will be evaluated in addition to the averaged absolute values. Secondly, all values for the differences in bond lengths and atomic displacement parameters are divided by the estimated standard deviations resulting from the neutron refinement. This scales the values to the achievable accuracy and makes discussion of the significance of the observed differences possible. The reference refinement will not be aggregated as they do not form a consistent set of values, having been refined with a multitude of options. However, the results from refinements using the SCAN functional without any consideration of the crystal environment were added to exclude the possibility that the functional itself is responsible for all the observed differences. The resulting aggregated values can be found in Figure 6.11.

The bond lengths agreement of the C−H bonds profits from a description of the crystal environment in general. However, the additional decrease in mean deviation from the cluster charge model to PAW-HAR is only 0.7 σ_n. The agreement to neutron anisotropic atomic displacement parameters decreases going from the neglect of the crystal environment to cluster charges and the deviation is no longer centred around zero. This effect is reversed, with the application of PAW-HAR and the mean agreement is again on the same level as the neglect of the crystal environment.

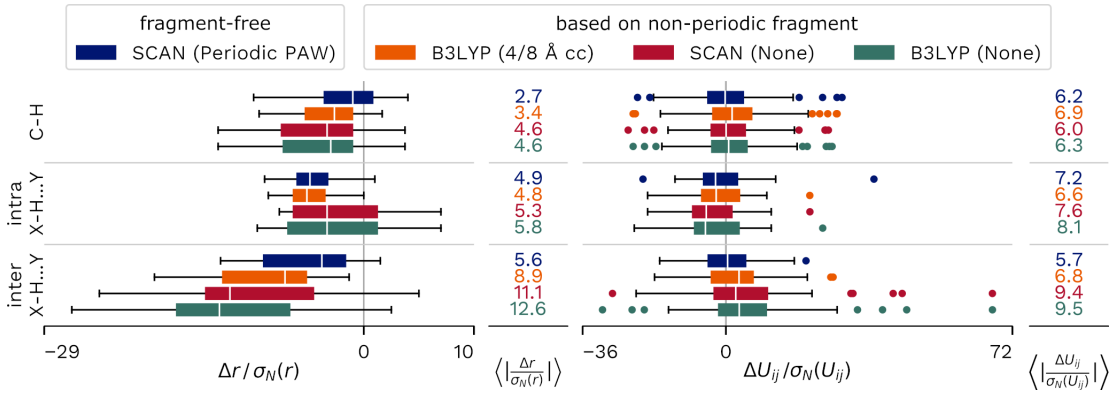


Figure 6.11.: Aggregated comparison of the X−H bond lengths and atomic displacement parameters derived by the different approaches to neutron-derived values. Periodic PAW calculations were conducted with GPAW, B3LYP (4/8 Å cc) were conducted with TONTO/NoSpherA2 in Olex2 and the calculations denominated with 'None' were conducted with Orca/NoSpherA2 in Olex2. Hydrogen atoms were categorised by bond type (C−H bonds or hydrogen bonds). The latter were sorted into hydrogen bonds where the acceptor atom was included in the calculated fragment (intra) or not included in the calculated fragment (inter) for the non-periodic calculations.

The hydrogen bonds located within the same fragment show similar performance for both methods including the crystal environment concerning the bond lengths. However, only the cluster charge approach shows a significant improvement in the agreement in atomic displacement parameters.

As expected, the largest differences are seen for the hydrogen bonds not located within the same fragment. The agreement improves from no crystal description to cluster charges by more than $2\,\sigma_n$. PAW-HAR yields a further improvement of $3.4\,\sigma_n$ over the cluster charge approach. The atomic displacement parameters show a smaller but still significant improvement of $2.6\,\sigma_n$ from no crystal environment to cluster charges and a further improvement of $1.1\,\sigma_n$ for the agreement from cluster charges to PAW-HAR.

6.7. Comparison of computation times for different approaches

After having established the practical benefit of PAW-HAR to the results, this section will show a comparison of the needed computation time to reach that result. Two different approaches have been employed to do this comparison. All the following calculations have used a Dell Precision 3640 system with the following specifications: Intel Core i9-10900K (used cores: 10) and 32 GB RAM.

6.7.1. Comparison to Olex2/NoSpherA2

The first comparison will use the structures and settings from the refinements conducted in Section 6.6. The comparison will be between the calculation using either TONTO or ORCA with NOSPHERA2 and XHARPY with GPAW. Note that this might mean a difference in convergence criteria. The calculations were originally done to compare both methods with a focus on quality instead of computing time. As such, the settings for all calculations are probably using a larger basis, than one would choose for a focus on performance. Additionally, the structures were taken as resulting from the independent atom model refinement. While both approaches might profit from refining with a cheaper functional or smaller basis first and only use more expensive methods starting from this first result, neither approach employed such a prerefinement. The number of variations to include in such an investigation would be unfeasible and in fact, one could also use a cheaper variation of one method as the prerefinement method for the other.

ORCA calculations showed a very good convergence, often within two cycles and in all cases except for **HMA-Mg** (36 min) reached convergence in under 10 minutes.

In contrast, TONTO calculations often showed small oscillations at the end of the refinement, while the calculated structural parameters entered a steady state within the estimated standard deviation. In this case, the first point reaching this steady state was interpreted as convergence and later subsequent points were ignored.

XHARPY/GPAW calculation again showed good convergence properties. The PAW-HAR calculation in GPAW and HAR calculation in TONTO showed lower computation times for different structures. For **A23K** the calculation time of TONTO is almost half (1:03 h against 1:45 h). The included shell of cluster charges was also smaller with 4 Å. Computation time for **Xy** is almost identical between the two methods (PAW-HAR: 3:20 h, cluster charge: 3:38 h). The hydrogen maleate structures provide a best-case scenario for PAW-HAR with computation times of 2:01 h for **HMa-8HQ** and 54 min for **HMA-Mg** compared to 8:23 h and 13:17 h for the calculation using TONTO. Here the periodic calculations profit from the fact, that the calculations were only conducted at the Γ-point, while TONTO suffers from comparatively large fragments.

On the other hand, TONTO shows a lower computation time for **urea** (16 min compared to 42 min). This is however accompanied by a significant benefit in the quality of the description.

This initial comparison showed a large dependence on the chosen settings for the calculation and fragments. Additionally, a potential difference in convergence criteria for the refinement or HAR procedure cannot be excluded. To address these open questions, an additional investigation was conducted using only the **A23K** structure within the XHARPY library.

6.7.2. Comparison of different methods within XHARPy

The last section compared the computation time to already established workflows for calculating with or without cluster charges in Olex2 with the methodology available at the time of the investigation. This section will do another comparison in a slightly more controlled environment. There are a number of key differences.

Refinement using the independent atom model is equal well within the estimated standard deviations between XHARPy and Olex2. It is however not completely identical. The HAR refinements use different convergence criteria. By doing the complete refinement and call of quantum chemical codes in XHARPy, both of these possible sources of differences are eliminated and comparison on a more equal footing can be achieved. While the refinement using XHARPy is currently the only option to drive both methods, the implementation of PAW-HAR in other refinement programs is already planned.

The cluster charge calculation in Olex2 was only available using tonto at the time of the investigation. The overall computation time in tonto is comparatively large, especially as the density from previous calculations is not recycled and Hirshfeld charges for the cluster charges are calculated in a separate calculation for each step. In contrast, this section will use Orca as the quantum chemistry code in the cluster charge calculation and the Hirshfeld charges used are reused from the Hirshfeld partitioning of the density from the previous step. As the positions converge, the differences between determined densities converge as well and therefore a possible influence of this approximation vanishes. No convergence problem could be observed during cluster charge HAR.

Quantum Espresso (QE) was also added as an additional source for a density calculated with PAWs. In contrast to GPAW, it represents the pseudo wave function with a plane wave basis instead of a real-space grid.

In the last section, options were chosen to be consistent with established practice and to reach an optimal quality for the individual methods. This section is focused on comparability. The influence of basis choices on the computation time is included. On the other hand, the functional in this first step is limited to only PBE, which is not the ideal choice but available without any complications for all three investigated quantum chemistry codes. The resulting computation times and quality indicators are listed in Table 6.1.

Comparison of the two PAW-HAR implementations

In direct comparison of the two PAW-HAR implementations using different programs as their backend, it can immediately be seen that the two choices converge to a different limit of $wR_2(F^2)$. While the calculations using GPAW seem to be converged at 0.18 Å for the wave function grid with a value of 3.22 %, QE still does not seem converged at 60 Ry with a value of 3.14 %.

Table 6.1.: Comparison of the performance of different methods for the **A23K** dataset using the PBE functional. The environment denotes the approximation of the crystal environment. A basis in Å denotes a real-space grid, Ry denotes a plane wave basis, a basis set starting with def2 denotes an atom centred basis. t_{XC}: Density calculation, t_{part}: Partitioning, t_{ref}: Refinement. All calculations are done with the PBE functional.

				$\left\langle \frac{\lvert\Delta r\rvert}{\sigma_n(r)} \right\rangle$ / mÅ			$\left\langle \frac{\lvert\Delta U_{ij}\rvert}{\sigma_n(U_{ij})} \right\rangle /10^{-2}\text{Å}^2$			time / min		
environment	program	basis	$\frac{wR_2(F^2)}{\%}$	NH_3	CH_3	CH	NH_3	CH_3	CH	t_{XC}	t_{part}	t_{ref}
periodic	GPAW	0.26 Å	3.25	10.5	3.1	2.2	11.6	7.7	8.8	1:47	0:58	0:37
		0.24 Å	3.24	9.9	3.3	2.2	11.2	7.7	9.2	1:57	1:04	0:39
		0.22 Å	3.23	10.1	3.4	2.2	11.5	7.6	8.8	2:15	1:40	0:42
		0.20 Å	3.23	10.1	3.4	2.2	11.6	7.6	8.4	2:23	1:45	0:41
		0.18 Å	3.22	10.3	3.3	2.2	11.6	7.7	8.9	3:03	2:45	0:41
		0.16 Å	3.22	10.3	3.4	2.2	11.6	7.9	8.9	4:02	3:38	0:42
		0.14 Å	3.22	10.3	3.4	2.2	11.6	7.8	8.9	5:32	5:40	0:40
		0.12 Å	3.22	10.3	3.4	2.2	11.6	7.8	8.9	7:59	8:39	0:39
	QE	40 Ry	3.22	11.2	3.3	2.9	11.9	7.5	9.0	2:09	3:45	0:39
		50 Ry	3.19	12.0	4.0	2.2	11.0	7.7	8.4	2:36	3:47	0:35
		60 Ry	3.14	10.5	3.3	2.2	11.0	7.6	8.6	4:04	5:08	0:45
None	Orca	def2-SVP	3.26	15.9	1.7	10.2	13.3	5.9	5.6	0:58	0:15	0:49
		def2-TZVP	3.25	15.6	3.0	8.2	14.8	6.8	5.1	1:32	0:23	0:55
4 Å cc	Orca	def2-SVP	3.22	14.4	3.7	6.9	11.9	6.0	6.6	1:03	0:14	1:02
		def2-TZVP	3.14	14.5	2.3	4.9	11.0	7.3	5.8	1:23	0:17	0:51
cluster	Orca	def2-SVP	3.20	11.0	6.5	2.2	9.5	6.2	7.8	24:06	1:56	0:52
		def2-TZVP	3.13	12.8	7.0	1.5	9.4	7.0	7.4	46:57	2:50	0:51

There are two possible explanations for this fact. The Hirshfeld partitioning in GPAW was done using the routine provided by GPAW for this purpose. At the same time, the procedure in Quantum Espresso was more custom. Atomic densities were generated by initialising a calculation without a single step from the atomic densities and then using the resulting density of this 'calculation' as the atomic density. The first possibility would be a difference resulting from these two approaches. The difference is noticeable in the computation times. Despite the fact, that this investigation was performed on an SSD, the input/output heavy approach yields a larger time for partitioning compared to the wave function optimisation. This is despite the fact, that the task was implemented in parallel for QE but not for GPAW.

The second possibility would be a difference resulting from the different representations of the wave functions and the resulting gradients, which leads to different densities even when using the same functional. A difference in performance for distances was basically non-existent for the largest basis of the respective calculations. At the same time, a small benefit below the significance of the neutron values was observed in the atomic displacement parameters.

In general, the computation time at convergence is lower for GPAW but does lead to worse results in crystallographic agreement.

Comparison of the different descriptions of the crystal environment in non-periodic calculations

In contrast to the difference in computation times observed between tonto and Orca in the last section, the description using cluster charges in Orca does not yield a higher overall computation time compared to the calculation neglecting the crystal environment completely. At the same time, the cluster charges improve the agreement of the X−H distances to the neutron refinement, while the effect on the atomic displacement parameters is more varied.

Performance using non-periodic calculation seems to be less stable overall: Larger basis sets do not lead to convergence at a better performance level but the relative performance between different basis set sizes is inconsistent. This points to error compensation in unpredictable ways. While it might seem positive to have both an improvement in computation time and agreement when using the smaller def2-SVP basis, the unpredictability of this effect means, that a lower agreement must be assumed without the ability to compare to neutron results, which will be the usual case for application instead of verification.

The description of the crystal surrounding by a cluster of hydrogen-bound molecules does yield an improvement in the agreement of bond lengths and a small improvement in atomic displacement parameters of the atoms gaining a bond partner this way. At the same time, the computation time increases from two minutes to 27 minutes for def2-SVP and from about 2:30 min to 51 minutes for def2-TZVP.

Comparison of periodic and non-periodic calculation

The computation times of the periodic calculation at convergence in GPAW or 60 Ry in Quantum Espresso were at about seven minutes for GPAW and ten minutes for Quantum Espresso. This is significantly larger than the two minutes and fifteen seconds for the cluster charge calculation using the def2-SVP basis set. On the other hand, it is also significantly faster than the calculation using a cluster to describe the crystal environment.

The $wR_2(F^2)$ reached by GPAW at convergence is also higher than the values reached by cluster charge or cluster descriptions using the def2-TZVP basis set. The density calculated with QE and a 60 Ry cut-off is basically identical to the $wR_2(F^2)$ of the non-periodic calculations in this case.

At the same time, the bond distances of the non-periodic calculations are usually higher in two or more of the categories compared to the periodic calculations.

Atomic displacement parameters are improved using non-periodic calculations with a description of the direct crystal environment. Neglect of the environment does lead to worse performance in the agreement of ADPs to neutron values.

Remarks on the comparison

It needs to be reiterated that PBE does not represent an ideal choice for any of the methods investigated. It is however available for all investigated quantum chemistry codes. As described in Section 6.4 good results can be achieved with functionals of the SCAN family. This also holds true for non-periodic calculations. However, metaGGA functionals are not available in Quantum Espresso and hybrid functionals are not available in GPAW. At the same time setups for PBE PAW calculation are available and well-tested. An investigation of additional codes to GPAW and Quantum Espresso might be worth considering as additional points might shine light on the nature of the performance difference between the two codes. The behaviour is consistent with the observations of Section 6.9.

6.8. Comment on the suggested structure factor by Wall

All presented results, as well as the corresponding publication (Ruth *et al.*, 2022), use a phase factor as established for crystallographic structure factor calculations (Coppens, 2010). This is despite the fact that Wall (2016) suggested that this is not appropriate in his pilot study of the application of PAW to the description of urea with HAR.

Instead, he suggested the replacement of the phase with a factor of the form, where $(u_0\, v_0\, w_0)$ are the integer and $(u_1\, v_1\, w_1)$ are the fractional parts of the transformation

of the coordinate shift x' to the grid coordinates such that $(u_1\, v_1\, w_1)$ lie in the range $0 \leq u_1\, v_1\, w_1 < 1$:

$$\begin{aligned}\frac{A_i'(hkl)}{A_i(hkl)} =& e^{-\frac{2\pi i h u_0}{N_1}} e^{-\frac{2\pi i k v_0}{N_2}} e^{-\frac{2\pi i l w_0}{N_3}} \\ &\times [(1-u_1)(1-v_1)(1-w_1) + e^{-\frac{2\pi i h}{N_1}} u_1(1-v_1)(1-w_1) \\ &+ e^{-\frac{2\pi i k}{N_2}} (1-u_1)v_1(1-w_1) + e^{-\frac{2\pi i l}{N_3}} (1-u_1)(1-v_1)w_1 \\ &+ e^{-\frac{2\pi i h}{N_1}} e^{-\frac{2\pi i k}{N_2}} u_1 v_1(1-w_1) + e^{-\frac{2\pi i h}{N_1}} e^{-\frac{2\pi i l}{N_3}} u_1(1-v_1)w_1 \\ &+ e^{-\frac{2\pi i k}{N_2}} e^{-\frac{2\pi i l}{N_3}} (1-u_1)v_1 w_1 + e^{-\frac{2\pi i h}{N_1}} e^{-\frac{2\pi i k}{N_2}} e^{-\frac{2\pi i l}{N_3}} u_1 v_1 w_1]\end{aligned} \tag{6.17}$$

In an implementation in XHARPY, the use of this structure factor did lead to less optimal convergence properties, even though in the iterative implementation of the HAR scheme, the difference at convergence should ultimately be zero. Subsequently, it was decided to use the traditional phase. There were two additional reasons to rationalise the decision.

6.8.1. The factor for Δx and $-\Delta x$ do not offset each other

Wall (2016) derived his factor in the one-dimensional case (equation 15 in the publication). This culminates in the equation.

$$\frac{A_{\Delta x}(h)}{A(h)} = e^{-\frac{2\pi i h u_o}{N}} \left[(1-u_1) + e^{-\frac{2\pi i h}{N}} u_1\right] \tag{6.18}$$

Here $u = N\Delta x$ with u_0 again being the integer component and u_1 being the fractional remainder in the range $0 \leq u_1 < 1$, h is the reciprocal index and N is the number of lattice points in the one-dimensional lattice. For the sake of investigation, it is now assumed that our difference between the calculated one-dimensional position and the original position is only small, *i. e.* $|\Delta x| \leq 1/N$. For a positive Δx this means $u_0 = 0$, $u_1 = N\Delta x$ and Equation 6.18 simplifies to:

$$\frac{A_{\Delta x}(h)}{A(h)} = (1-u_1) + e^{-\frac{2\pi i h}{N}} u_1 = 1 - N\Delta x + N\Delta x \cdot e^{-\frac{2\pi i h}{N}} \tag{6.19}$$

For a sufficiently small negative Δx, $u_0 = -1$ and $u_1 = 1 - N\Delta x$. As such, the equation simplifies to

$$\frac{A_{-\Delta x}(h)}{A(h)} = 1 - N\Delta x + N\Delta x \cdot e^{\frac{2\pi i h}{N}} \tag{6.20}$$

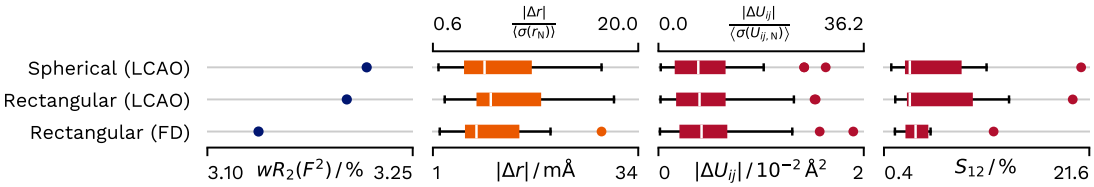


Figure 6.12.: Comparison of the performance of an LCAO density in the spherical expansion to a rectangular grid expansion of the same density and a density calculated with a real-space grid basis. All calculations were done using the RPBE functional in GPAW.

Intuitively, the product $A_{\Delta x} \cdot A_{-\Delta x}$ should be one for the operation to be reversible. However, this is not the case. This might lead to the experienced convergence problems. Additionally, the phase does allow for easier access to atomic form factors, where the atomic position is defined to be at the origin of the reference system.

6.8.2. Evaluation of LCAO-derived densities Fourier transformed on spherical and rectangular grids do not show a large difference

The quantum mechanical code GPAW enables the use of three different bases for the description of the wave function: plane waves, real-space grids and LCAO. Therefore, it enables the comparison of the LCAO density expanded and Fourier transformed on a rectangular grid to an atom-centred grid. As such, a direct comparison of the two expansions using the traditional phase factor is possible and it can be verified, whether the issue raised by Wall can be observed in our implementation.

The expansion and Fourier transform on a spherical grid was implemented in PYTHON using the grid implementation and integration procedures of HORTON (Verstraelen *et al.*, 2015) with grids of the 'insane' preset. The comparison was conducted using the **A23K** dataset. In GPAW, calculations using an LCAO basis are only available with GGA functionals. As such the RPBE functional was used with a grid spacing for the wave function calculation chosen as 0.12 Å. Expansion on the rectangular grid was done with the four-fold interpolation.

As can be seen in Figure 6.12, the difference between the two grid types (spherical and rectangular) is small when both use the LCAO basis. While the $wR_2(F^2)$ is slightly higher for the spherical expansion, the X−H bond distances are improved by an amount below the significance of the neutron-derived distances. As such, the evaluation does not favour the spherical expansion of the LCAO density. At the same time, the rectangular grid wave function (FD) calculation shows a larger improvement in $wR_2(F^2)$ compared to the LCAO expansion, even if the differences for the hydrogen parameters are still small and inconclusive, as for the ADPs the difference to the neutron-derived value

gets larger, while the overlap between the probability distributions as indicated by S_{12} improves.

In conclusion, there is no practical drawback to using the rectangular grid expansion, while at the same time, using it enables us to support a much larger range of software packages and methodologies within those packages. Additionally, the procedure is interoperable with established crystallographic software such as Olex2 (Dolomanov *et al.*, 2009).

6.9. Reducing the number of parameters for DFT benchmarking

So far, atomic Hirshfeld densities were used to determine atomic parameters, with a focus on hydrogen atoms. In other words, the chapter presented a method for achieving aims from the field of crystallography. The ability to achieve this has a systematic dependence on the functional used to obtain the atomic densities and has been tracked using the agreement to neutron diffraction, which is considered the gold standard for the obtained parameters.

While the performance can directly be used as a measure of the quality, the refinement of atomic parameters might influence the result in unexpected ways. This section aims to switch the train of thought around. How can the quality of the atomic density of a given functional be isolated from the other refined parameters and thereby how can PAW-HAR be used as a prototype for density benchmarking for periodic boundary DFT calculations?

Recall that Blessing (1995) showed that in order to apply atomic displacement parameters from neutron diffraction to X-ray diffraction experiments they need to be scaled by the following formula introduced earlier in the theory section:

$$U^{ij}_{\mathrm{c,n}} = q_{\mathrm{ADP}} \cdot U^{ij}_{\mathrm{c,n}} + \Delta U^{ij} \tag{2.82}$$

Accordingly, these up to seven parameters overall can also be used to adapt the atomic displacement parameters of the neutron refinement to the X-ray experiment. A free crystallographic refinement would mean 6 atomic displacement parameters *per atom*. Additionally, the positions from neutron refinement are directly applicable to our X-ray experiment and do not need any modification, which reduces the number of parameters per atom by another three. As neutrons are scattered from the atomic cores there is much less probability to introduce a bias into our density evaluation.

The resulting method will be called '**X**-ray And **N**eutron diffraction data for benchmarking via **TH**eoretically **O**btained **S**tructure factors' or XANTHOS.[3]

In theory, this should provide a benchmark, which is orthogonal to the usual employed energy-based benchmarks in the solid state. A theory-theory benchmark on the question of whether density functionals still reproduce the density has been tackled by Medvedev *et al.* (2017) for isolated atoms and cations. Theory benchmarks for the solid state are generally hard to achieve. At the same time comparing to experimental results has its own challenges. Most benchmarks on the solid state evaluate properties such as equilibrium geometries (Björkman *et al.*, 2012), electronic band-gaps (Borlido *et al.*, 2019), cell volumes and lattice energies (Otero-de-la Roza and Johnson, 2012*b*; Hoja *et al.*, 2017; Dolgonos *et al.*, 2019). The idea is to provide a diffraction-based benchmark in addition and certainly not instead of these well-established datasets. The general idea is similar to the one laid out by Medvedev *et al.* (2017). The minimisation of energy should lead to a correct density, which with a correct functional should lead to the correct energy. By staying closer to the physical background of the method, the method should generalise to a larger number of systems.

6.9.1. Considerations for practical application

Data quality is always the most fundamental consideration in investigations of electron density from X-ray diffraction. This becomes even more important if something as fundamental as a density functional is to be evaluated. As such, structures should be investigated with a normal HAR first. If there are strong features in the difference electron density, significant effort should be spent on searching for systematic deviations stemming from data processing or the crystal itself. The application of crystallographic knowledge to mediate systematic deviations is paramount. Especially resolution-dependent effects could lead to misleading results. As usual, the inner data is most crucial. As most DFT calculations use some form of frozen core, the valence density encoded in the inner reflections is the actual target of our benchmark.

Even though there is no need to decouple the density from the vibration to the same degree as in a normal refinement, a reasonable resolution still ensures that subtle effects can be detected. On the other hand, a resolution, which is too high can bias the quality indicators to the core density, which is not the target of our investigations, as they cannot be affected by the functional itself in the investigated implementations of density functional theory. A similar effect has been observed for X-ray wave function fitting, where Genoni *et al.* (2017) found, that with a too-high resolution, subtle effects in theoretically calculated densities were no longer retrieved. As such an *ad-hoc* procedure was implemented for the prototype of XANTHOS, where the resolution-dependent absolute theoretical structure factor was determined. The cutoff was then set to recover 95 % of

[3]This is even more severe example of a 'meaning-follows-acronym' naming approach, but in one Greek tradition Xanthos or Xanthus is one of the horses of Achilles and son to the harpy Podarge. The opportunity to link the XHARPY package and the derived method was too good not to take it.

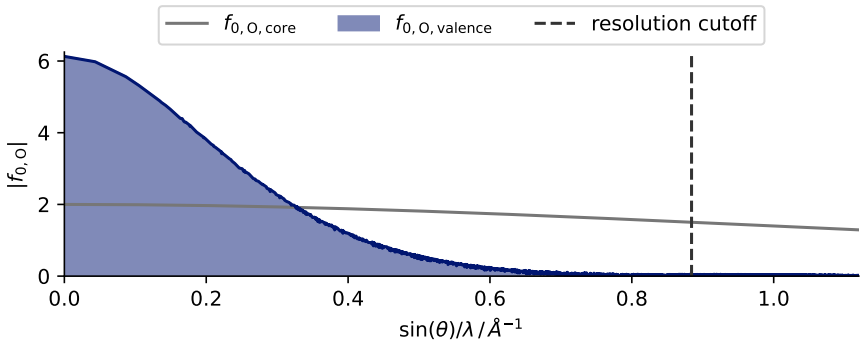


Figure 6.13.: Illustration of the calculation of the resolution cutoff on the fictitious example of a single oxygen atom in a structure.

the integral. A detailed investigation of this procedure was beyond the scope of this work. A graphical illustration for a single oxygen atom in a larger structure is illustrated in Figure 6.13. It is quite obvious, that the core atomic form factor is still very much significant at the cutoff, while the valence atomic form factor is negligible in comparison. If the performance of interest is mainly related to the valence density, the outer data can obscure the desired test.

All the caveats for HAR are also applicable to the XANTHOS method. Hirshfeld atom partitioning is only one of several fuzzy partitioning methods. Which of these provides the best representation of a static vibrating atomic density remains an open question. While Chodkiewicz *et al.* (2020) provided a first study tackling this problem, a canon of work exploring this subject is still missing. A first step to go from prototype to universally applicable benchmark would be to check, whether the relative performance of functionals is robust with respect to the atomic partitioning method. The lower the deviation from average density, *i.e.* the lower the temperature, the smaller this influence should get.

This leads nicely into the still lingering problem, that in its current application, the density of the vibrationally excited system is represented by a modified ground-state density. The effect of this approximation needs to be assessed on each combination of system and temperature range, but if a specific measurement has been shown to be a possible candidate, there should be no dependency on the specific functional. In other words, the preparation of a database of prevetted structures should be possible.

6.9.2. Evaluating the robustness with individual crystal choice

The first step is to evaluate the robustness of the results with the individual measurement. The **OxyH** dataset provides an excellent basis to tackle this problem as the thirteen crystals measured with different setups should provide similar results. A refinement was

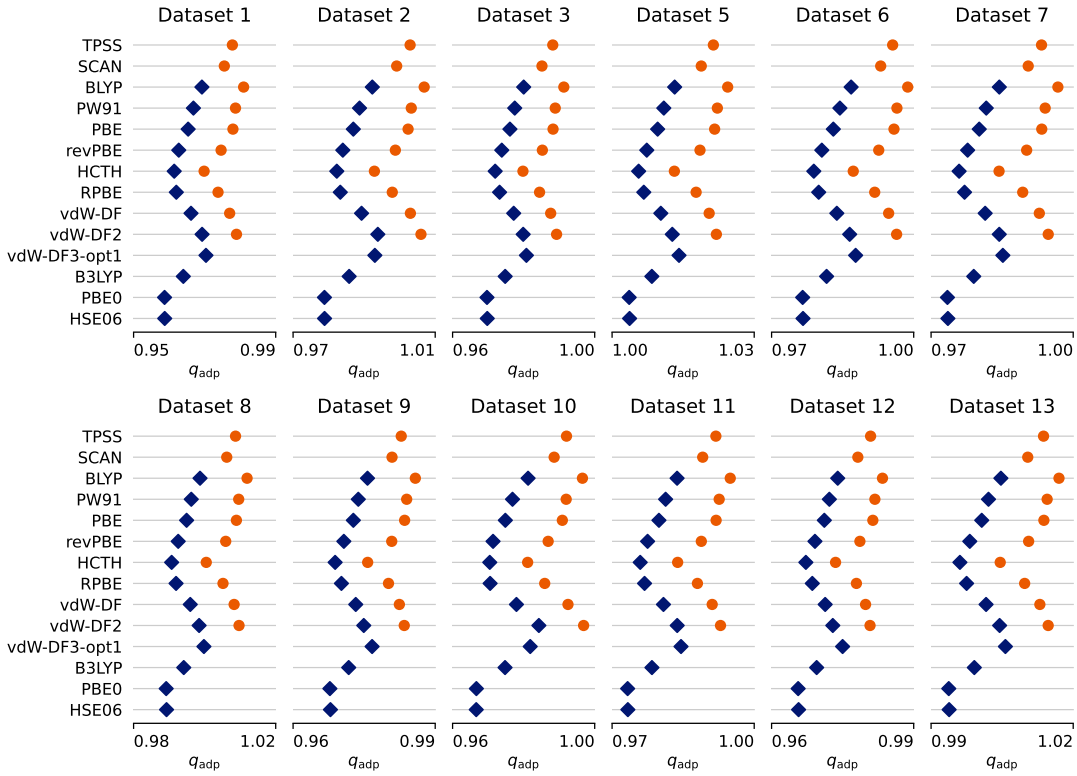


Figure 6.14.: Resulting q_{ADP} for a XANTHOS refinement in dependence of functional and crystal of **OxyH**. Blue diamonds: Results calculated from a QUANTUM ESPRESSO density; Orange spheres: Results calculated from GPAW density.

done for all datasets. However, dataset 4 was excluded from the overall investigation as the mean q_{ADP} of 0.888(8) indicated a measurement temperature significantly different from the reference neutron data collection. The remaining 12 q_{ADP} values are depicted in Figure 6.14.

All determined values lie within 0.95 and 1.03 indicating a good agreement between the measurement temperature of the neutron experiment and the X-ray experiment. The refined q_{ADP} is dependent on the functional in a consistent manner for all datasets. Higher q_{ADP} values should indicate a higher localisation of electrons, which the correction redistributes away from the atomic positions. Interestingly, the two software packages used for calculating the Hirshfeld density show a systematic difference in the relative size of q_{ADP}, with the values of densities calculated with QUANTUM ESPRESSO (QE) being systematically lower than the values from GPAW. The question, of whether this difference stems from the different descriptions of the pseudo wave function (QE: plane-waves, GPAW: real-space grid) or some approximation used in the respective Hirshfeld partitioning routines remains open.

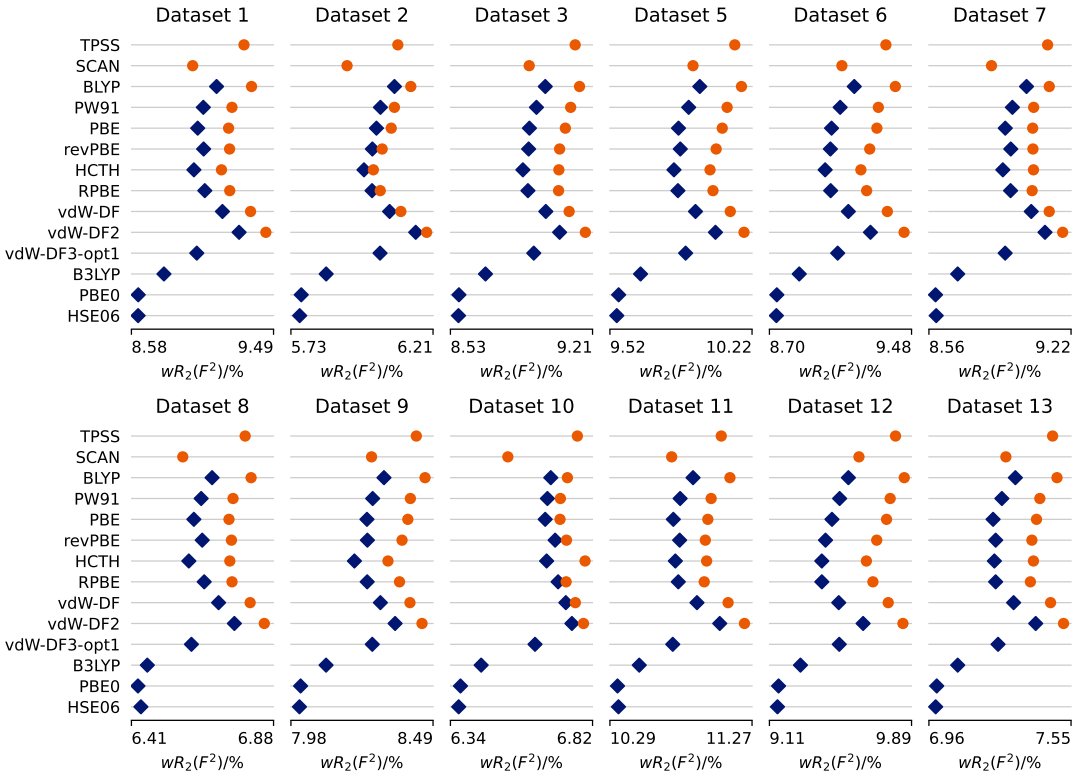


Figure 6.15.: Resulting $wR_2(F^2)$ for a XANTHOS refinement in dependence of functional and crystal of **OxyH**. Blue diamonds: Results calculated from a QUANTUM ESPRESSO density; Orange spheres: Results calculated from GPAW density.

A similar trend can be seen with the crystallographic agreement factor $wR_2(F^2)$ displayed in Figure 6.15. The relative performance of functionals is well preserved over the datasets, especially for the performance of functionals compared to other functionals used in the same program. There is a systematic difference between the two programs used for calculating the densities, where QUANTUM ESPRESSO performs superior to GPAW, which is also consistent with the evaluations from Section 6.7.2.

There are two quality indicators which are directly calculated from the difference electron density: e_{gross} and the difference $\Delta\rho_{\text{diff}} = \Delta\rho_{\text{max}} - \Delta\rho_{\text{min}}$. The preliminary investigation presented here exhibited a significant drop in the consistency of the performance with functional. At first glance, this fact is puzzling, as the difference electron density should be directly connected to the quality of the refinement. However, this is only the case if there is no correlation of the agreement with the resolution of the individual reflections.

This work will now focus on two datasets out of the 13 datasets presented in **OxyH**. One where the connection between the density-based quality indicators and the agreement factor holds (Dataset 2) and one where it does not hold (Dataset 13). The quality

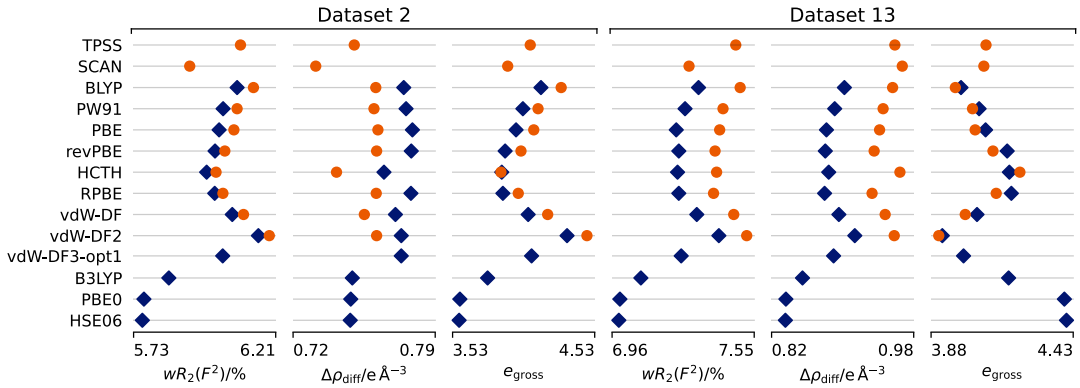


Figure 6.16.: Resulting $wR_2(F^2)$, e_{gross} and $\Delta\rho_{\text{diff}}$ for a XANTHOS refinement in dependence of functional and for the datasets 2 and 13 of **OxyH**. Blue diamonds: Results calculated from a QUANTUM ESPRESSO density; Orange spheres: Results calculated from GPAW density.

indicators for both datasets are depicted in Figure 6.16. For dataset 2 the behaviour of the e_{gross} is as expected. An improvement in $wR_2(F^2)$ also results in an improvement in the overall difference electron density. At the same time, the difference between the maximum and minimum difference electron density roughly follows the same trend but is much more flattened. This is the behaviour that might be expected from a high-quality dataset.

On the other hand, dataset 13 does show the same relative performance in $wR_2(F^2)$, but the e_{gross} runs opposite to the performance in the data agreement. At the same time, the difference between maximum and minimum difference electron density shows larger differences and is systematically following the performance in $wR_2(F^2)$.

An indicator of successful refinement is the resolution-dependent scaling plot commonly known as DRKplot (Stash, 2007) from the program, where it is usually displayed with. A resolution dependency of the agreement can stem from either modelling problems or problems during data collection and processing. The resulting plots and the reference plots from a multipolar refinement of Kaminski *et al.* (2014) are displayed in Figure 6.17. Two things are immediately obvious.

Firstly, the inner data of Dataset 13 is not adequately described in the very flexible multipolar reference refinement. A check for problems in the data of this dataset is therefore warranted.

Secondly, for dataset 2, the XANTHOS refinement yields modelled intensities, which are too low on average, while dataset 13 does not show this behaviour. This behaviour is consistent with all datasets, which show a well-behaved DRKPlot in the reference multipolar refinement. This warrants further future investigation.

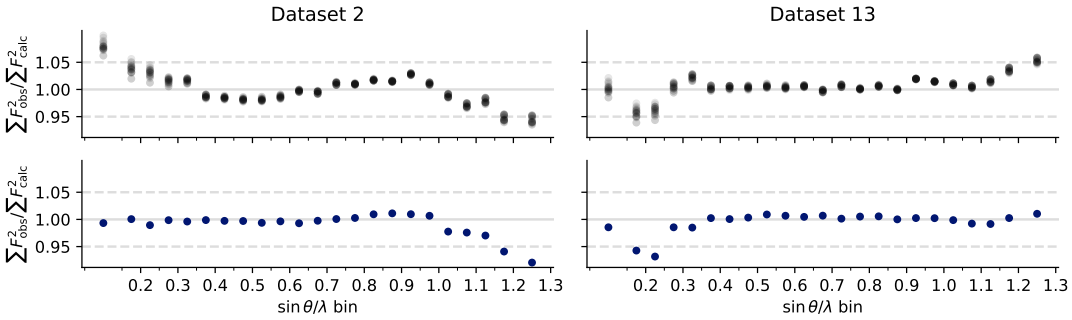


Figure 6.17.: Resolution-dependent quotient of sums of observed and modelled values for F^2 as usually done with DRKPlot (Stash, 2007). The top row is an overlay of the values for all XANTHOS refinement of the two datasets. The bottom row is the result of the reference multipolar refinement as published by Kaminski *et al.* (2014). The first and last bins are larger to avoid single reflections in a bin.

In summary, the XANTHOS method is a way to isolate the performance of functionals from the additional parameters, which are fitted in a crystallographic refinement. It is still in an experimental stage, where limits of the method in structures and data quality need to be tested in future investigations. In its current state, it provides a performance of the $wR_2(F^2)$ with crystal quality, while density-dependent parameters require a more rigorous quality assessment of the crystal quality and data processing. In my personal opinion, this rigorous approach in combination with interpretable difference electron density quality indicators is to be strongly preferred.

6.9.3. Applying the method to A23K and Xy and discussion

As written before, care needs to be applied in choosing candidate datasets for the XANTHOS method. As such a HAR without a structured difference electron density should be an essential starting point. If the density cannot be fitted with additional degrees of freedom, it will probably not be well-behaved with the limited set of parameters used in XANTHOS. There are two different reasons, why datasets included in the evaluation so far may be unsuitable.

In the case of the **urea** dataset, Gram-Charlier features were seen and successfully described against X-ray data. Ideally, these would be fitted against neutron data. However, Gram-Charlier does need a considerably higher resolution for refinement (Kuhs, 1992). While the Gram-Charlier parameters could also be refined against the X-ray data, this would defeat the purpose of reducing the number of parameters.

In the case of the hydrogen maleate structures, the difference electron density is much more likely to stem from problems in the data collection or processing, as current

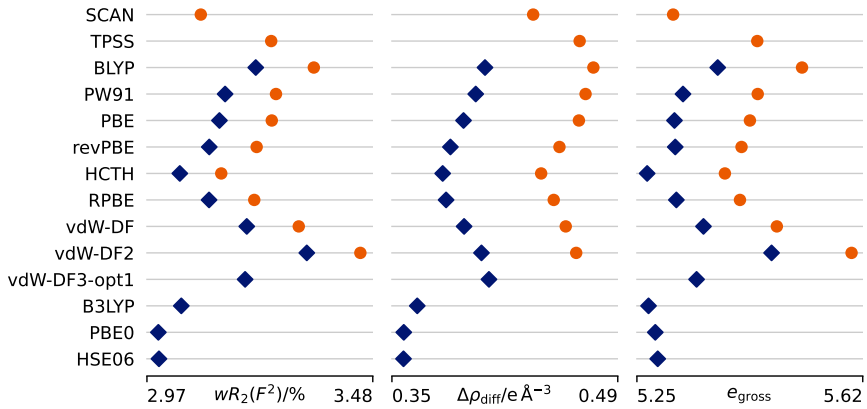


Figure 6.18.: Resulting $wR_2(F^2)$, e_{gross} and $\Delta\rho_{diff}$ for a XANTHOS refinement of **A23K** in dependence of functional. Blue diamonds: Results calculated from a QUANTUM ESPRESSO density; Orange spheres: Results calculated from GPAW density.

crystallographic approaches did not reduce the difference electron density. Therefore, the strong possibility of a systematic influence in the data precludes the use of these datasets. Subsequently, the final application will focus on two datasets: **A23K** and **Xy**. The respective results are depicted in Figures 6.18 and 6.19.

Again, it is obvious, that the two codes employed perform differently, where QUANTUM ESPRESSO shows a performance, which is systematically improved over GPAW. Interestingly the relative difference seems to be smaller for the refinement of **A23K**, than for **Xy**. The source of the difference cannot be identified at this moment and may be searched in either the description of the wave functions (real-space grid for GPAW, plane-waves for GPAW) or in the internal routine of GPAW for the determination of the Hirshfeld densities.

The functional dependence for the resulting quality indicators is very well retained between the two datasets. Within the metaGGA functionals, which could only be evaluated in GPAW, the SCAN functional also proves to be superior to TPSS with the reduced set of parameters in the XANTHOS refinement. Within the GGA functionals, the best performance in both programs is achieved using the HCTH/120 functional. The performance of HCTH is to be expected to a certain degree as it is fitted to reproduce potentials. Again functionals including van-der-Waals interactions lead to a lower agreement to the X-ray diffraction intensities. It seems that the approaches are not able to achieve a closer agreement in both density and energy, but have sacrificed some accuracy in the resulting density to get an improvement in the agreement to energy. The highest agreement is shown by the Hybrid functionals with PBE0 and HSE06 showing an almost identical performance, which is slightly superior to the performance of B3LYP. Hybrid functionals could only be evaluated in QUANTUM ESPRESSO.

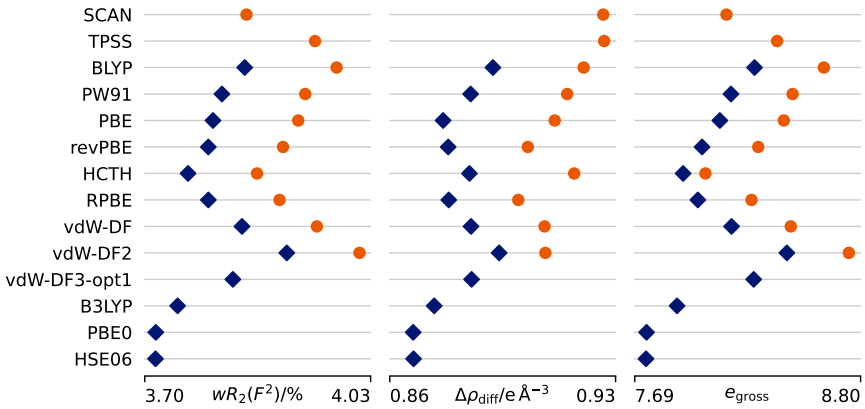


Figure 6.19.: Resulting $wR_2(F^2)$, e_{gross} and $\Delta\rho_{diff}$ for a XANTHOS refinement of **Xy** in dependence of functional. Blue diamonds: Results calculated from a QUANTUM ESPRESSO density; Orange spheres: Results calculated from GPAW density.

6.10. Conclusion and Outlook

This chapter established using PAW densities in periodic boundary conditions in HAR, which so far had only been evaluated in non-iterative HAR in a single pilot study. This approach removes one source of ambiguity in the setup of these calculations, namely, the decision of how a representative fragment from the unit cell looks like. There is no potential bias from fortunate or unfortunate selections of the calculated fragment and/or cluster radii. In this investigation, this concern applies to the **A23K** structure, where a cluster of 4 Å of Hirshfeld charges provided an improved agreement to diffraction data and neutron results over the larger 8 Å cluster in the fragment-based method.

The benefits for the agreement of the model to the data, as well as the improved reproduction of X−H bond length derived from neutron diffraction, were successfully demonstrated, both over results from literature as well as results from provided non-periodic comparison calculations. Only embedding in quantum mechanical clusters did provide similar results. However, at this point, the PAW-HAR implementation is no longer the more computationally demanding method, at least in our comparison investigation. An exhaustive investigation, where each participant contributes their own method using a set of compromises of their own choice might be worthwhile.

The implementation was not limited to custom code for the investigated systems, but instead, a PYTHON package was provided to the scientific community, which should enable interested readers to both reproduce the findings presented here, as well as offer the possibility to apply the methodology for the refinement of their own structures. The ability to combine the method with more established refinement codes is provided on a basic level by the ability to calculate atomic form factor 'tsc' files from given 'cif'

files. A more convenient way to access the calculation of atomic form factors using the implemented PAW codes is already planned.

The interaction with the crystal environment is included in the theoretical method. This facilitates the use of this method for benchmarking purposes. There is no longer the question of whether a difference between theory and experiment stems from an inadequate description of the surrounding of the molecule or other structure within the crystal. If everything worked perfectly, the calculated density would be the ground state density of the unit cell within the crystal and therefore the agreement between calculated and measured X-ray diffraction intensities can be used as a quality indicator for the given method of calculation. This chapter has also presented how such a benchmark could be streamlined using neutron data. The fact that positions are no longer refined, reduces the number of quantum chemical calculations needed to one. The ability of the refinement to attribute deficiencies of a given functional into crystallographic parameters has been reduced.

What remains an open question is the quality of the deconvolution of vibrational and density influences. Different density partitioning methods should be evaluated to evaluate the performance of their partitioning in finding atomic contributions compatible with the crystallographic model for vibration within the crystal. Additionally, more sophisticated models of displacement could reduce systematic errors from inadequate descriptions. The influence of the current models should be evaluated. The current obstacle to such an investigation is the unavailability of a theoretical model of the mean density at finite temperature, which is not in itself based on atomic partitioning. By removing noise present in the experimental structure factors, these influences could be studied in more detail. Such an investigation could also evaluate the size of the benefit of the Xanthos procedure over conventional HAR and thus determine whether having the neutron data is only beneficial or necessary. After such an investigation, given that it is successful, the benchmark of functionals against carefully vetted X-ray and neutron data should prove worthwhile as a measure of functional performance that is orthogonal to the energy.

Finally, as this chapter was focused on validation, the application was limited to structures and thereby questions which could be subjected to HAR in the past. However, the periodic basis also enables refinement for structures where a fragment of reasonable size could not be chosen. Using the aspheric atomic form factors calculated for chiral grey selenium it could be shown that the fact that the two enantiomeric forms are not distinguishable by resonance scattering, might not be true for refinements using aspherical form factors. An experimental evaluation of this fact, which contradicts an exercise long established in Göttingen's teaching of crystallography, was not possible in the time frame of this thesis.

In summary, this chapter succeeded on two points. The interaction of the crystal environment could be included in the calculation of atomic densities for HAR and the fact that this is included enabled me to build a prototype for a structure-factor-based benchmark.

7. Summary and Outlook

This thesis aimed to find systems and methods for benchmarking quantum chemical methods using quantum crystallographic methodology. While some of the systems and approaches proved to be unsuitable for this purpose for a variety of reasons, a methodology could ultimately be established with the use of periodic boundary conditions for Hirshfeld atom refinement as well as the derived '**X**-ray And **N**eutron diffraction data for benchmarking via **TH**eoretically **O**btained **S**tructure factors' or XANTHOS method and the evaluation of the MetalJet diffractometer for the purpose of quantum crystallographic measurements.

The first approach, presented in Chapter 3 was the original idea of this project, namely using the luminescence properties of crystals of anthracene derivatives. While the data proved to be unsuitable in its current form for the strict requirements enforced by benchmarking, valuable information could be obtained about the luminescence of the system class in general. For the first group of compounds, intermolecular interactions had been proposed as a factor determining the strength of this luminescence. This could be shown to be highly unlikely due to the difference between the low-temperature and room-temperature structures determined by X-ray diffraction. Ultimately, the preformed bending of the anthracene molecule in addition to the flexibility of the system proved to be the most promising candidates for explaining the differences.

The following Chapter 4 investigated the case of decamethylcyclopentadienyl silicon. This system was chosen for investigation as the reported X-ray structure contained the molecule in two very distinct conformations, which had been the topic of quantum chemical investigations in the past. The answer to why this structure behaved differently from its heavier analogues, proved to be a phase transition and as such, this chapter illustrated how the neglect of entropic effects can lead discussions about most stable states astray. The newly observed phase change was evaluated using geometric and crystallographic quantities and methods of computational chemistry were employed to underpin the idea that dispersive interactions play a large role in determining, which conformer is the most stable.

While the previous projects provided valuable insights, the first project to advance towards a benchmark using X-ray diffraction was presented in Chapter 5. The indium K_α radiation provided by the new Excillum MetalJet in Göttingen has the highest photon energy of a commercial in-house source. Thereby, it can potentially improve the study of high electron density materials, such as inorganic solid-state compounds from heavy elements held together by strong ionic interactions. To unearth the potential of

these compounds often poses a challenge for the quantum crystallographic evaluation, as the scattering of the involved elements often enables the measurement to the resolution limit of a given wavelength, but strong resonance scattering effects and absorption can introduce systematic biases. The evaluation of the detector now completes the assessment of parts of the setup of the machine. The chapter could successfully establish the Eiger2 1M CdTe as the detector meeting the unique challenges posed by that system. In addition to independent atom model refinements of the target compounds, the investigation included a first charge density refinement with very promising results. However, these need to be repeated for the target ionic compounds. Extrapolating from the data of this evaluation, the MetalJet should fare well for these applications.

After tackling how certain compounds are measured with an in-house source with the data quality required for a benchmark, Chapter 6 tackled one of the possibilities for approaching such a benchmark. First, it showed how including the interaction with the full crystal environment in the calculation of the density of a given system benefits the crystallographic Hirshfeld Atom Model. Here, the performance of the different functionals was already reproducible for the structures within the investigation. It then showed how instead of refining positions and atomic displacement parameters and comparing them to neutron refinement, available neutron data can be used directly to limit the number of calculation steps, as well as, possible biases where the refinement of crystallographic parameters compensates for the deficiency of a given functional. While the presented methodology warrants further validation, as well as the evaluation of an optimal partitioning method, the systematic reproduction of the relative performance with different high-quality datasets, shows great promise for a benchmark not directly dependent on energy but making the independent evaluation of the produced density possible for methods in density functional theory.

As such, this thesis has opened the door for further explorations with a solid set of methods. Even though I might have wished to explore the room behind it further, this was beyond the time constraints of a thesis.

Bibliography

Andersen, O. K. (1975). *Phys. Rev., B*, **12**(8), 3060–3083.

Aragó, J. and Troisi, A. (2015). *Phys. Rev. Lett.* **114**(2), 026402.

Atwood, J. L., Hunter, W. E., Cowley, A. H., Jones, R. A. and Stewart, C. A. (1981). *J. Chem. Soc., Chem. Commun.* pp. 925–927.

Bader, R. F. W. (1985). *Acc. Chem. Res.* **18**(1), 9–15.

Becker, P. J. and Coppens, P. (1974*a*). *Acta Cryst. A*, **30**(2), 129–147.

Becker, P. J. and Coppens, P. (1974*b*). *Acta Cryst. A*, **30**(2), 148–153.

Birkedal, H., Madsen, D., Mathiesen, R. H., Knudsen, K., Weber, H.-P., Pattisond, P. and Schwarzenbach, D. (2004). *Acta Cryst. A*, **60**(5), 371–381.

Bishop, C. M. (2009). *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York: Springer-Verlag, 1st ed.

Björkman, T., Gulans, A., Krasheninnikov, A. V. and Nieminen, R. M. (2012). *J. Phys. Condens. Matter*, **24**(42), 424218.

Blessing, R. H. (1995). *Acta Cryst. B*, **51**(5), 816–823.

Blöchl, P. E. (1994). *Phys. Rev., B*, **50**(24), 17953–17979.

Blöchl, P. E., (2022). Cp-paw hands-on course on first-principles calculations - theory of first principles calculations.

Borlido, P., Aull, T., Huran, A. W., Tran, F., Marques, M. A. L. and Botti, S. (2019). *J. Chem. Theory Comput.* **15**(9), 5069–5079.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S. and Zhang, Q., (2018). Jax: composable transformations of python+numpy programs.

Bragg, W. H. and Bragg, W. L. (1913). *Proc. R. Soc. Lond. A*, **88**(605), 428–438.

Bragg, W. L., James, R. W. and Bosanquet, C. H. (1922). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **44**(261), 433–449.

Brandenburg, J. G. and Grimme, S. (2016). *Acta Cryst. B*, **72**(4), 502–513.

Brandenburg, J. G., Maas, T. and Grimme, S. (2015). *J. Chem. Phys.* **142**(12), 124104.

Bricogne, G. (2010). In *International Tables for Crystallography Volume B: Reciprocal space*, edited by U. Shmueli, vol. B of *International Tables for Crystallography*, pp. 24–113. second online edition: Springer Netherlands.

Bronstein, I. N., Semendjajew, K. A., Musiol, G. and Mühlig, H. (1993). *Taschenbuch der Mathematik*. Frankfurt am Main: Harri Deutsch, 1st ed.

Broyden, C. G. (1970). *IMA J. Appl. Math.* **6**(1), 76–90.

Bruker AXS Inc., (2019). Saint v8.40b.

Capelli, S. C., Bürgi, H.-B., Dittrich, B., Grabowsky, S. and Jayatilaka, D. (2014). *IUCrJ*, **1**(5), 361–379.

Chodkiewicz, M., Pawlędzio, S., Woińska, M. and Woźniak, K. (2022). *IUCrJ*, **9**(2), 298–315.

Chodkiewicz, M. L., Woińska, M. and Woźniak, K. (2020). *IUCrJ*, **7**(6), 1199–1215.

Clegg, W. and Harrington, R. W., (2021). Ccdc 2059713: Experimental crystal structure determination.

Clementi, E. and Raimondi, D. L. (1963). *J. Chem. Phys.* **38**(11), 2686–2689.

Compton, A. H. (1915). *Nature*, **95**(2378), 343–344.

Contreras-García, J., Johnson, E. R., Keinan, S., Chaudret, R., Piquemal, J.-P., Beratan, D. N. and Yang, W. (2011). *J. Chem. Theory Comput.* **7**(3), 625–632.

Cooley, J. W. and Tukey, J. W. (1965). *Math. Comput.* **19**(90), 297–301.

Coolidge, W. D., (1916). Vacuum-tube, united states patent us 1,203,495.

Coolidge, W. D. (1925). *J. Franklin Inst.* **199**(5), 619–648.

Coppens, P. (2010). In *International Tables for Crystallography Volume B: Reciprocal space*, edited by U. Shmueli, vol. B of *International Tables for Crystallography*, pp. 10–23. second online edition: Springer Netherlands.

Coppens, P., Guru Row, T. N., Leung, P., Stevens, E. D., Becker, P. J. and Yang, Y. W. (1979). *Acta Cryst. A*, **35**(1), 63–72.

Creagh, D. C. (2010). In *International Tables for Crystallography Volume C: Mathematical, physical and chemical tables*, edited by E. Prince, vol. C of *International Tables for Crystallography*, pp. 241–258. Third edition: Wiley.

Crespo-Otero, R., Li, Q. and Blancafort, L. (2019). *Chem. Asian J.* **14**(6), 700–714.

Crookes, W. (1879). *Philos. Trans. R. Soc. Lond.* **170**, 135–164.

Czernek, J., Pawlak, T. and Potrzebowski, M. J. (2012). *Chem. Phys. Lett.* **527**, 31–35.

Davydov, A. S. (1964). *Sov. phys., Usp.* **7**(2), 145–178.

Dawson, B. (1967). *Proc. R. Soc. Lond. A*, **298**(1454), 255–263.

De Meulenaer, J. and Tompa, H. (1965). *Acta Cryst.* **19**(6), 1014–1018.

Debye, P. (1915). *Ann. Phys.* **351**(6), 809–823.

Destro, R., Marsh, R. E. and Bianchi, R. (1988). *J Phys. Chem.* **92**(4), 966–973.

Dion, M., Rydberg, H., Schröder, E., Langreth, D. C. and Lundqvist, B. I. (2004). *Phys. Rev. Lett.* **92**(24), 246401.

Diop, M. B., Diop, L. and Maris, T. (2016). *Acta Cryst. E*, **72**(4), 482–485.

Dittrich, B. and Jayatilaka, D. (2012). In *Electron Density and Chemical Bonding II*, pp. 27–45. Springer, Berlin, Heidelberg.

Dittrich, B., Sze, E., Holstein, J. J., Hübschle, C. B. and Jayatilaka, D. (2012). *Acta Cryst. A*, **68**(4), 435–442.

Dolgonos, G. A., Hoja, J. and Boese, A. D. (2019). *Phys. Chem. Chem. Phys.* **21**(44), 24333–24344.

Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. and Puschmann, H. (2009). *J. Appl. Cryst.* **42**(2), 339–341.

Dommet, M., (2022). Exciton coupling. `https://github.com/mdommett/exciton_coupling`.

Doyle, P. A. and Turner, P. S. (1968). *Acta Cryst. A*, **24**(3), 390–397.

Dunning, T. H. (1989). *J. Chem. Phys.* **90**(2), 1007–1023.

Díaz Mirón, J. E. Z. and Stein, M. (2022). *Phys. Chem. Chem. Phys.* **24**(48), 29338–29349.

Eickerling, G., Hauf, C., Scheidt, E.-W., Reichardt, L., Schneider, C., Muñoz, A., Lopez-Moreno, S., Humberto Romero, A., Porcher, F., André, G., Pöttgen, R. and Scherer, W. (2013). *Z. Anorg. Allg. Chem.* **639**(11), 1985–1995.

Enkovaara, J., Rostgaard, C., Mortensen, J. J., Chen, J., Dułak, M., Ferrighi, L., Gavnholt, J., Glinsvad, C., Haikola, V., Hansen, H. A., Kristoffersen, H. H., Kuisma, M., Larsen, A. H., Lehtovaara, L., Ljungberg, M., Lopez-Acevedo, O., Moses, P. G., Ojanen, J., Olsen, T., Petzold, V., Romero, N. A., Stausholm-Møller, J., Strange, M., Tritsaris, G. A., Vanin, M., Walter, M., Hammer, B., Häkkinen, H., Madsen, G. K. H., Nieminen, R. M., Nørskov, J. K., Puska, M., Rantala, T. T., Schiøtz, J., Thygesen, K. S. and Jacobsen, K. W. (2010). *J. Phys. Condens. Matter*, **22**(25), 253202.

Fei, Z., Kocher, N., Mohrschladt, C. J., Ihmels, H. and Stalke, D. (2003). *Angew. Chem.* **42**(7), 783–787.

Finkelmeier, N. (2013). *Synthesis and Fluorescence Properties of Anthracene Derivatives and their Metal Complexes*. Dissertation, Göttingen, Georg-August-Universität, Göttingen.

Fischer, E. O. and Grubert, H. (1956*a*). *Z. Naturforsch. B*, **11**(7), 423–424.

Fischer, E. O. and Grubert, H. (1956*b*). *Z. Anorg. Allg. Chem.* **286**(5-6), 237–242.

Fletcher, R. (1970). *Comput. J.* **13**(3), 317–322.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman,

J. B. and Fox, D. J., (2016). Gaussian~16 Revision C.01. Gaussian Inc. Wallingford CT.

Fugel, M., Jayatilaka, D., Hupf, E., Overgaard, J., Hathwar, V. R., Macchi, P., Turner, M. J., Howard, J. A. K., Dolomanov, O. V., Puschmann, H., Iversen, B. B., Bürgi, H.-B. and Grabowsky, S. (2018). *IUCrJ*, **5**(1), 32–44.

Gauß, C. F. (1866). In *Carl Friedrich Gauss Werke*, vol. 3, pp. 265–303. Göttingen: Königlichen Gesellschaft der Wissenschaften.

Genoni, A., Dos Santos, Leonardo H. R., Meyer, B. and Macchi, P. (2017). *IUCrJ*, **4**(2), 136–146.

Genoni, A., Franchini, D., Pieraccini, S. and Sironi, M. (2018). *Chem. Eur. J.* **24**(58), 15507–15511.

Genoni, A. and Macchi, P. (2020). *Crystals*, **10**(6).

Giacovazzo, C., Monaco, H., Artioli, G., Viterbo, D., Milanesio, M., Gilli, G., Gilli, P., Zanotti, G., Ferraris, G. and Catti, M. (2011). *Fundamentals of Crystallography*. International Union of Crystallography Texts on Crystallography. OUP Oxford.

Giannozzi, P., Andreussi, O., Brumme, T., Bunau, O., Buongiorno Nardelli, M., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Cococcioni, M., Colonna, N., Carnimeo, I., Dal Corso, A., de Gironcoli, S., Delugas, P., DiStasio, R. A., Ferretti, A., Floris, A., Fratesi, G., Fugallo, G., Gebauer, R., Gerstmann, U., Giustino, F., Gorni, T., Jia, J., Kawamura, M., Ko, H.-Y., Kokalj, A., Küçükbenli, E., Lazzeri, M., Marsili, M., Marzari, N., Mauri, F., Nguyen, N. L., Nguyen, H.-V., Otero-de-la Roza, A., Paulatto, L., Poncé, S., Rocca, D., Sabatini, R., Santra, B., Schlipf, M., Seitsonen, A. P., Smogunov, A., Timrov, I., Thonhauser, T., Umari, P., Vast, N., Wu, X. and Baroni, S. (2017). *J. Phys. Condens. Matter*, **29**(46), 465901.

Goldfarb, D. (1970). *Math. Comput.* **24**(109), 23–26.

Gonze, X., Amadon, B., Antonius, G., Arnardi, F., Baguet, L., Beuken, J.-M., Bieder, J., Bottin, F., Bouchet, J., Bousquet, E., Brouwer, N., Bruneval, F., Brunin, G., Cavignac, T., Charraud, J.-B., Chen, W., Côté, M., Cottenier, S., Denier, J., Geneste, G., Ghosez, P., Giantomassi, M., Gillet, Y., Gingras, O., Hamann, D. R., Hautier, G., He, X., Helbig, N., Holzwarth, N., Jia, Y., Jollet, F., Lafargue-Dit-Hauret, W., Lejaeghere, K., Marques, M. A., Martin, A., Martins, C., Miranda, H. P., Naccarato, F., Persson, K., Petretto, G., Planes, V., Pouillon, Y., Prokhorenko, S., Ricci, F., Rignanese, G.-M., Romero, A. H., Schmitt, M. M., Torrent, M., van Setten, M. J., van Troeye, B., Verstraete, M. J., Zérah, G. and Zwanziger, J. W. (2020). *Comput. Phys. Commun.* **248**, 107042.

Graw, N. (2022). *Bis(benzoxazol-2-yl)methanides in main group organometallic syntheses and research based education*. Phd thesis, Georg-August-Universität, Göttingen.

Gray, W. R. (1930). *Br. J. Radiol.* **3**(28), 171–177.

Grenz, M., Hahn, E., Du Mont, W.-W. and Pickardt, J. (1984). *Angew. Chem.* **23**(1), 61–63.

Grimme, S. (2006). *J. Comput. Chem.* **27**(15), 1787–1799.

Grimme, S., Antony, J., Ehrlich, S. and Krieg, H. (2010). *J. Chem. Phys.* **132**(15), 154104.

Grimme, S., Ehrlich, S. and Goerigk, L. (2011). *J. Comput. Chem.* **32**(7), 1456–1465.

Grosse-Kunstleve, R. W. and Adams, P. D. (2002). *J. Appl. Cryst.* **35**(4), 477–480.

Guzei, I. A., Bikzhanova, G. A., Spencer, L. C., Timofeeva, T. V., Kinnibrugh, T. L. and Campana, C. F. (2008). *Cryst. Growth Des.* **8**(7), 2411–2418.

Hamilton, W. C. (1965). *Acta Cryst. A*, **18**(3), 502–510.

Hansen, N. K. and Coppens, P. (1978). *Acta Cryst. A*, **34**(6), 909–921.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. and Oliphant, T. E. (2020). *Nature*, **585**(7825), 357–362.

Hartree, D. R. (1925). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **50**(295), 289–306.

Hartree, D. R. and Waller, I. (1929). *Proc. R. Soc. Lond. A*, **124**(793), 119–142.

He, M., Wong, C. H., Shi, D., Tse, P. L., Scheidt, E.-W., Eickerling, G., Scherer, W., Sheng, P. and Lortz, R. (2015). *J. Phys. Condens. Matter*, **27**(7), 075702.

Hemberg, O., Otendal, M. and Hertz, H. M. (2003). *Appl. Phys. Lett.* **83**(7), 1483–1485.

Hirshfeld, F. (1971). *Acta Cryst. B*, **27**(4), 769–781.

Hjorth Larsen, A., Jørgen Mortensen, J., Blomqvist, J., Castelli, I. E., Christensen, R., Dułak, M., Friis, J., Groves, M. N., Hammer, B., Hargus, C., Hermes, E. D., Jennings, P. C., Bjerre Jensen, P., Kermode, J., Kitchin, J. R., Leonhard Kolsbjerg, E., Kubal, J., Kaasbjerg, K., Lysgaard, S., Bergmann Maronsson, J., Maxson, T., Olsen, T., Pastewka, L., Peterson, A., Rostgaard, C., Schiøtz, J., Schütt, O., Strange, M., Thygesen, K. S., Vegge, T., Vilhelmsen, L., Walter, M., Zeng, Z. and Jacobsen, K. W. (2017). *J. Phys. Condens. Matter*, **29**(27), 273002.

Hoffman, M. D., Gelman, A. *et al.* (2014). *J. Mach. Learn. Res.* **15**(1), 1593–1623.

Hohenberg, P. and Kohn, W. (1964). *Phys. Rev.* **136**(3B), B864–B871.

Hoja, J., Reilly, A. M. and Tkatchenko, A. (2017). *WIREs Comput. Mol. Sci.* **7**(1), e1294.

Huang, L., Massa, L. and Karle, J. (1999). *J. Mol. Struct.* **474**(1-3), 9–12.

Hübschle, C. B., Sheldrick, G. M. and Dittrich, B. (2011). *J. Appl. Cryst.* **44**(6), 1281–1284.

Hutton, A. T. (1986). *Annu. Rep. Prog. Chem., Sect. B*, **83**, 221–239.

Jayatilaka, D. and Dittrich, B. (2008). *Acta Cryst. A*, **64**(3), 383–393.

Jayatilaka, D. and Grimwood, D. J. (2001). *Acta Cryst. A*, **57**(1), 76–86.

Jayatilaka, D. and Grimwood, D. J. (2003). *Computational Science - ICCS 2003*, **2660**, 142–151.

Jeitschko, W., Gerss, M. H., Hoffmann, R.-D. and Lee, S. (1989). *J. Less Common Met.* **156**(1-2), 397–412.

Johnson, C. K. and Levy, H. A. (1974). In *International Tables for X-ray Crystallography*, vol. 4, pp. 311–336. Birmingham: Kynoch Press.

Jutzi, P. and Hielscher, B. (1985). *J. Organomet. Chem.* **291**(3), c25–c27.

Jutzi, P., Holtmann, U., Kanne, D., Krüger, C., Blom, R., Gleiter, R. and Hyla-Kryspin, I. (1989). *Chem. Ber. Recl.* **122**(9), 1629–1639.

Jutzi, P., Kanne, D. and Krüger, C. (1986). *Angew. Chem.* **25**(2), 164.

Jutzi, P., Kohl, F., Hofmann, P., Krüger, C. and Tsay, Y.-H. (1980). *Chem. Ber. Recl.* **113**(2), 757–769.

Kaminski, R., Domagala, S., Jarzembska, K. N., Hoser, A. A., Sanjuan-Szklarz, W. F., Gutmann, M. J., Makal, A., Malinska, M., Bak, J. M. and Wozniak, K. (2014). *Acta Cryst. A*, **70**(1), 72–91.

Kleemiss, F., Dolomanov, O. V., Bodensteiner, M., Peyerimhoff, N., Midgley, L., Bourhis, L. J., Genoni, A., Malaspina, L. A., Jayatilaka, D., Spencer, J. L., White, F., Grundkötter-Stock, B., Steinhauer, S., Lentz, D., Puschmann, H. and Grabowsky, S. (2021). *Chem. Sci.* **12**(5), 1675–1692.

Klein, Y. M., Prescimone, A., Neuburger, M., Constable, E. C. and Housecroft, C. E. (2017). *CrystEngComm*, **19**(21), 2894–2902.

Kohn, W. and Sham, L. J. (1965). *Phys. Rev.* **140**(4A), A1133–A1138.

Kottke, T. and Stalke, D. (1993). *J. Appl. Cryst.* **26**(4), 615–619.

Krause, L. (2017). *Assessment of Single Crystal X-Ray Diffraction Data Quality*. Phd thesis, Georg-August-Universität, Göttingen.

Krause, L., Herbst-Irmer, R., Sheldrick, G. M. and Stalke, D. (2015). *J. Appl. Cryst.* **48**(1), 3–10.

Krzywinski, M. and Altman, N. (2014). *Nat. Methods*, **11**(2), 119–120.

Kuhs, W. F. (1992). *Acta Cryst. A*, **48**(2), 80–98.

Langmann, J., Vöst, M., Schmitz, D., Haas, C., Eickerling, G., Jesche, A., Nicklas, M., Lanza, A., Casati, N., Macchi, P. and Scherer, W. (2021). *Phys. Rev., B*, **103**(18).

Larsson, D. H., Takman, P. A. C., Lundström, U., Burvall, A. and Hertz, H. M. (2011). *Rev. Sci. Instrum.* **82**(12), 123701.

Lee, K., Murray, É. D., Kong, L., Lundqvist, B. I. and Langreth, D. C. (2010). *Phys. Rev., B*, **82**(8).

Li, C., Hanif, M., Li, X., Zhang, S., Xie, Z., Liu, L., Yang, B., Su, S. and Ma, Y. (2016). *J. Mater. Chem. C*, **4**(31), 7478–7484.

Lippmann, T., Blaha, P., Andersen, N. H., Poulsen, H. F., Wolf, T., Schneider, J. R. and Schwarz, K.-H. (2003). *Acta Cryst. A*, **59**(5), 437–451.

Macchi, P. (2020). *Crystallogr. Rev.* **26**(4), 209–268.

Macchi, P. and Coppens, P. (2001). *Acta Cryst. A*, **57**(6), 656–662.

Madsen, A. Ø., Mason, S. and Larsen, S. (2003). *Acta Cryst. B*, **59**(5), 653–663.

Madsen, A. Ø., Sørensen, H. O., Flensburg, C., Stewart, R. F. and Larsen, S. (2004). *Acta Cryst. A*, **60**(6), 550–561.

Malaspina, L. A., Edwards, A. J., Woińska, M., Jayatilaka, D., Turner, M. J., Price, J. R., Herbst-Irmer, R., Sugimoto, K., Nishibori, E. and Grabowsky, S. (2017). *Cryst. Growth Des.* **17**(7), 3812–3825.

Malaspina, L. A., Genoni, A., Jayatilaka, D., Turner, M. J., Sugimoto, K., Nishibori, E. and Grabowsky, S. (2021). *J. Appl. Cryst.* **54**(3), 718–729.

Malaspina, L. A., Hoser, A. A., Edwards, A. J., Woińska, M., Turner, M. J., Price, J. R., Sugimoto, K., Nishibori, E., Bürgi, H.-B., Jayatilaka, D. and Grabowsky, S. (2020). *CrystEngComm*, **22**(28), 4778–4789.

Malaspina, L. A., Wieduwilt, E. K., Bergmann, J., Kleemiss, F., Meyer, B., Ruiz-López, M. F., Pal, R., Hupf, E., Beckmann, J., Piltz, R. O., Edwards, A. J., Grabowsky, S. and Genoni, A. (2019). *J. Phys. Chem. Lett.* **10**(22), 6973–6982.

Medvedev, M. G., Bushmarinov, I. S., Sun, J., Perdew, J. P. and Lyssenko, K. A. (2017). *Science*, **355**(6320), 49–52.

Mei, J., Hong, Y., Lam, J. W. Y., Qin, A., Tang, Y. and Tang, B. Z. (2014). *Adv. Mater.* **26**(31), 5429–5479.

Meindl, K. and Henn, J. (2008). *Acta Cryst. A*, **64**(3), 404–418.

Meurer, F., Dolomanov, O. V., Hennig, C., Peyerimhoff, N., Kleemiss, F., Puschmann, H. and Bodensteiner, M. (2022). *IUCrJ*, **9**(5), 604–609.

Minyaev, R. M. and Gribanova, T. N. (2005). *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **54**(3), 533–546.

Monkhorst, H. J. and Pack, J. D. (1976). *Phys. Rev., B*, **13**(12), 5188–5192.

Neese, F. (2018). *WIREs Comput. Mol. Sci.* **8**(1).

Niepötter, B., Herbst-Irmer, R. and Stalke, D. (2015). *J. Appl. Cryst.* **48**(5), 1485–1497.

Otendal, M., Tuohimaa, T. and Hertz, H. M. (2007). *J. Appl. Phys.* **101**(2), 026102.

Otendal, M., Tuohimaa, T., Vogt, U. and Hertz, H. M. (2008). *Rev. Sci. Instrum.* **79**(1), 016102.

Pastore, Smargiassi and Buda (1991). *Phys. Rev., A*, **44**(10), 6334–6347.

Peng, X.-L., Ruiz-Barragan, S., Li, Z.-S., Li, Q.-S. and Blancafort, L. (2016). *J. Mater. Chem. C*, **4**(14), 2802–2810.

Prince, E., Fuess, H., Hahn, T., Wondratschek, H., Müller, U., Shmueli, U., Authier, A., Kopský, V., Litvin, D. B., Rossmann, M. G., Arnold, E., Hall, S. and McMahon, B. (2006). *International Tables for Crystallography, Volumes 1-8 (5th edition)*, vol. C. Chester, England and New York: International Union of Crystallography and Springer, 1st ed.

Rajak, S., Schott, O., Kaur, P., Maris, T., Hanan, G. S. and Duong, A. (2020). *Polyhedron*, **180**, 114412.

Rohrmoser, B., Eickerling, G., Presnitz, M., Scherer, W., Eyert, V., Hoffmann, R.-D., Rodewald, U. C., Vogt, C. and Pöttgen, R. (2007). *J. Am. Chem. Soc.* **129**(30), 9356–9365.

Röntgen, W. C. (1895). *Sitzungsberichte der Würzburger physik.-med. Gesellschaft*, **137**, 132–141.

Otero-de-la Roza, A. and Johnson, E. R. (2012*a*). *J. Chem. Phys.* **137**(5), 054103.

Otero-de-la Roza, A. and Johnson, E. R. (2012*b*). *J. Chem. Phys.* **137**(5), 054103.

Ruth, P. N., Herbst-Irmer, R. and Stalke, D. (2022). *IUCrJ*, **9**(2), 286–297.

Sabatini, R., Gorni, T. and de Gironcoli, S. (2013). *Phys. Rev., B*, **87**(4), 041108.

Salvatier, J., Wiecki, T. V. and Fonnesbeck, C. (2016). *PeerJ Comput. Sci.* **2**, e55.

Sanjuan-Szklarz, W. F., Woińska, M., Domagała, S., Dominiak, P. M., Grabowsky, S., Jayatilaka, D., Gutmann, M. and Woźniak, K. (2020). *IUCrJ*, **7**(5), 920–933.

Sapunov, V. N., Kirchner, K. and Schmid, R. (2001). *Coord. Chem. Rev.* **214**(1), 143–185.

Schillmöller, T. (2022). *Structure-Property Correlations in Luminescent Anthracene Derivatives*. Doktorarbeit, Georg-August-Universität, Göttingen.

Schillmöller, T., Ruth, P. N., Herbst-Irmer, R. and Stalke, D. (2020*a*). *Chem. Eur. J.* **26**(72), 17390–17398.

Schillmöller, T., Ruth, P. N., Herbst-Irmer, R. and Stalke, D. (2020*b*). *Chem. Comm.* **56**(54), 7479–7482.

Schoeller, W. W., Friedrich, O., Sundermann, A. and Rozhenko, A. (1999). *Organometallics*, **18**(11), 2099–2106.

Schulz, T., Meindl, K., Leusser, D., Stern, D., Graf, J., Michaelsen, C., Ruf, M., Sheldrick, G. M. and Stalke, D. (2009). *J. Appl. Cryst.* **42**(5), 885–891.

Schwab, G. (2004). *Fluoreszenz von Wirt/Gast-Komplexen der (Di-)-Phosphinylanthracene*. Diploma thesis, Julius-Maximilians-Universität, Würzburg.

Schwab, G. (2008). *Die Strukturabhängigkeit der Festkörperfluoreszenz bei Phosphanylanthracenen*. Dissertation, Georg-August-Universität, Göttingen.

Schwarzenbach, D. (1996). *Crystallography*. Chichester, England: John Wiley & Sons.

Schäfer, A., Horn, H. and Ahlrichs, R. (1992). *J. Chem. Phys.* **97**(4), 2571–2577.

Scibelli, J. V. and Curtis, M. D. (1973). *J. Am. Chem. Soc.* **95**(3), 924–925.

Sevvana, M., Ruf, M., Usón, I., Sheldrick, G. M. and Herbst-Irmer, R. (2019). *Acta Cryst. D*, **75**(12), 1040–1050.

Shanno, D. F. (1970). *Math. Comput.* **24**(111), 647–656.

Sheldrick, G. M. (2015*a*). *Acta Cryst. C*, **71**(1), 3–8.

Sheldrick, G. M. (2015*b*). *Acta Cryst. A*, **71**, 3–8.

Smith, D. G. A., Burns, L. A., Patkowski, K. and Sherrill, C. D. (2016). *J. Phys. Chem. Lett.* **7**(12), 2197–2203.

Smith, J. D. and Hanusa, T. P. (2001). *Organometallics*, **20**(14), 3056–3062.

Stalke, D. (1998). *Chem. Soc. Rev.* **27**, 171–178.

Stash, A., (2007). Drkplot.

Stern, D. (2006). *Synthesestrategien zu asymmetrisch substituierten Phosphanylanthracenen*. Diploma thesis, Georg-August-Universität, Göttingen.

Stern, D. (2009). *Fluoreszenzchemosensoren auf Basis des Anthracen-Fluorophors*. Phd thesis, Georg-August-Universität, Göttingen.

Stewart, R. (1976). *Acta Cryst. A*, **32**(4), 565–574.

Stewart, R. F. (1969). *J. Chem. Phys.* **51**(10), 4569–4577.

Stewart, R. F. (1973). *J. Chem. Phys.* **58**(4), 1668–1676.

Su, Z. and Coppens, P. (1998). *Acta Cryst. A*, **54**(5), 646–652.

Swaminathan, S., Craven, B. M. and McMullan, R. K. (1984). *Acta Cryst. B*, **40**(3), 300–306.

Thonhauser, T., Zuluaga, S., Arter, C. A., Berland, K., Schröder, E. and Hyldgaard, P. (2015). *Phys. Rev. Lett.* **115**(13), 136402.

Verstraelen, T., Tecmer, P., Heidar-Zadeh, F., Boguslawski, K., Chan, M., Zhao, Y., Kim, T. D., Vandenbrande, S., Yang, D., González-Espinoza, C. E., Fias, S., Limacher, P. A., Berrocal, D., Malek, A. and Ayers, P. W., (2015). Horton 2.0.1.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F. and van Mulbregt, P. (2020). *Nat. Methods*, **17**(3), 261–272.

Volkov, A., Abramov, Y., Coppens, P. and Gatti, C. (2000). *Acta Cryst. A*, **56**(4), 332–339.

Volkov, A., Macchi, P., Farrugia, L. J., Gatti, C., Mallinson, P. R., Richter, T. and Koritsanszky, T., (2016). Xd2016: A computer program package for multipole refinement, topological analysis of charge densities and evaluation of intermolecular energies from experimental and theoretical structure factors.

Wall, M. E. (2016). *IUCrJ*, **3**(4), 237–246.

Waller, I. t. and Hartree, D. R. (1929). *Proc. R. Soc. Lond. A*, **124**(793), 119–142.

Weiss, M. S. (2001). *J. Appl. Cryst.* **34**(2), 130–135.

Whitten, A. E. and Spackman, M. A. (2006). *Acta Cryst. B*, **62**(5), 875–888.

Wieduwilt, E. K., Macetti, G. and Genoni, A. (2021). *J. Phys. Chem. Lett.* **12**(1), 463–471.

Wieduwilt, E. K., Macetti, G., Malaspina, L. A., Jayatilaka, D., Grabowsky, S. and Genoni, A. (2020). *J. Mol. Struct.* **1209**, 127934.

Wilson, A. J. C. (1942). *Nature*, **150**(3796), 152.

Woińska, M., Jayatilaka, D., Spackman, M. A., Edwards, A. J., Dominiak, P. M., Woźniak, K., Nishibori, E., Sugimoto, K. and Grabowsky, S. (2014). *Acta Cryst. A*, **70**(5), 483–498.

Zhang, J., Song, L., Sist, M., Tolborg, K. and Iversen, B. B. (2018). *Nat. Commun.* **9**(1), 1–10.

Zhang, L., Fehse, C., Eckert, H., Vogt, C., Hoffmann, R.-D. and Pöttgen, R. (2007). *Solid State Sci.* **9**(8), 699–705.

A. Determination of X-ray Data

A.1. Measurement and Refinement of temperature-dependent SPAnPS/toluene data

The temperature-dependent measurement of SPAnPS was conducted as follows. Suitable crystals were coated in perfluorinated oil and mounted on a MiTeGen crystal loop. Crystal data was recorded using an Incoatec Mo Microsource (Schulz *et al.*, 2009) and an APEX II detector. Data was integrated with SAINT (Bruker AXS Inc., 2019) and corrected using the multiscan absorption correction in SADABS (Krause *et al.*, 2015). The structure was solved once for the data at 100 K using SHELXT (Sheldrick, 2015*b*) and all the datasets were refined on F^2 using SHELXL (Sheldrick, 2015*a*) in the graphical user interface ShelXle (Hübschle *et al.*, 2011). Using the respective cell constants and diffraction data this model was then fitted to each dataset for the individual temperatures.

In the HAR conducted in Olex2 (Dolomanov *et al.*, 2009) using NoSpherA2 (Kleemiss *et al.*, 2021), atoms in the SPAnPS molecule were refined freely including hydrogen atoms with isotropic displacement parameters, while the toluene guest molecule was refined using restraints. Carbon-carbon and carbon-hydrogen distances as well as angles within the aromatic ring were restrained to be similar, as were carbon-carbon and carbon-hydrogen distances to and within the methyl groups. Atomic displacement parameters were restrained to be similar to nearby atoms and follow a rigid-bond model to bound atoms. All carbon atoms, as well as hydrogen atoms connected to the aromatic ring, were restrained to lie in a common plane. The methyl hydrogen atoms 1,3 distance to the C_α was restrained to 2.056 Å to provide further stability to the model.

A.2. Measurement of decamethylsilicocene data

For all measurements, suitable crystals were coated in perfluorinated oil and mounted on a MiTeGen crystal loop under cooling using an XTemp2 device (Stalke, 1998; Kottke and Stalke, 1993).

The initial measurement which lead to the discovery of the low-temperature phase was measured using a first crystal on a Bruker AXS rotating anode using an APEX II detector with a resolution of $d_{\min} = 0.6$ Å and a measurement temperature of 90 K. The

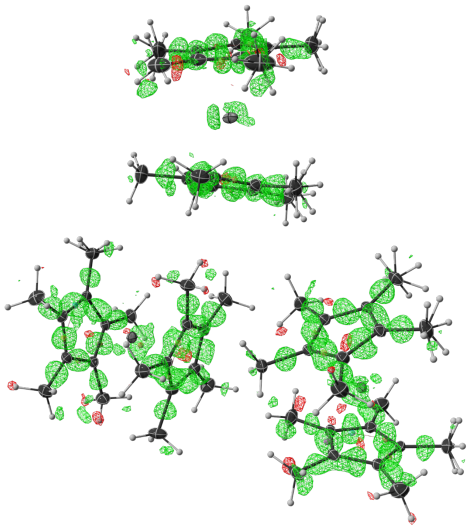


Figure A.1.: Crystal structure derived from the first crystal at 90 K, with the difference electron density being depicted at an iso-level of $\pm 0.2 e\,\text{Å}^{-3}$. Atomic displacements are depicted at 50 % probability level. Hydrogen positions were refined disordered where necessary and all hydrogen atoms were placed using idealised geometries in a riding model.

difference electron density derived from this structure contained the main maximum difference electron density on the bonds with the exception of a maximum of $0.55\,e\,\text{Å}^{-3}$ near Si2 (see Figure A.1).

For further investigation, a second crystal was subjected to a set of measurements for the listed different measurement temperatures where the resolution was set to $d_{\text{min}} = 0.75\,\text{Å}$ to assure reasonable diffraction up to the resolution limit for all temperatures. The setup itself was retained, however, the X-ray detector had been changed to a Photon III detector. Unfortunately, the crystal quality seems to have deteriorated between the two measurements, as indicated by the occurrence of additional maxima in the difference electron density for all temperature-dependent measurements. The size of the maxima as depicted in Figure A.2 are as follows: two maxima near Si1, 1.40 and 1.24 e, the maximum near Si2 has increased to 1.23 e, and four maxima of approximately 0.5 e near Si3. These features could neither be explained with twinning, as no twin matrix could be derived nor with a disorder of chemically reasonable geometry.

However, an overlay of the structures determined at 90 K from the first and the second crystal reveals no influence of the geometrical parameters (Figure A.3). The geometric considerations can therefore be considered valid.

All data was integrated with SAINT (Bruker AXS Inc., 2019) and corrected using the multiscan absorption correction in SADABS (Krause *et al.*, 2015). SHELXT (Sheldrick, 2015*b*) was used to solve the structures once for the data from crystal one at 90 K for

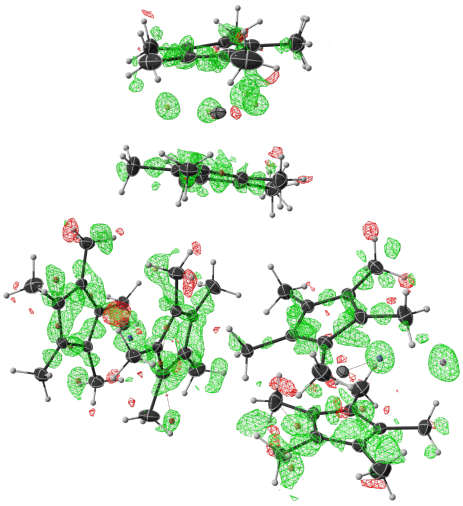


Figure A.2.: Crystal structure derived from the first crystal at 90 K, with the difference electron density being depicted at an iso-level of $\pm 0.2e\,\text{Å}^{-3}$. Silicon atoms are numbered anti-clockwise with the silicon atom at the lower right being denoted Si1. Atomic displacements are depicted at 50 % probability level. Hydrogen positions were refined disordered where necessary and all hydrogen atoms were placed using idealised geometries in a riding model.

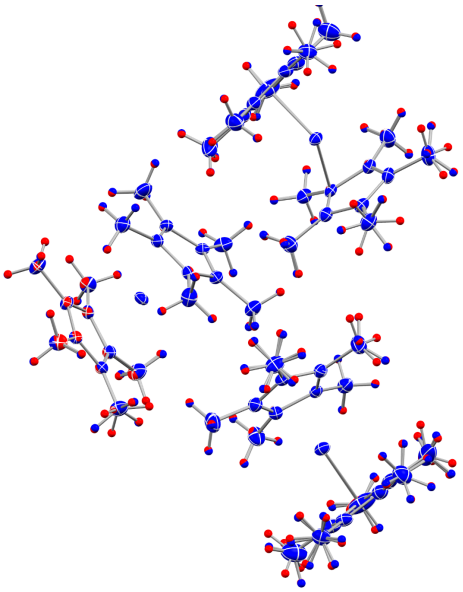


Figure A.3.: Overlay of the structures at 90 K derived from crystal 1 (red) and crystal 2 (blue). For most of the structure, no red thermal ellipsoids are visible as they are completely contained within the blue ellipsoids. All ellipsoids are depicted at 50 % probability level.

the structure in $P\,2_1/c$ and once from the data at 298 K for the structure in $C\,2/c$. Both structures were subsequently refined on F^2 using SHELXL (Sheldrick, 2015*a*) in the graphical user interface ShelXle (Hübschle *et al.*, 2011). Using the respective cell constants and diffraction data this model was then fitted to each dataset for the individual temperatures. Hydrogen atoms were placed in idealised geometries using a riding model and disorder between hydrogen atoms was refined when needed, where the number of disordered hydrogen groups changed with temperature.

A.3. Measurement details and Data Processing for the evaluation of the MetalJet diffractometer

A.3.1. Implementation of the Eiger II CdTe in the Bruker D8 Enclosure

While the overall results of the investigation of the Eiger2 1M CdTe on the MetalJet induced a more close implementation that is currently in progress, the evaluations were done on a more custom implementation. In general, this had two aspects: The implementation needed to be done in hardware and software. The mounting of the detector had been developed by Nico Graw in cooperation with the precision engineering workshop. The second challenge was to synchronise the triggering of the Eiger2 to the movement and data collection and frame collection of the machine. The basic approach was to not let the machine itself know that anything was different. The Photon III remained in the enclosure and therefore one part of the puzzle were 'sfrm' frame files, which contained no intensity information but the goniometer angle information at the time of collection. In cooperation with the electronics workshop in Göttingen, a trigger box was developed that combined signals from the enclosure to achieve a clean signal ultimately triggering the Eiger2 detector at the correct time. This now enabled the collection of the X-ray intensities from the hardware side of things.

The Eiger2 can be easily programmed to collect a set of identical frames and react to multiple triggers to collect identical runs. However, a typical set of runs on a Bruker instrument does not contain runs with identical numbers of frames. Additionally, high-resolution X-ray diffraction experiments typically contain multiple different exposure times to collect the weak outer reflections while also determining the very strong inner reflections without overexposure. As such a python script was written, which read a measurement strategy in the form of a Bruker .exp file and fed the necessary information to the Eiger2 detector during the measurement.

To develop the necessary strategy the approach was as follows. First, a new strategy was created from the measured cell constants and crystal orientation. After Apex4 had written the necessary mask files for the corresponding angles, these were then modified to create an Eiger2-1M-shaped set of valid pixels, while all other pixels were masked invalid. Using these a new strategy was created. In case new angles were included and therefore new mask files were created, the process was repeated until convergence.

After the measurement was finished the two sets of frames were then combined using another script, therefore enabling the comparison of the two detectors on equal footing.

A.3.2. Data processing for the Independent Atom Model investigation

The intensity of the Bragg reflections on the collected frames was integrated using SAINT (Bruker AXS Inc., 2019). While the refinement of the data of the Photon III detector was integrated using defaults, the profile fitting was deactivated for the integration of the data on the Eiger 2 CdTe detector. Absorption correction and scaling were done in SADABS (Krause *et al.*, 2015), where the radius for the absorption correction was assumed to be half of the lowest crystal dimension (Krause *et al.*, 2015). Structure solutions were obtained using SHELXT (Sheldrick, 2015*b*). Structure refinement was conducted against F^2 using SHELXL (Sheldrick, 2015*a*). While initial refinement was done using the graphical user interface ShelXle (Hübschle *et al.*, 2011), subsequent data processing was automated using a script created in PYTHON.

The comparison was done once against the complete collected data up to the given resolution. In addition, refinements were conducted against datasets, which were multiplicity normalised. To this end, sets of symmetry equivalent reflections of equivalent exposure time were created from the measured reflections. For a given symmetry equivalent reflection and exposure time, the sets from both detectors were compared. While the smaller set was included completely, the size of the larger set was reduced by taking a random sample. This procedure was then continued until new sets had been created for both of the detectors. To quantify the statistical variation resulting from this procedure, it was repeated 100 times. This enabled the calculation of mean values and their standard deviation from the refinement against these 100 sets.

A.3.3. Data processing for the refinement of aspherical atom models

Again, integration was conducted using SAINT (Bruker AXS Inc., 2019), while absorption correction and scaling were done in SADABS (Krause *et al.*, 2015). For the multipole refinement, it has been shown that thermal diffuse scattering can skew the results Niepötter *et al.* (2015), which can be corrected by using a resolution-dependent scaling. Accordingly, the intensity within the YLIDE measurements was corrected using the described method, by applying the formula.

$$I_{\mathrm{corr}} = \frac{I_{\mathrm{meas}}}{1 + a\left[\frac{\sin(\theta)}{\lambda}\right]^2 + b\left[\frac{\sin(\theta)}{\lambda}\right]^3} \tag{A.1}$$

The determined values for the correction parameters a/b were 0.13/0.76 for the indium/Photon III dataset, -0.08/1.74 for the indium/Eiger2 CdTe and 0.15/1.07 for the silver/Photon III dataset.

Instead of solving the structures from scratch, a presolved model was refined against the collected data using SHELXL (Sheldrick, 2015*a*) using again the graphical user interface ShelXle (Hübschle *et al.*, 2011). Subsequent multipolar refinement was conducted in XD2016 (Volkov *et al.*, 2016), while HAR used the methodology described in Chapter 6. PAW-DFT calculations were done in GPAW (Hjorth Larsen *et al.*, 2017) at the Γ-point using the SCAN functional with grid spacings of 0.16 / 0.08 / 0.04 Å (wavefunction / density / FFT).

B. Additional Details and Verification

B.1. Additional information concerning the evaluation and experiments on fluorescence

B.1.1. Additional implemented features

XE-plane angle This is the angle between the mean plane of the anthracene molecule and the vector from the connected phosphorous or silicon atom (E) and the oxidising atom (X), which is usually an oxygen atom, a sulphur atom or if E is a silicon atom it consists of a chlorine atom.

XEEX torsion angle and XEEX torsion angle on CC These two indicators are only evaluated if the molecule in question has oxidised substituents on opposing sides of the anthracene molecule.

The first of the two indicators is simply the torsion angle between the substituent atoms directly connected to the anthracene E (in most cases phosphorous) and the oxidising element X (usually oxygen or sulphur).

The second angle is determined by determining the vector between the opposing carbon atoms, where the substituents are attached (*e. g.* the 9 and 10 positions). The two E-X vectors are projected onto the plane perpendicular to the carbon-carbon vector. The given angle is the angle between these two projections.

C-E distance This is simply the distance between the substituent directly connected to the anthracene molecule. Only phosphorous and silicon are currently evaluated

XE distance This is simply the distance between the phosphorous or silicon directly connected to the anthracene molecule and the oxidising atom.

Wilson $\overline{B}$: This descriptor can be calculated directly from the X-ray intensities from the diffraction experiment. Wilson (1942) could show that a mean isotropic displacement parameter can be calculated by the fit.

$$\ln\left(\langle I \rangle\right) = \frac{-B\left(\sin(\theta)/\lambda\right)^2}{2} + K' \tag{B.1}$$

where B is the desired parameter. The evaluation is done in resolution shells and $\overline{B}$ is the average over these resolution shells.

B.1.2. Trace and posterior values for the Bayesian fit in PyMC

To check convergence the trace of the No-U-Turn sampler should be indistinguishable from noise for all parameters. The trace plot is depicted in Figure B.1.

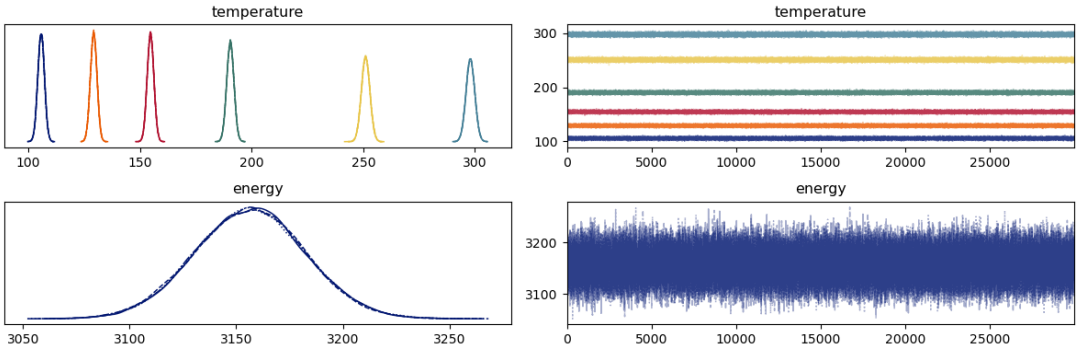


Figure B.1.: Trace plot for the Bayesian Boltzmann fit.

B.2. Difference electron density SPAnPS 298K

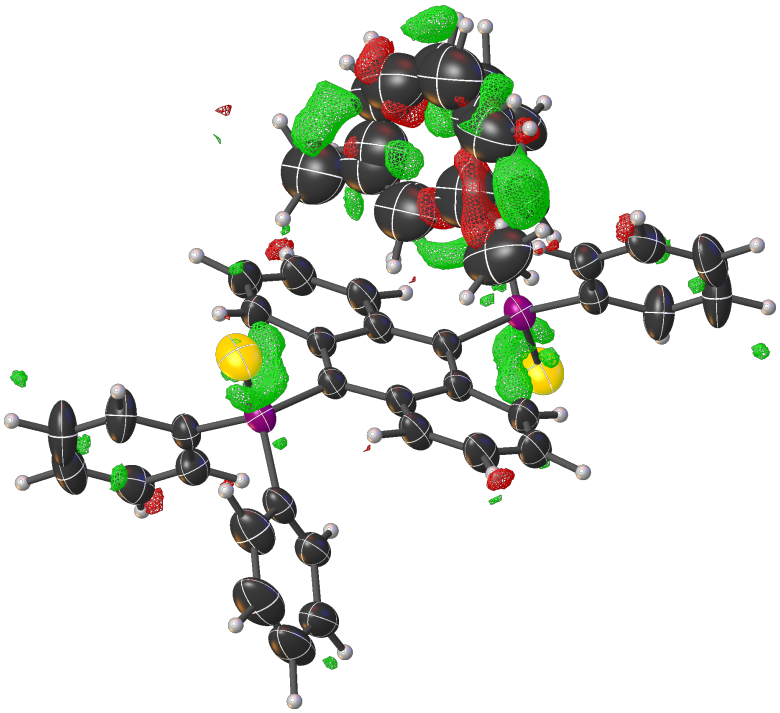


Figure B.2.: Molecular structure of 9,10-SPAnPS. Atomic displacement parameters are depicted at the 50 % probability level. The difference electron density is displayed at an iso-level of $\pm 0.133\, e\, \text{Å}^{-3}$.

B.3. Additional information on Decamethylsilicocene

Table B.1.: Value for the interplanar angle α for the given temperatures. While α_1 and α_3 are symmetry equivalent at high temperature, α_2 belongs to the molecule adopting an inversion symmetric conformation at high temperature

T / K	α_1 / °	α_2 / °	α_3 / °
80(2)	29.31(10)	20.40(10)	27.96(11)
85(2)	29.14(10)	19.80(10)	27.91(11)
90(2)	28.96(12)	18.78(11)	27.84(12)
95(2)	28.77(12)	17.17(12)	27.77(12)
100(2)	28.41(14)	13.88(14)	27.82(14)
105(2)	28.17(16)	8.47(18)	27.77(16)
110(2)	27.93(16)	5.13(18)	27.87(16)
120(2)	27.85(16)	2.83(19)	27.72(16)
130(2)	27.51(17)	2.18(19)	27.84(16)
140(2)	27.64(17)	1.38(19)	27.52(17)
150(2)	27.44(18)	1.25(20)	27.53(17)
160(2)	27.59(18)	0.8(2)	27.15(18)

Table B.2.: Temperature dependence of the unit cell lengths and the $R(F)$ for refinement in the low-temperature or high-temperature phase.

T / K	a / Å	b / Å	c / Å	V / Å^3	$\frac{R(F)}{\%}$ ($P\,\frac{2_1}{c}$)	$\frac{R(F)}{\%}$ ($C\,\frac{2}{c}$)
80(2)	30.885(3)	8.400(2)	22.997(2)	5636.4(17)	6.44	10.05
85(2)	30.896(3)	8.402(2)	23.010(2)	5642.9(17)	6.91	10.04
90(2)	30.908(3)	8.405(2)	23.022(2)	5650.1(17)	7.96	9.73
95(2)	30.915(3)	8.408(2)	23.039(2)	5657.6(17)	8.10	8.82
100(2)	30.917(3)	8.409(2)	23.063(2)	5664.5(17)	9.73	8.18
105(2)	30.915(3)	8.410(2)	23.081(2)	5669.6(17)	11.69	7.98
110(2)	30.929(3)	8.410(2)	23.093(2)	5674.7(17)	12.01	7.54
120(2)	30.961(3)	8.415(2)	23.104(2)	5686.7(17)	12.12	7.27
130(2)	31.001(3)	8.420(2)	23.112(2)	5699.4(17)	12.07	7.19
140(2)	31.038(3)	8.425(2)	23.123(2)	5712.6(17)	12.15	7.15
150(2)	31.077(3)	8.430(2)	23.130(2)	5724.9(17)	12.36	7.20
160(2)	31.119(3)	8.435(2)	23.140(2)	5738.9(17)	12.25	7.19

B.4. Additional Information concerning the MetalJet

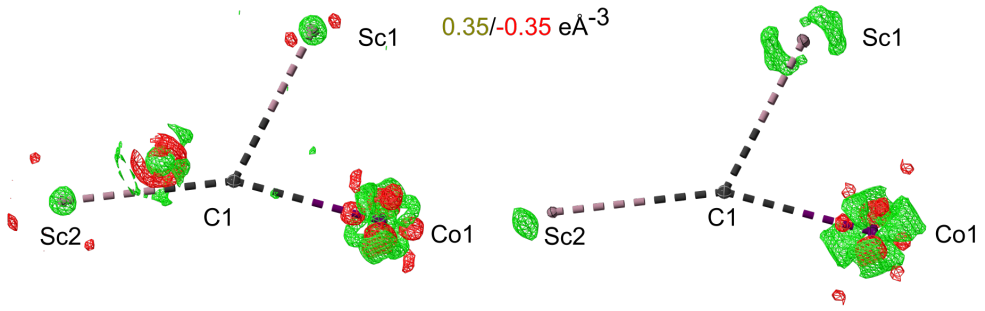


Figure B.3.: Difference electron density for **1** in dependence of the detector from all collected data. The Photon III result is depicted on the left, the Eiger2 CdTe is depicted on the right. ADPs are depicted at the 50 % probability level.

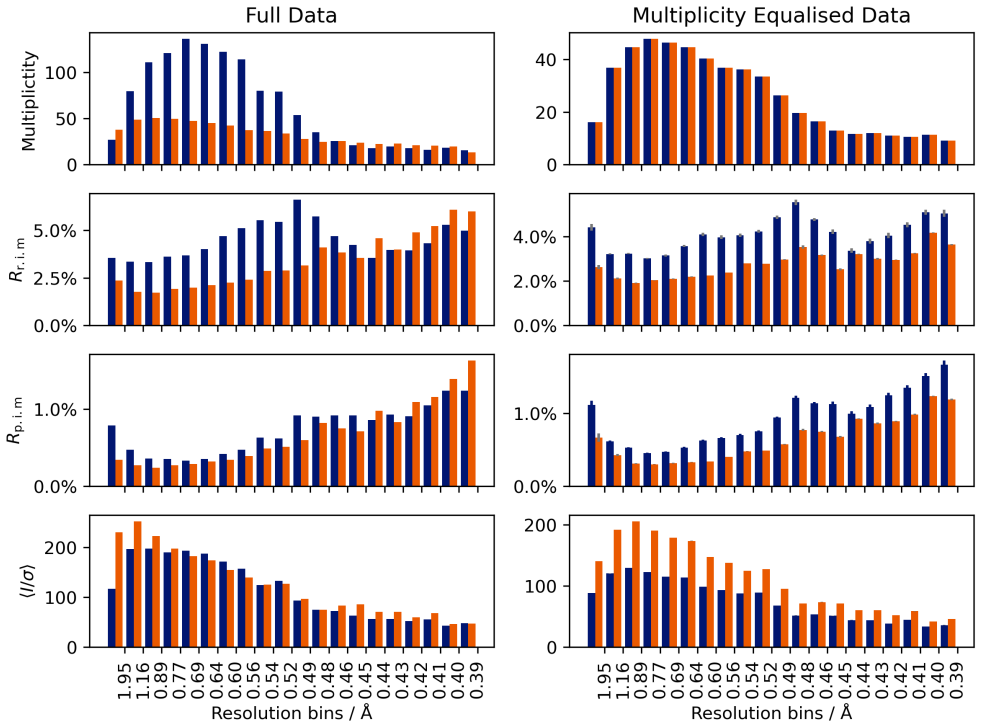


Figure B.4.: Multiplicity and precision indicating quality indicators for the data of **1** as calculated by XPREP in comparison between the Photon III (blue) and the Eiger2 CdTe (orange). Grey indicators denote the standard deviation of the 100 drawn sets but are often below the visible threshold.

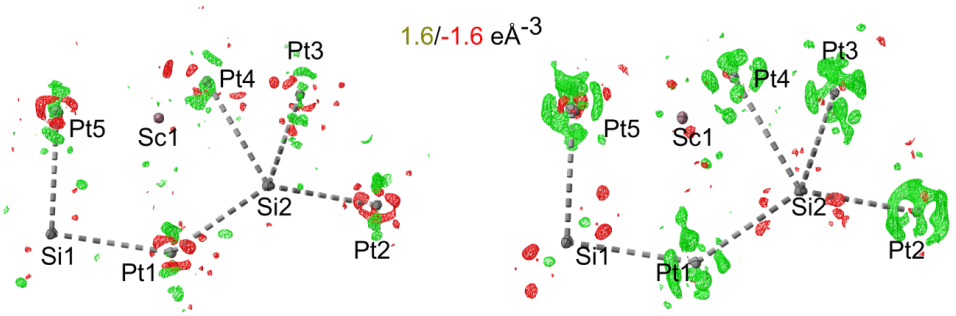


Figure B.5.: Difference electron density for **2** in dependence of the detector from all collected data. The Photon III result is depicted on the left, the Eiger2 CdTe is depicted on the right. ADPs are depicted at the 50 % probability level.

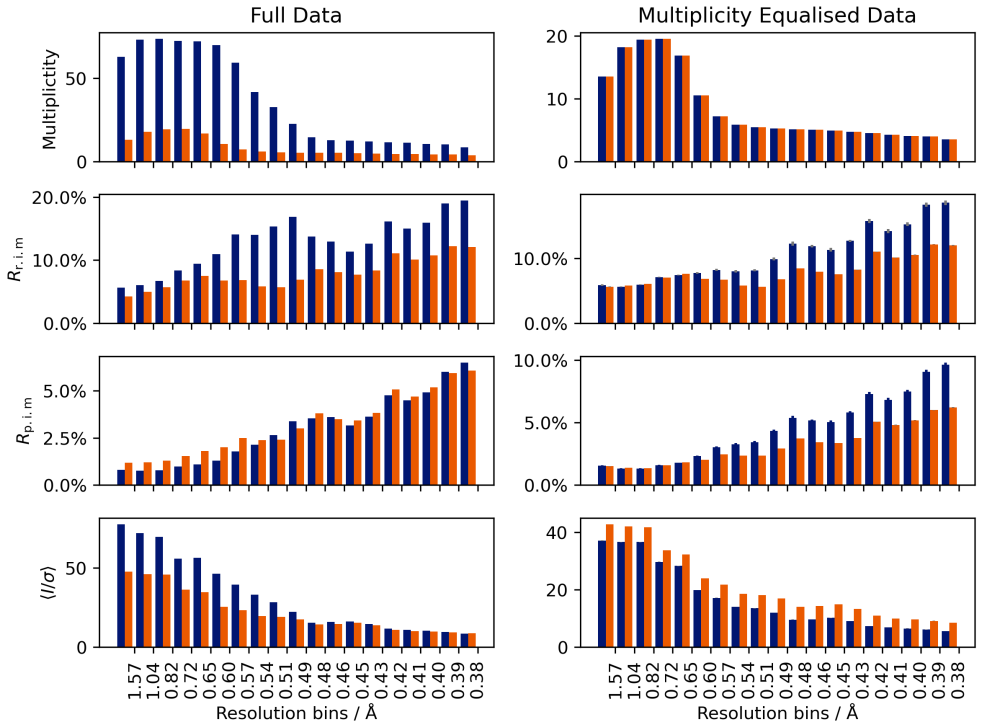


Figure B.6.: Multiplicity and precision indicating quality indicators for the data of **2** as calculated by XPREP in comparison between the Photon III (blue) and the Eiger2 CdTe (orange). Grey indicators denote the standard deviation of the 100 drawn sets but are often below the visible threshold.

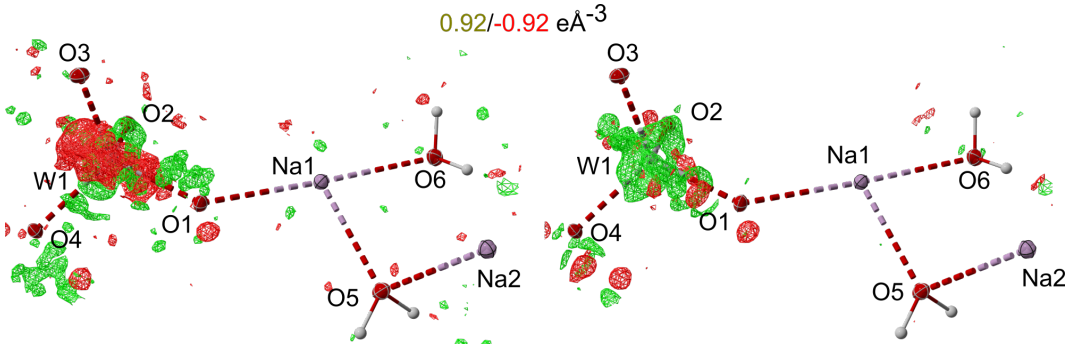


Figure B.7.: Difference electron density for **3** in dependence of the detector from all collected data. The Photon III result is depicted on the left, the Eiger2 CdTe is depicted on the right. ADPs are depicted at the 50 % probability level.

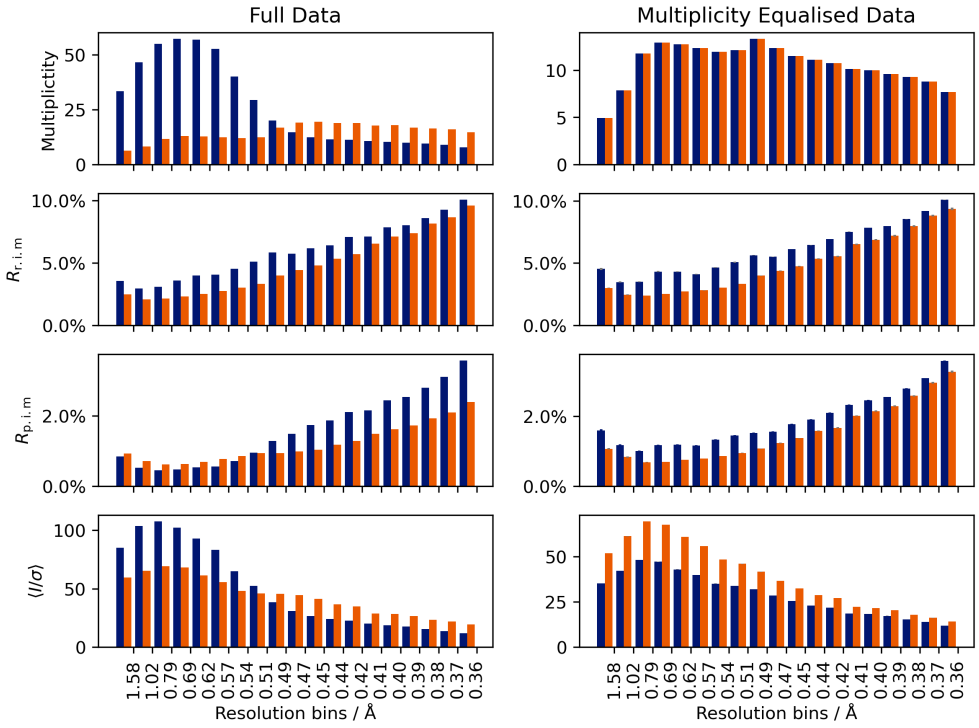


Figure B.8.: Multiplicity and precision indicating quality indicators for the data of **3** as calculated by XPREP in comparison between the Photon III (blue) and the Eiger2 CdTe (orange). Grey indicators denote the standard deviation of the 100 drawn sets but are often below the visible threshold.

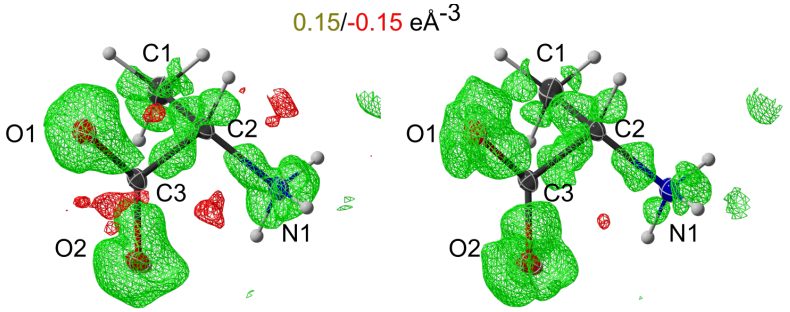


Figure B.9.: Difference electron density for **4** in dependence of the detector from all collected data. The Photon III result is depicted on the left, the Eiger2 CdTe is depicted on the right. ADPs are depicted at the 50 % probability level.

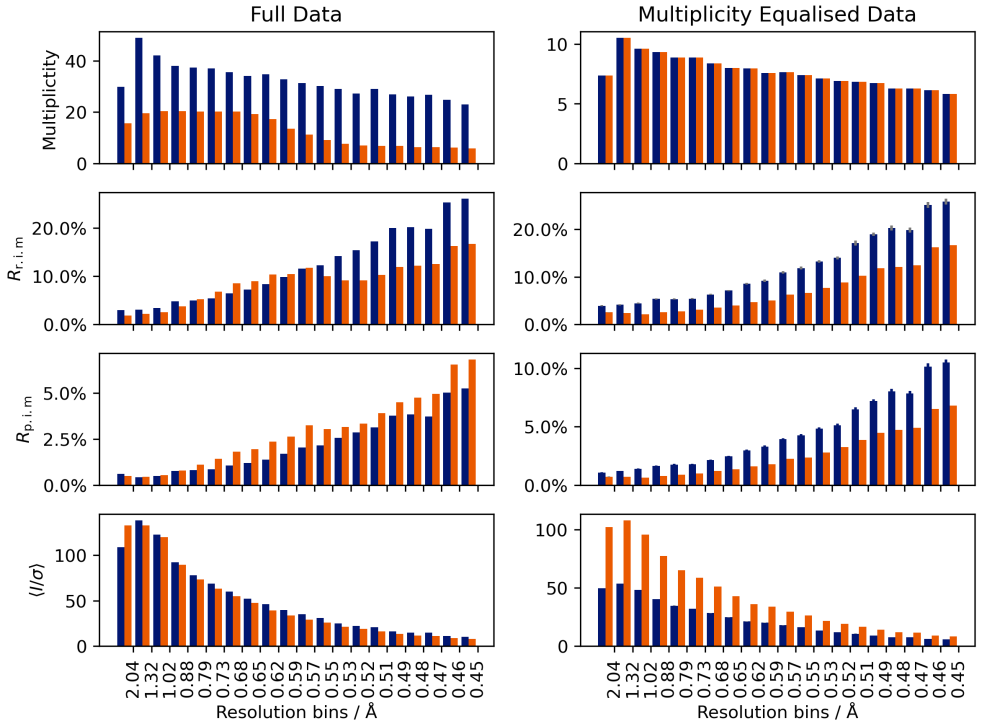


Figure B.10.: Multiplicity and precision indicating quality indicators for the data of **4** as calculated by XPREP in comparison between the Photon III (blue) and the Eiger2 CdTe (orange). Grey indicators denote the standard deviation of the 100 drawn sets but are often below the visible threshold.

B.5. Additional Information for the PAW-HAR evaluations

Table B.3.: Evaluated crystallographic agreement factor $wR_2(F^2)$ in dependence of functional and crystal environment description given in mÅ. Periodic PAW calculations were done in XHARPy using GPAW and the settings given for the evaluation in Chapter 6. The cluster charge environment was calculated in TONTO, and the calculation without consideration of the environment was calculated in ORCA, both using NoSpherA2 in Olex2.

environment	dataset / functional	A23K	HMa-8HQ	HMa-Mg	Xy	Urea
Periodic PAW	PW	3.31	4.11	3.27	2.35	1.97
	BLYP	3.20	4.08	3.20	2.24	1.89
	PW91	3.17	4.07	3.06	2.23	1.85
	PBE	3.16	4.07	3.04	2.22	1.83
	revPBE	3.12	4.06	2.96	2.19	1.80
	RPBE	3.11	4.06	2.94	2.18	1.79
	TPSS	3.14	4.05	3.00	2.21	1.80
	SCAN	3.06	4.03	2.85	2.14	1.76
	revSCAN	3.03	4.04	2.78	2.12	1.77
	vdW-DF	3.14	4.07	3.08	2.20	1.82
	vdW-DF2	3.22	4.12	3.33	2.26	1.97
	BEEF-vdW	3.10	4.07	3.04	2.20	1.80
8 Å cc	B3LYP	3.06	4.07	2.91	2.12	1.80
None	PBE	3.24	4.09	3.11	2.22	2.20
	SCAN	3.12	4.05	2.96	2.14	2.11
	B3LYP	3.17	4.06	3.05	2.18	2.16

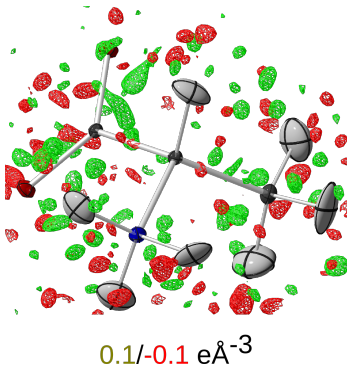


Figure B.11.: Difference electron density resulting from the PAW-HAR refinement with XHARPy/GPAW for **A23K**. ADPs are depicted at 50 % probability level.

Table B.4.: Evaluated absolute differences of X−H bond length determined from HAR and neutron diffraction ($\langle|\Delta r|\rangle$) in dependence of functional and crystal environment description given in mÅ. Periodic PAW calculations were done in XHARPy using GPAW and the settings given for the evaluation in Chapter 6. The cluster charge environment was calculated in TONTO, and the calculation without consideration of the environment was calculated in Orca, both using NoSpherA2 in Olex2. Values in brackets are not uncertainties but sample standard deviations.

environment	dataset / functional	A23K	HMa-8HQ	HMa-Mg	Xy	Urea
Periodic PAW	PW	11(11)	15(15)	7(3)	15(11)	5(1)
	BLYP	13(11)	15(14)	10(5)	16(12)	6(3)
	PW91	11(10)	13(13)	9(4)	14(11)	3(3)
	PBE	11(10)	13(13)	9(4)	14(11)	4(2)
	revPBE	11(10)	12(12)	9(4)	14(10)	3(1)
	RPBE	11(9)	12(12)	9(5)	14(10)	3(1)
	TPSS	11(10)	12(12)	9(5)	14(11)	4(3)
	SCAN	9(8)	9(9)	6(4)	11(8)	2(1)
	revSCAN	8(6)	7(7)	5(4)	9(7)	4(0)
	vdW-DF	14(10)	14(12)	11(5)	17(11)	8(3)
	vdW-DF2	16(12)	16(15)	11(5)	18(13)	13(3)
	BEEF-vdW	12(9)	12(11)	10(5)	14(10)	5(3)
4/8 Å cc	B3LYP	12(9)	13(10)	7(5)	15(11)	7(2)
None	PBE	12(10)	17(15)	15(7)	24(15)	7(1)
	SCAN	10(7)	13(11)	14(8)	20(14)	10(1)
	B3LYP	12(9)	15(13)	14(7)	22(15)	9(3)

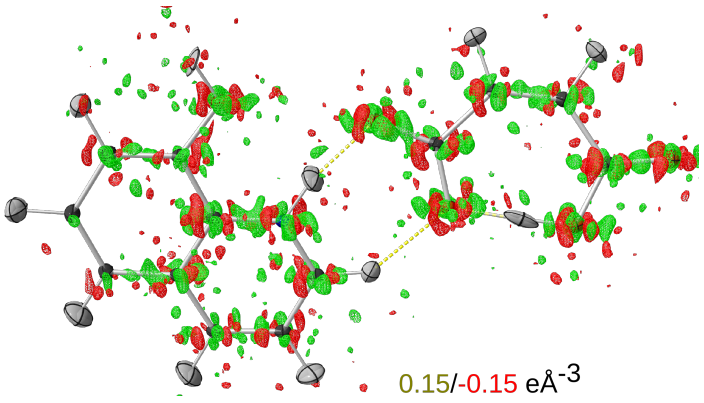


Figure B.12.: Difference electron density resulting from the PAW-HAR refinement with XHARPy/GPAW for **HMa-8HQ**. ADPs are depicted at 50 % probability level.

Table B.5.: Evaluated absolute differences of hydrogen atomic displacement parameters determined from HAR and neutron diffraction ($\langle|\Delta U_{ij}|\rangle$) in dependence of functional and crystal environment description given in 10^{-2} Å^2. Periodic PAW calculations were done in XHARPy using GPAW and the settings given for the evaluation in Chapter 6. The cluster charge environment was calculated in TONTO, and the calculation without consideration of the environment was calculated in ORCA, both using NoSpherA2 in Olex2. Values in brackets are not uncertainties but sample standard deviations.

environment	dataset / functional	A23K	HMa-8HQ	HMa-Mg	Xy	Urea
Periodic PAW	PW	0.5(3)	0.7(6)	0.4(4)	0.6(5)	0.6(3)
	BLYP	0.5(4)	0.5(5)	0.3(5)	0.6(5)	0.6(5)
	PW91	0.5(3)	0.5(5)	0.3(4)	0.5(5)	0.6(4)
	PBE	0.5(3)	0.5(5)	0.3(4)	0.5(5)	0.6(4)
	revPBE	0.5(4)	0.5(5)	0.3(4)	0.6(5)	0.6(5)
	RPBE	0.5(4)	0.5(5)	0.3(4)	0.6(5)	0.6(5)
	TPSS	0.5(4)	0.5(5)	0.3(4)	0.6(5)	0.6(5)
	SCAN	0.4(3)	0.5(4)	0.3(4)	0.5(5)	0.5(4)
	revSCAN	0.4(3)	0.5(4)	0.3(3)	0.5(5)	0.5(4)
	vdW-DF	0.5(4)	0.5(5)	0.4(5)	0.6(5)	0.7(5)
	vdW-DF2	0.6(5)	0.6(5)	0.4(5)	0.7(5)	0.7(5)
	BEEF-vdW	0.5(4)	0.5(4)	0.4(4)	0.6(5)	0.7(5)
8 Å cc	B3LYP	0.5(4)	0.5(4)	0.3(2)	0.6(5)	0.6(5)
None	PBE	0.6(6)	0.6(5)	0.4(4)	0.7(5)	0.8(8)
	SCAN	0.6(6)	0.5(5)	0.4(3)	0.6(5)	0.7(8)
	B3LYP	0.6(6)	0.5(5)	0.4(3)	0.7(5)	0.7(8)

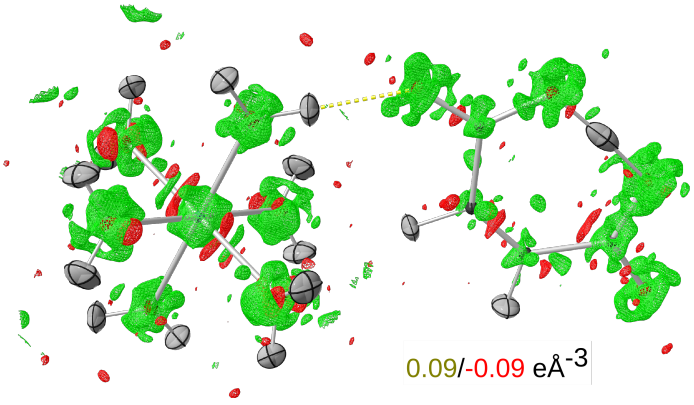


Figure B.13.: Difference electron density resulting from the PAW-HAR refinement with XHARPy/GPAW for **HMa-Mg**. ADPs are depicted at 50 % probability level.

Table B.6.: Evaluated overlap of hydrogen atomic displacement determined from HAR and neutron diffraction ($\langle S_{12} \rangle$) in dependence of functional and crystal environment description given in %. Periodic PAW calculations were done in XHARPy using GPAW and the settings given for the evaluation in Chapter 6. The cluster charge environment was calculated in TONTO, and the calculation without consideration of the environment was calculated in ORCA, both using NoSpherA2 in Olex2. Values in brackets are not uncertainties but sample standard deviations.

environment	dataset / functional	A23K	HMa-8HQ	HMa-Mg	Xy	Urea
Periodic PAW	PW	NPD	NPD	3(2)	6(5)	4.7(1.1)
	BLYP	5(3)	5(6)	2(2)	6(4)	2.8(1.3)
	PW91	5(3)	6(8)	1.7(1.6)	6(4)	2.8(8)
	PBE	5(2)	6(8)	1.6(1.6)	6(4)	2.8(9)
	revPBE	3.8(1.3)	5(7)	1.5(1.5)	5(4)	2.7(1.2)
	RPBE	4.1(1.5)	5(7)	1.4(1.5)	5(4)	2.7(1.2)
	TPSS	3.7(1.3)	4(6)	1.5(1.5)	6(4)	2.5(1.2)
	SCAN	3.9(1.7)	5(7)	1.4(1.1)	5(4)	1.8(9)
	revSCAN	3.5(1.5)	6(9)	1.3(9)	5(4)	1.5(7)
	vdW-DF	3.7(1.2)	4(5)	1.6(1.7)	6(4)	2.6(1.6)
	vdW-DF2	4.5(2)	4(5)	2(2)	6(4)	2.7(1.7)
	BEEF-vdW	3.6(1.1)	4(6)	1.5(1.5)	6(4)	2.4(1.5)
8 Å cc	B3LYP	3.8(1.9)	3(3)	1.5(9)	6(4)	1.8(1.0)
None	PBE	8(8)	5(6)	4(4)	6(3)	4.2(1.3)
	SCAN	7(9)	4(4)	4(6)	6(3)	3.0(1.4)
	B3LYP	8(10)	4(4)	3(4)	6(3)	3.2(1.5)

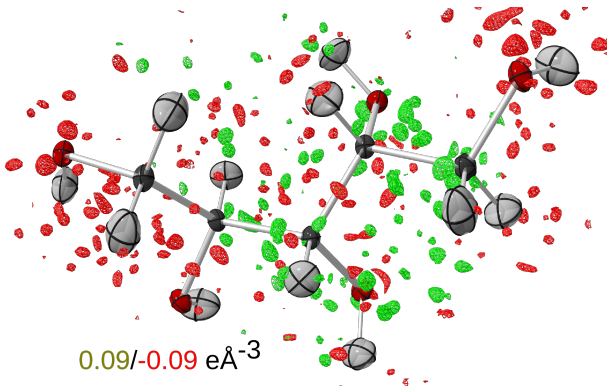


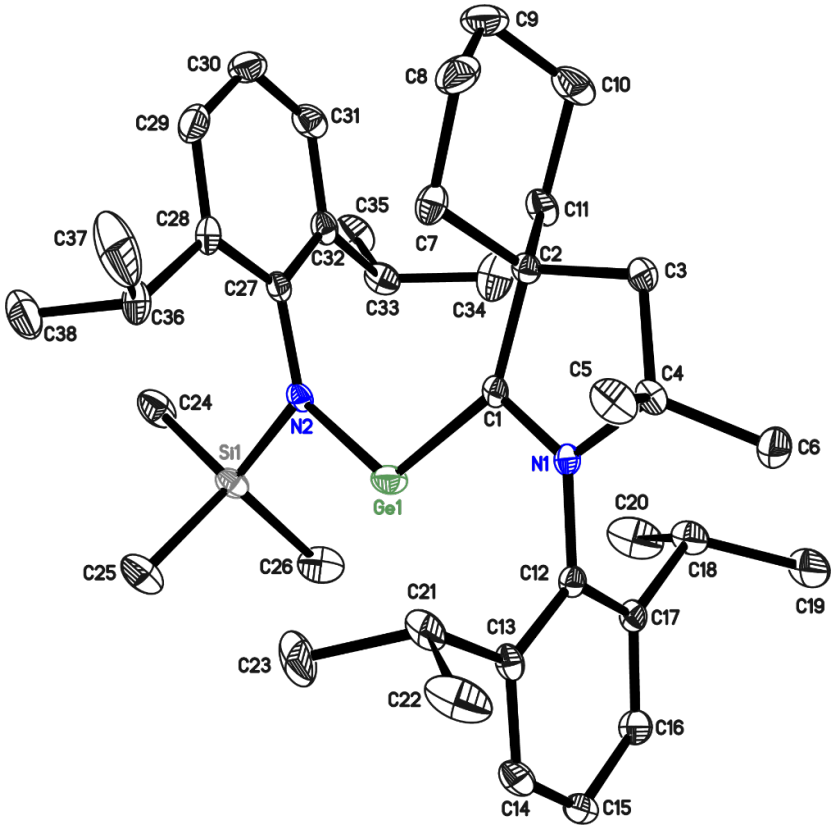
Figure B.14.: Difference electron density resulting from the PAW-HAR refinement with XHARPy/GPAW under inclusion of extinction for **Xy**. ADPs are depicted at 50 % probability level.

C. X-ray structure determinations in cooperation with synthetic groups

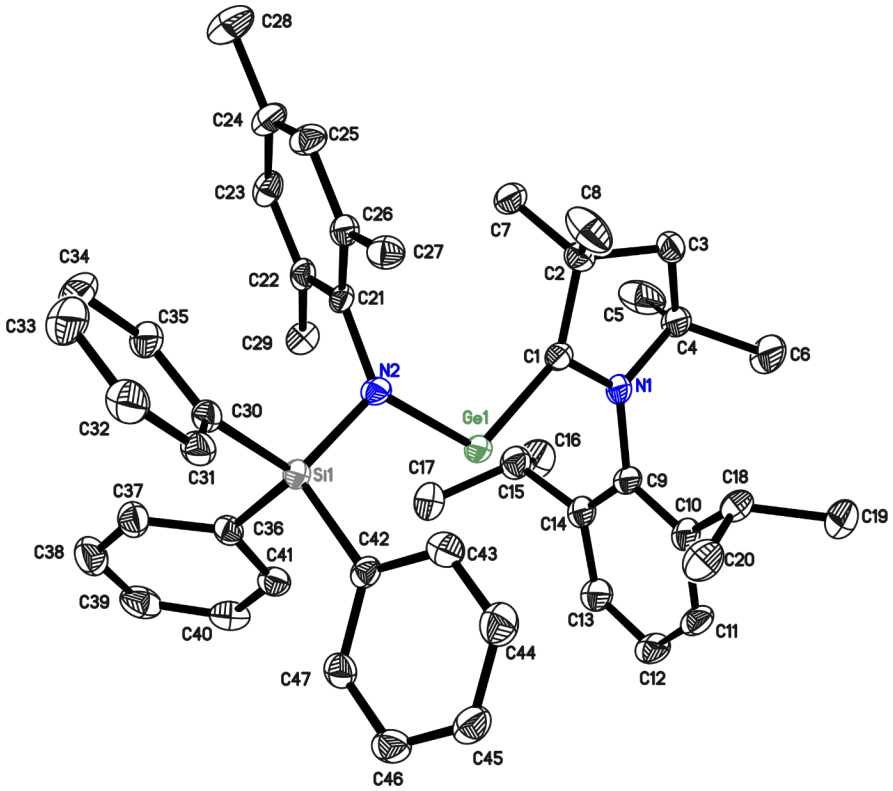
Part of a crystallographer's research is the structure determination of compounds from synthetic groups *via* X-ray diffraction and I was fortunate enough to also take part in such investigations. The obtained structural information will be listed here for reference. For any more details, please consult the respective publications, which are:

- CCDC 1880920 / 1880921: M. M. Siddiqui, S. K. Sarkar, S. Sinhababu, P. N. Ruth, R. Herbst-Irmer, D. Stalke, M. Ghosh, M. Fu, L. Zhao, D. Casanova, G. Frenking, B. Schwederski, W. Kaim, H. W. Roesky, *J. Am. Chem. Soc.* **2019**, *141*, 1908.
- CCDC 1935771: C. P. Sindlinger, P. N. Ruth, *Angew. Chem.* **2019**, *58*, 15051.
- CCDC 2049785 – 2049791: Y. Ding, P. N. Ruth, R. Herbst-Irmer, D. Stalke, Z. Yang, H. W. Roesky, *Dalton Trans.* **2021**, *50*, 2067.
- CCDC 2067321: Y. Ding, S. K. Sarkar, M. Nazish, S. Muhammed, D. Lüert, P. N. Ruth, C. M. Legendre, R. Herbst-Irmer, P. Parameswaran, D. Stalke, Z. Yang, H. W. Roesky, *Angew. Chem.* **2021**, *60*, 27206.
- CCDC 2086522 / 2086522: Ding, S. K. Sarkar, M. Nazish, P. N. Ruth, R. Herbst-Irmer, S. Muhammed, P. Parameswaran, D. Stalke, H. W. Roesky, *Inorg. Chem.* **2022**, *48*, 19067.

C.1. Cooperation with Mujahuddin M. Siddiqui (Roesky group)

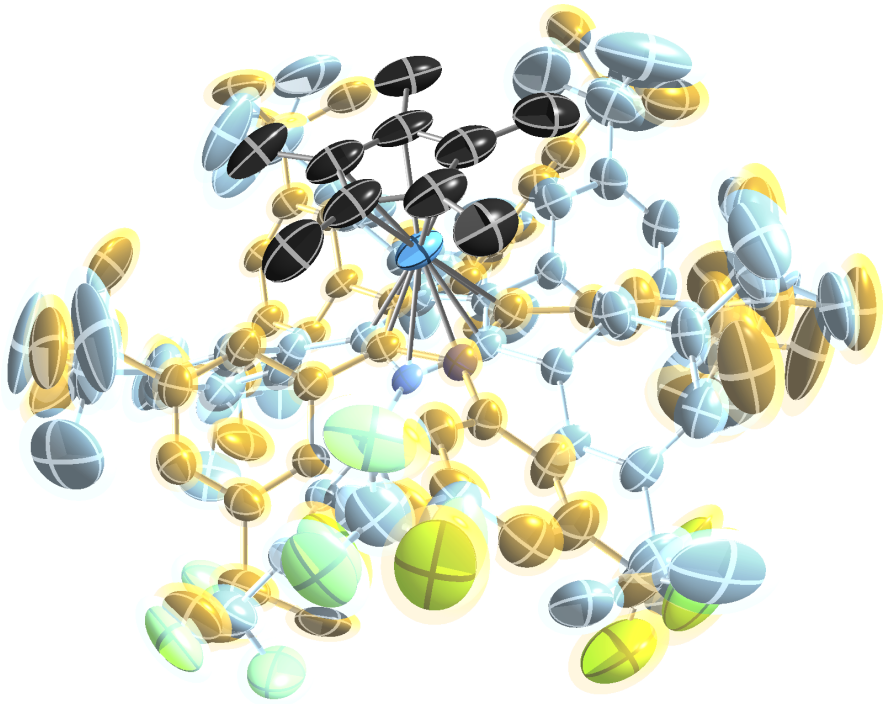


CCDC no.	1880920	formula	$C_{38}H_{61}GeN_2Si$
crystal system	triclinic	space group	$P\,\bar{1}$
a / Å	9.570(2)	α / °	104.34(2)
b / Å	11.955(2)	β / °	93.68(2)
c / Å	17.699(3)	γ / °	108.63(3)
V / Å^{-1}	1836.3(7)	Z	2
M /g mol^{-1}	646.56	ρ_{mass} / Mg m^{-1}	1.169
d_{cryst} / mm	0.49, 0.34, 0.27	μ_{abs} / mm^{-1}	0.476
λ / Å	0.56086	d_{max}, d_{min} / Å	8.96, 0.79
$N_{reflections}$ (all)	62594	$N_{reflections}$ (unique)	7706
$N_{parameters}$	392	$N_{restraints}$	0
R_{merge} / %	7.12	$\Delta\rho_{max}$,$\Delta\rho_{min}$	1.133 / –1.003
$R(F)$ $[I > 2\sigma(I)]$ / %	4.36	$wR_2(F^2)$ / % (all)	11.92



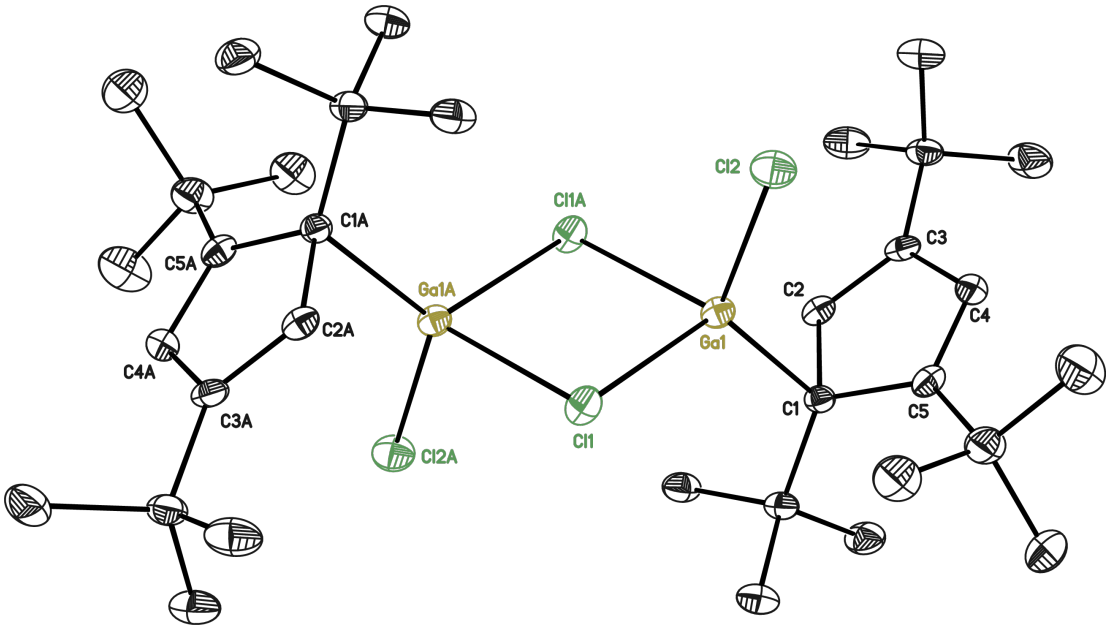
CCDC no.	1880921	formula	$C_{47}H_{57}GeN_2Si$
crystal system	triclinic	space group	$P\,\overline{1}$
a / Å	13.375(2)	α / °	112.83(2)
b / Å	13.416(2)	β / °	100.59(2)
c / Å	14.015(3)	γ / °	108.21(3)
V / Å^{-1}	2063.8(8)	Z	2
M /g mol^{-1}	750.62	ρ_{mass} / Mg m^{-1}	1.208
d_{cryst} / mm	0.51, 0.22, 0.14	μ_{abs} / mm^{-1}	0.43
λ / Å	0.56086	d_{max}, d_{min} / Å	10.74, 0.80
$N_{reflections}$ (all)	72471	$N_{reflections}$ (unique)	8474
$N_{parameters}$	471	$N_{restraints}$	0
R_{merge} / %	5.07	$\Delta\rho_{max}$,$\Delta\rho_{min}$	1.166 / –1.047
$R(F)\,[I > 2\sigma(I)]$ / %	4.59	$wR_2(F^2)$ / % (all)	12.51

C.2. Cooperation with Christian P. Sindlinger

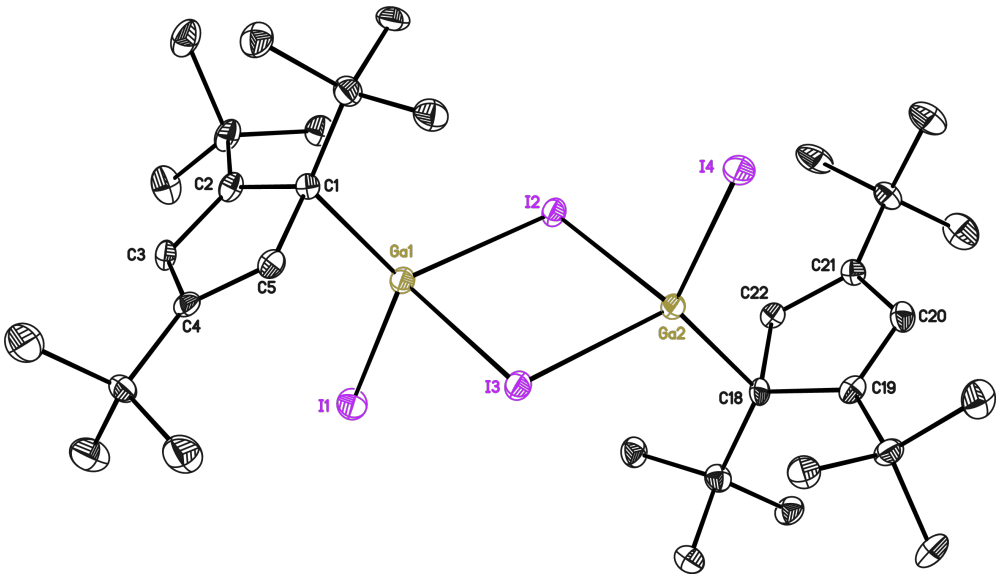


CCDC no.	1935771	formula	$C_{78}H_{102}AlBF_6$
crystal system	monoclinic	space group	$P\,2_1/n$
a / Å	14.770(4)		
b / Å	17.075(5)	β / °	97.27(2)
c / Å	34.408(10)		
V / Å^{-1}	8608(4)	Z	4
M /g mol^{-1}	1191.38	ρ_{mass} / Mg m^{-1}	0.919
d_{cryst} / mm	0.41, 0.31, 0.10	μ_{abs} / mm^{-1}	0.07
λ / Å	0.71073	d_{max}, d_{min} / Å	17.07, 1.14
$N_{reflections}$ (all)	75486	$N_{reflections}$ (unique)	5992
$N_{parameters}$	1276	$N_{restraints}$	8425
R_{merge} / %	6.92	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.399 / –0.264
$R(F)\,[I > 2\sigma(I)]$ / %	8.30	$wR_2(F^2)$ / % (all)	21.93

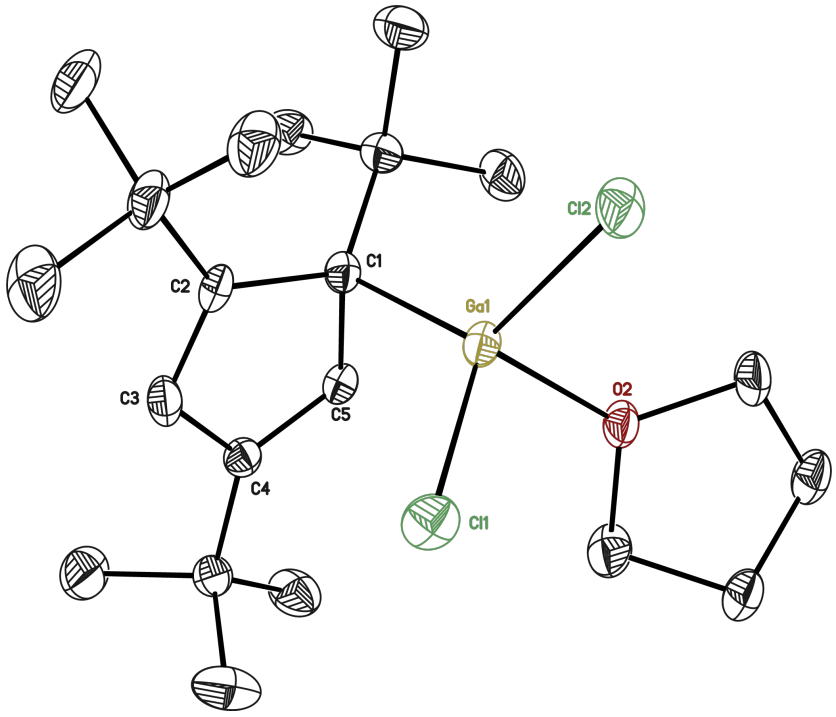
C.3. Cooperation with Yi Ding (Roesky group)



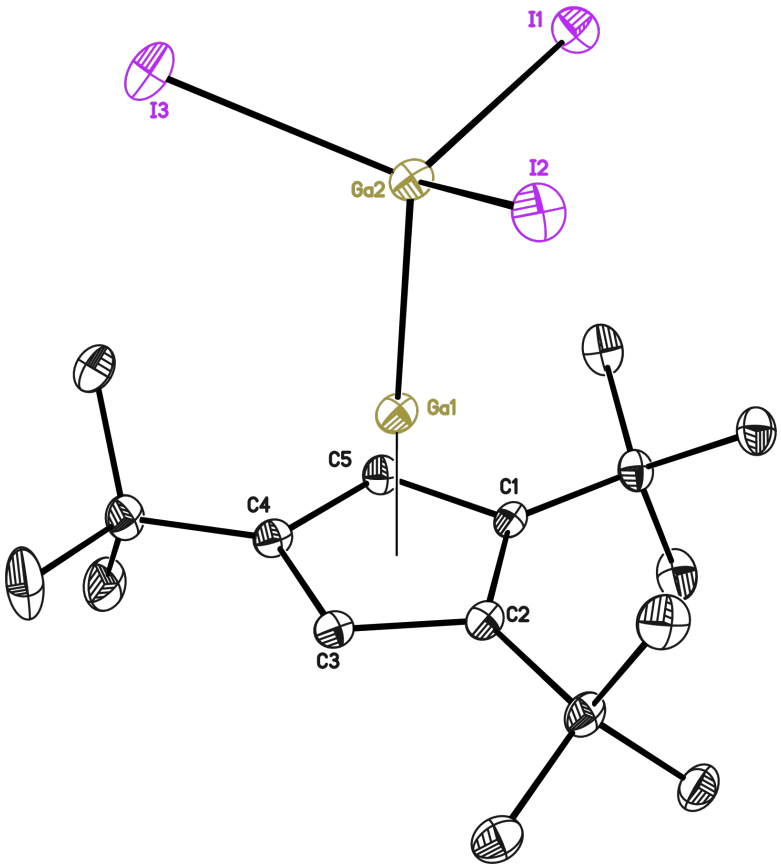
CCDC no.	2049785	formula	$C_{34}H_{58}Cl_4Ga_2$
crystal system	triclinic	space group	$P\,\bar{1}$
a / Å	11.601(2)	α / °	91.96(2)
b / Å	12.454(2)	β / °	101.72(3)
c / Å	13.378(3)	γ / °	90.14(2)
V / Å^{-1}	1891.3(6)	Z	2
M /g mol^{-1}	748.04	ρ_{mass} / g m^{-1}	1.314
d_{cryst} / mm	0.25, 0.12, 0.10	μ_{abs} / mm^{-1}	1.728
λ / Å	0.71073	d_{max}, d_{min} / Å	13.09, 0.77
$N_{reflections}$ (all)	74912	$N_{reflections}$ (unique)	8777
$N_{parameters}$	380	$N_{restraints}$	0
R_{merge} / %	5.32	$\Delta\rho_{max}$,$\Delta\rho_{min}$	1.19 / –0.983
$R(F)$ $[I > 2\sigma(I)]$ / %	3.85	$wR_2(F^2)$ / % (all)	10.36



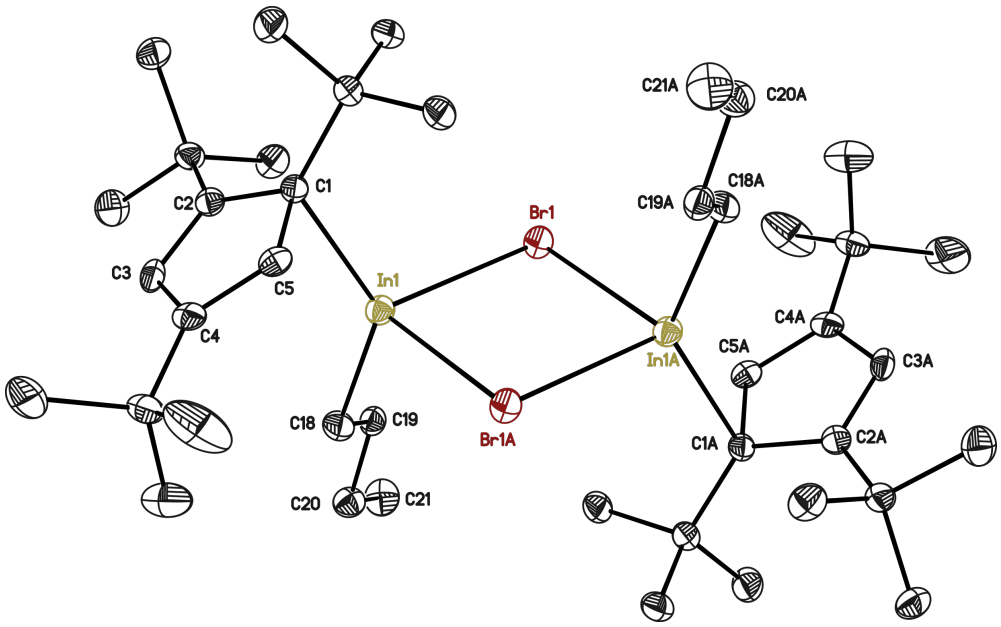
CCDC no.	2049786	formula	$C_{34}H_{58}Ga_2I_4$
crystal system	monoclinic	space group	$P\,2_1/n$
a / Å	9.984(2)		
b / Å	23.624(3)	β / °	99.89(2)
c / Å	17.260(2)		
V / Å^{-1}	4010.5(11)	Z	4
M /g mol^{-1}	1113.84	ρ_{mass} / Mg m^{-1}	1.845
d_{cryst} / mm	0.20, 0.14, 0.10	μ_{abs} / mm^{-1}	4.443
λ / Å	0.71073	d_{max}, d_{min} / Å	11.81, 0.77
$N_{reflections}$ (all)	91909	$N_{reflections}$ (unique)	9248
$N_{parameters}$	379	$N_{restraints}$	0
R_{merge} / %	6.96	$\Delta\rho_{max}$,$\Delta\rho_{min}$	1.686 / –0.602
$R(F)\,[I > 2\sigma(I)]$ / %	2.91	$wR_2(F^2)$ / % (all)	6.16



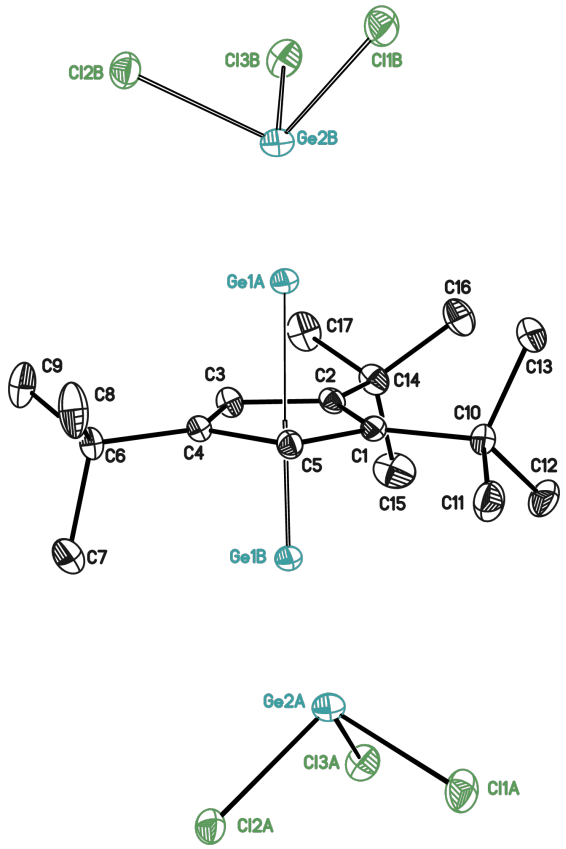
CCDC no.	2049787	formula	$C_{21}H_{37}Cl_2GaO$
crystal system	orthorhombic	space group	$P\,2_12_12$
a / Å	15.744(2)		
b / Å	24.553(4)		
c / Å	11.926(2)		
V / Å^{-1}	4610.1(12)	Z	8
M /g mol^{-1}	446.12	ρ_{mass} / Mg m^{-1}	1.286
d_{cryst} / mm	0.13, 0.12, 0.11	μ_{abs} / mm^{-1}	1.432
λ / Å	0.71073	d_{max}, d_{min} / Å	11.92, 0.77
$N_{reflections}$ (all)	96451	$N_{reflections}$ (unique)	10598
$N_{parameters}$	470	$N_{restraints}$	0
R_{merge} / %	4.93	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.603 / –0.437
$R(F)\,[I > 2\sigma(I)]$ / %	2.93	$wR_2(F^2)$ / % (all)	7.24



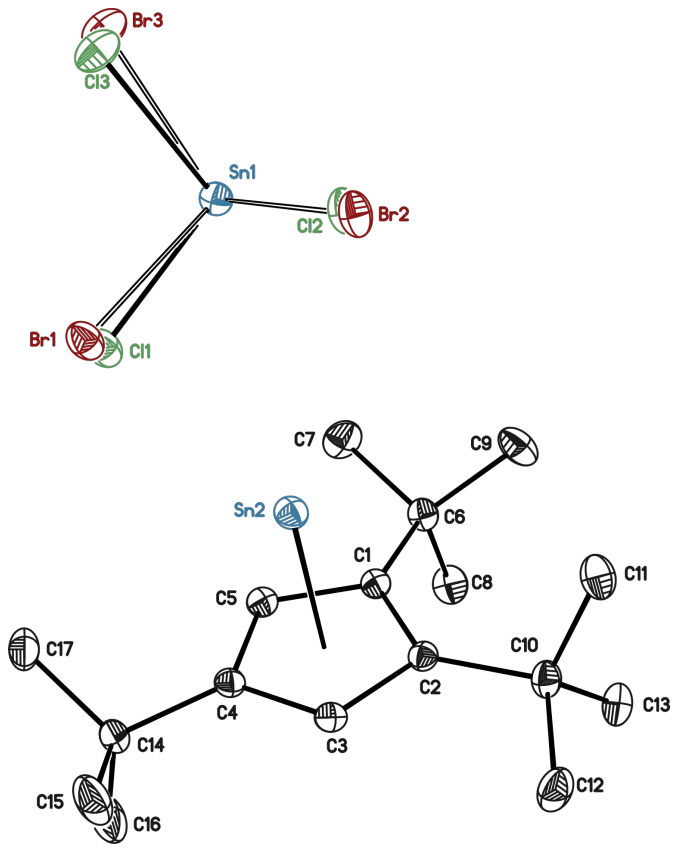
CCDC no.	2049788	formula	$C_{17}H_{29}Ga_2I_3$
crystal system	triclinic	space group	$P\,\bar{1}$
a / Å	10.845(2)	α / °	82.69(2)
b / Å	11.250(3)	β / °	61.37(2)
c / Å	11.306(3)	γ / °	74.03(3)
V / Å^{-1}	1164.0(6)	Z	2
M /g mol^{-1}	753.54	ρ_{mass} / Mg m^{-1}	2.15
d_{cryst} / mm	0.24, 0.20, 0.10	μ_{abs} / mm^{-1}	6.289
λ / Å	0.71073	d_{max}, d_{min} / Å	10.81, 0.76
$N_{reflections}$ (all)	20363	$N_{reflections}$ (unique)	5543
$N_{parameters}$	208	$N_{restraints}$	0
R_{merge} / %	3.85	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.995 / –0.961
$R(F)$ $[I > 2\sigma(I)]$ / %	2.38	$wR_2(F^2)$ / % (all)	5.81



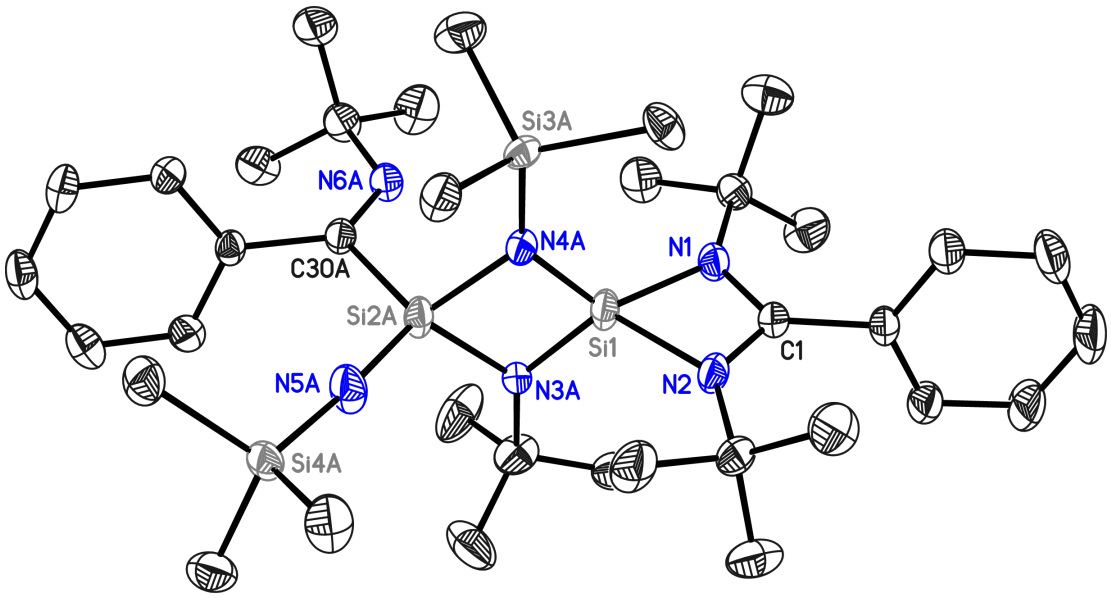
CCDC no.	2049789	formula	$C_{42}H_{76}Br_2In_2$
crystal system	triclinic	space group	$P\,\overline{1}$
a / Å	9.494(2)	α / °	71.33(2)
b / Å	10.197(2)	β / °	73.77(2)
c / Å	12.616(3)	γ / °	77.36(3)
V / Å^{-1}	1099.7(5)	Z	1
M /g mol^{-1}	970.48	ρ_{mass} / Mg m^{-1}	1.465
d_{cryst} / mm	0.15, 0.15, 0.05	μ_{abs} / mm^{-1}	2.891
λ / Å	0.71073	d_{max}, d_{min} / Å	11.64, 0.77
$N_{reflections}$ (all)	22845	$N_{reflections}$ (unique)	5060
$N_{parameters}$	218	$N_{restraints}$	0
R_{merge} / %	7.33	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.712 / –0.525
$R(F)$ $[I > 2\sigma(I)]$ / %	3.50	$wR_2(F^2)$ / % (all)	7.48



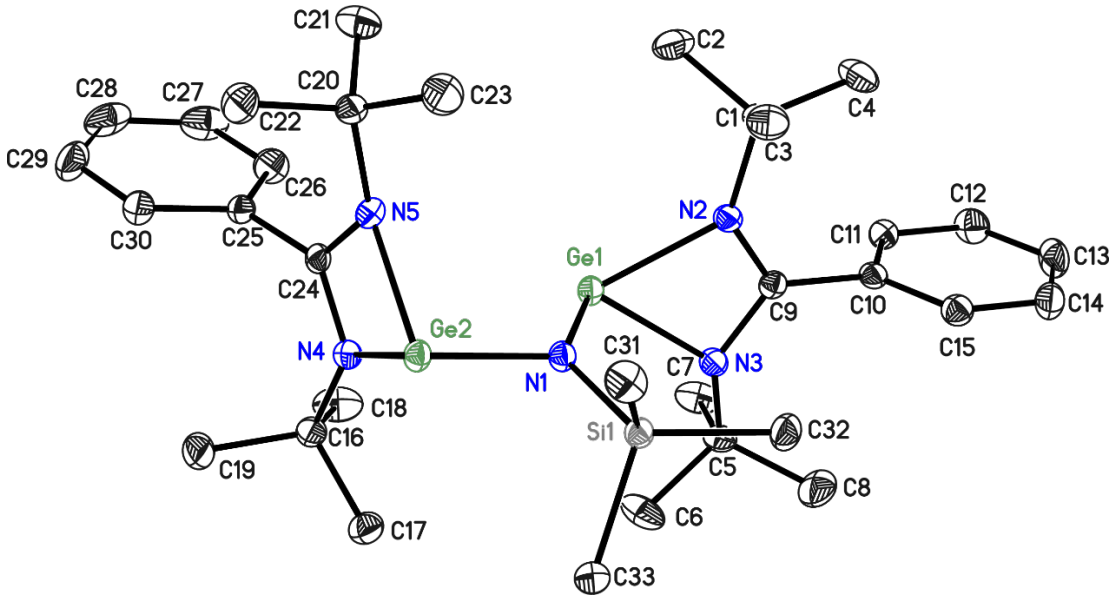
CCDC no.	2049790	formula	$C_{17}H_{29}Cl_3Ge_2$
crystal system	orthorhombic	space group	*P bca*
a / Å	17.206(2)		
b / Å	9.715(2)		
c / Å	25.399(3)		
V / Å^{-1}	4245.6(11)	Z	8
M /g mol^{-1}	484.93	ρ_{mass} / Mg m^{-1}	1.517
d_{cryst} / mm	0.27, 0.22, 0.17	μ_{abs} / mm^{-1}	3.205
λ / Å	0.71073	d_{max}, d_{min} / Å	12.70, 0.70
$N_{reflections}$ (all)	86549	$N_{reflections}$ (unique)	6504
$N_{parameters}$	224	$N_{restraints}$	89
R_{merge} / %	2.94	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.481 / –0.523
$R(F)$ $[I > 2\sigma(I)]$ / %	2.71	$wR_2(F^2)$ / % (all)	6.22



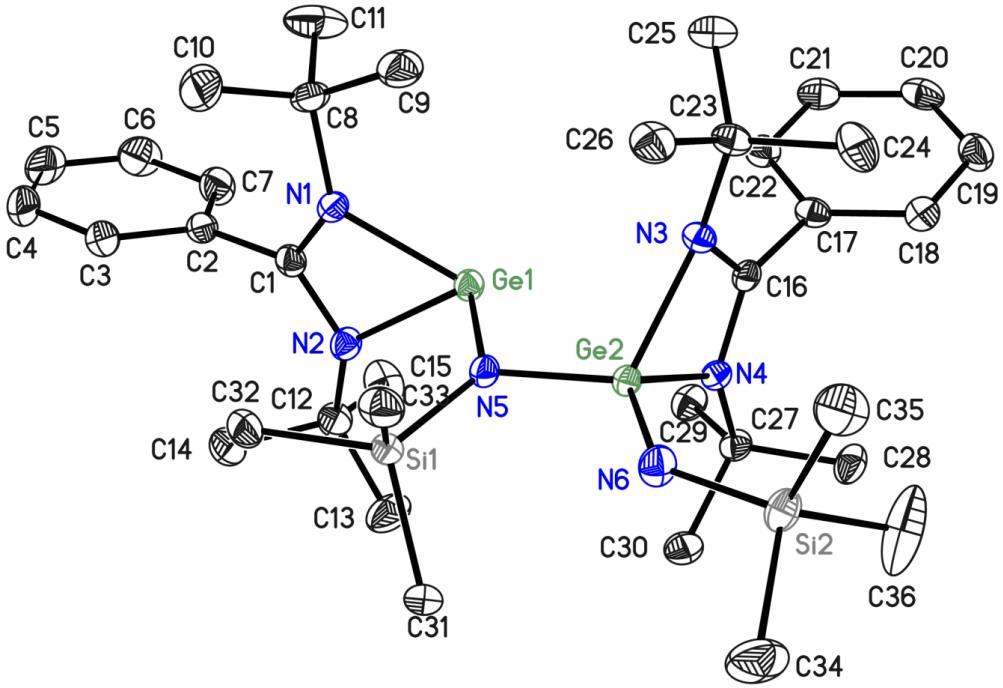
CCDC no.	2049791	formula	$C_{17}H_{29}Br_{0.3}Cl_{2.7}Sn_2$
crystal system	monoclinic	space group	$C\,2/c$
a / Å	27.222(3)		
b / Å	8.752(2)	β / °	106.84(2)
c / Å	19.398(2)		
V / Å^{-1}	4423.3(13)	Z	8
M /g mol^{-1}	590.89	ρ_{mass} / Mg m^{-1}	1.775
d_{cryst} / mm	0.35, 0.25, 0.14	μ_{abs} / mm^{-1}	3.138
λ / Å	0.71073	d_{max}, d_{min} / Å	13.03, 0.75
$N_{reflections}$ (all)	33569	$N_{reflections}$ (unique)	5495
$N_{parameters}$	220	$N_{restraints}$	6
R_{merge} / %	2.05	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.426 / –0.214
$R(F)\,[I > 2\sigma(I)]$ / %	1.40	$wR_2(F^2)$ / % (all)	3.29



CCDC no.	2067321	formula	$C_{36}H_{64}N_6Si_4$
crystal system	monoclinic	space group	*C* c
a / Å	11.985(2)		
b / Å	22.595(3)	*β* / °	94.22(2)
c / Å	15.218(2)		
V / Å^{-1}	4109.9(10)	*Z*	4
M /g mol^{-1}	693.29	ρ_{mass} / Mg m^{-1}	1.12
d_{cryst} / mm	0.42, 0.40, 0.22	μ_{abs} / mm^{-1}	0.176
λ / Å	0.71073	d_{max}, d_{min} / Å	11.29, 0.73
$N_{reflections}$ (all)	85318	$N_{reflections}$ (unique)	11096
$N_{parameters}$	652	$N_{restraints}$	1644
R_{merge} / %	3.45	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.37 / –0.196
$R(F)$ $[I > 2\sigma(I)]$ / %	3.40	$wR_2(F^2)$ / % (all)	9.28



CCDC no.	2086521	formula	$C_{33}H_{55}Ge_2N_5Si$
crystal system	triclinic	space group	$P\,\bar{1}$
a / Å	14.858(2)	α / °	87.39(2)
b / Å	14.933(2)	β / °	88.93(3)
c / Å	16.872(3)	γ / °	78.48(2)
V / Å^{-1}	3664.1(10)	Z	4
M /g mol^{-1}	695.09	ρ_{mass} / Mg m^{-1}	1.26
d_{cryst} / mm	0.23, 0.19, 0.15	μ_{abs} / mm^{-1}	1.701
λ / Å	0.71073	d_{max}, d_{min} / Å	16.86, 0.77
$N_{reflections}$ (all)	118941	$N_{reflections}$ (unique)	16776
$N_{parameters}$	769	$N_{restraints}$	0
R_{merge} / %	3.26	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.346 / –0.293
$R(F)\,[I > 2\sigma(I)]$ / %	2.22	$wR_2(F^2)$ / % (all)	5.61



CCDC no.	2086522	formula	$C_{39.5}H_{68}Ge_2N_6Si_2$
crystal system	triclinic	space group	$P\,\bar{1}$
a / Å	12.576(2)	α / °	112.60(3)
b / Å	13.182(2)	β / °	93.17(2)
c / Å	14.738(3)	γ / °	96.32(2)
V / Å^{-1}	2229.2(8)	Z	2
M /g mol^{-1}	828.35	ρ_{mass} / Mg m^{-1}	1.234
d_{cryst} / mm	0.27, 0.19, 0.12	μ_{abs} / mm^{-1}	1.435
λ / Å	0.71073	d_{max}, d_{min} / Å	12.04, 0.78
$N_{reflections}$ (all)	53118	$N_{reflections}$ (unique)	9852
$N_{parameters}$	560	$N_{restraints}$	555
R_{merge} / %	3.35	$\Delta\rho_{max}$,$\Delta\rho_{min}$	0.82 / –0.475
$R(F)$ $[I > 2\sigma(I)]$ / %	3.05	$wR_2(F^2)$ / % (all)	7.83

D. Acknowledgements

No project can be completed without help from many other people and so it is customary but also always the right thing to do to thank these people who were essential for the overall success of this PhD period.

First, I want to thank my supervisor Prof. Dietmar Stalke for giving me the opportunity to work on interesting projects for encouragement and focus, but also for the freedom to pursue approaches of my own choosing. I feel that the degree to which this was possible should not be taken for granted. At the same time, he also reminded me when collecting and writing up the results was in order.

I also want to thank the other two members of my thesis advisory committee: Prof. Martin Suhm and Dr Carlo Gatti. I always had the feeling you were taking the word advice from the naming of this group seriously and I benefited from that substantially. Additionally, I want to thank the other four members of my examination board for agreeing to be a part of my thesis defence.

Additionally, I want to thank all the people participating on the MetalJet project. First and foremost Tobias and Nico, who worked with me on and thought with me about the machine and how to bring it to where we want it to be. But I also want to thank Holger Ott, Julius Hållstedt, Martin Adam, Jürgen Graf, Michael Peeters and Irene Cimatti for lots of discussions and help on this journey.

I want to thank the synthetic chemists who cooperated with me on projects. The first to name is Timo Schillmöller. Thank you for sharing your thoughts on luminescence and listening to mine. Second in this text but equally important is Christian Sindlinger. Thanks for all the discussions, advice on and beyond chemistry and the two cooperation projects which were part of this thesis. Finally, I also want to thank my two cooperation partners from the group of Prof. Herbert Roesky, Mujahuddin M. Siddiqui and Yi Ding, for very productive cooperation.

I want to thank Carlo Gatti and Davide Ceresoli for their hospitality during my research stay in Milan, both while at work and during leisure time. Even though the luminescence calculations might have been less yielding in concrete results, I feel that the basics and approaches I learned in Italy were the foundation of the successes achieved afterwards. Here, I also want to thank them for giving feedback on my ideas during and after the research stay. Some of these ideas turned into Chapter 6, while others turned out to be rubbish. From the same research stay, I also want to thank Bjørn Egede Grønbech

for being a great flatmate for that time. I appreciated the company, as well as the conversations about crystallography and all the other topics we discussed.

This trip was made possible because I was part of the BENCh RTG. Additionally, I want to thank all the members for sharing insight into their research and approaches and for social activities, even if my participation in these stalled after COVID. I also want to thank the planners of that RTG. I felt the overall conception set us up for success and I enjoyed and profited from the goals, insights and opportunities the research training group provided. I hope there is enough BENCh in this thesis in return.

I want to thank Christian Schürmann. Having been a friend and neighbour, he lured me into the topic of crystallography and also became a great colleague. Additionally, I want to thank all the other people that shared my office and passion for crystallography, with nice conversations on and off-topic. To be named here in the cop-out alphabetical order are Anna Lübben, Annika Münch, Christian Köhler, Corinna Willenberg, Felix Engelhardt, Helena Keil, Katharina Rachuy and Lennard Krause.

Also thanks to the remaining Stalke group and the adjacent Krawczuk group as well. I always felt welcome, enjoyed all the group activities be it in university, online during the pandemic or in town and hope to see a lot of you again. I started writing out names, but every line that could be drawn who to mention and where to stop felt too arbitrary to me.

I want to thank my two flatmates, for saving my sanity during otherwise social distancing, nice kitchen discussions and sharing a home with me.

Also keeping my sanity during COVID but especially during non-COVID times were the people from the Postorchester Göttingen. Thank you for being such a large and persistent part of my life in Göttingen. I will miss meeting so many nice people and making music together.

And last but the opposite of least, I want to thank my family. First and foremost my mother who always supported me during my life. But I also want to thank my brother, step-father, step-brothers and everyone who always has been part of or has joined this family.