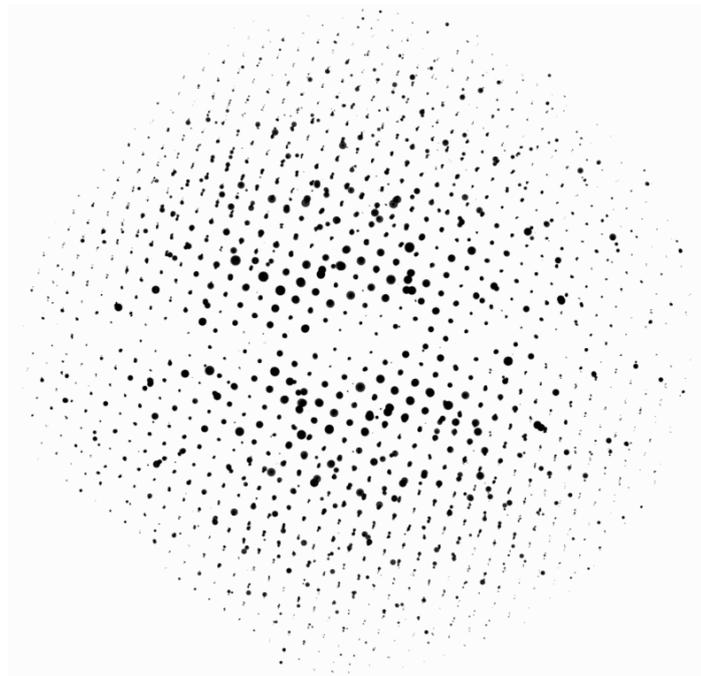Aurora Teien

Optimizing the data collection and data processing procedure for 3D electron diffraction crystal structure analysis

Master's thesis in Nanotechnology
Supervisor: Antonius T. J. van Helvoort
Co-supervisor: Tina Bergh, Oskar Ryggetangen, Emil F. Christiansen
June 2024

Master's thesis

◻ NTNU
Norwegian University of
Science and Technology

Aurora Teien

Optimizing the data collection and data processing procedure for 3D electron diffraction crystal structure analysis

Master's thesis in Nanotechnology
Supervisor: Antonius T. J. van Helvoort
Co-supervisor: Tina Bergh, Oskar Ryggetangen, Emil F. Christiansen
June 2024

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

Abstract

Determining the structure of crystalline materials is essential to understand and optimize their properties. Over the recent years, the efficiency and quality of structure analysis of small inorganic materials and organic crystalline molecules have been revolutionized by the introduction of rotational 3D electron diffraction (3D ED) methods. Rotational 3D ED techniques are based on the concept of orienting the crystal along an arbitrary low-symmetry direction, rotating it about a fixed tilt axis, and collecting a series of 2D diffraction patterns at regular intervals.

This work documents the first steps toward establishing 3D ED as a technique at the Norwegian University of Science and Technology (NTNU), where data is collected on an existing transmission electron microscope (TEM) with a direct electron detector (DED). The data are obtained through the continuous rotation electron diffraction (cRED) method, where the crystal is continuously tilted, with the dedicated software Instamatic controlling the TEM and the 2D DED. Different variants of the zeolite mordenite (MOR) are analyzed. Proven protocols for data acquisition and processing for 3D ED have been developed by several research groups abroad, however, there is currently no standard processing protocol. The data processing workflow investigated in this work is based on the programs REDp, XDS, XPREP, EDtools, SHELXT, SHELXL, and Olex2, of which most are well-established in the X-ray structural analysis community. Structure analysis consists of crystal structure determination, involving reciprocal reconstruction from the 2D pattern series, peak indexing, structure solution, and further structure refinement. The final result is a fully reconstructed 3D model of the studied crystal.

The present work has two main objectives. The first aim is to establish a data processing procedure for 3D ED crystal structure analysis. The workflow was successfully tested by application to published MOR reference datasets. A user manual covering the extensive data processing procedure was then developed based on the results and experiences from testing the reference data. The second aim is to optimize the data acquisition routine for the experimental setup at NTNU. During processing of this experimental data, errors were identified related to the observed intensities, such as saturation, in addition to timing errors in capturing defocused tracking frames with Instamatic, and an inhomogeneous tilt step. Nevertheless, the analyses resulted in useful input regarding data collection optimization. Concrete suggestions are made for future work based on the results of the present study.

To summarize, this work is an important step towards implementing 3D ED data acquisition and processing for structure analysis at NTNU. The findings show that even though a successful data processing procedure is established, further improvements to the data collection procedure still remain to reach the end goal of a robust 3D ED method.

Sammendrag

Å bestemme strukturen til krystallinske materialer er avgjørende for å forstå og optimalisere egenskapene deres. I løpet av de siste årene har effektiviteten og kvaliteten på strukturanalyse av små uorganiske materialer og organiske krystallinske molekyler blitt revolusjonert og betydelig forbedret ved introduksjonen av rotasjon 3D elektron diffraksjons (3D ED) metoder. 3D ED-teknikker er basert på konseptet om å orientere krystallen langs en vilkårlig retning med lav symmetri, rotere den rundt en fast vippeakse og samle inn en serie 2D-diffraksjonsmønstre med regelmessige intervaller.

Dette arbeidet dokumenterer de første stegene mot å etablere 3D ED som en teknikk ved Norges Teknisk-Naturvitenskapelige Universitet (NTNU), der data samles inn på et eksisterende transmisjonselektronmikroskop (TEM) med en direkte elektron detektor (DED). Dataene innhentes ved hjelp av kontinuerlig rotasjonselektrondiffraksjon (cRED)-metoden, der krystallen roteres kontinuerlig, med den dedikerte programvaren Instamatic som kontrollerer TEMen og 2D DEDen. Ulike varianter av zeolitten mordenite (MOR) vil analyseres ved hjelp av 3D ED. Flere utenlandske forskergrupper har utviklet velprøvde protokoller for datainnsamling og -behandling for 3D ED, men det finnes foreløpig ingen standard behandlingsprotokoll. Databehandlingsarbeidsflyten som er undersøkt i dette arbeidet, er basert på programmene REDp, XDS, XPREP/EDtools, SHELXT, SHELXL og Olex2, hvorav de fleste av disse er veletablerte i røntgenstrukturanalysemiljøet. Strukturanalysen består av krystallstrukturbestemmelse, som omfatter resiprok rekonstruksjon fra 2D-mønsterserien, indeksering av diffraksjonspunkter, strukturløsning og videre strukturforbedring. Det endelige resultatet er en fullstendig rekonstruert 3D-modell av den studerte krystallen.

Det foreliggende arbeidet har to hovedmål. Det første målet er å etablere en databehandlingsprosedyre for 3D ED-krystallstrukturanalyse. Arbeidsflyten ble testet ved å bruke den på publiserte MOR-referansedatasett. Deretter ble det utviklet en brukerhåndbok som dekker den omfattende databehandlingsprosedyren, basert på resultatene og erfaringene fra testingen av referansedataene. Det andre målet er å optimalisere datainnsamlingsrutinen for det eksperimentelle oppsettet ved NTNU. Under prosesseringen av disse dataene ble det oppdaget at de eksperimentelle intensitetene var kompromittert, muligens på grunn av metning. Andre feil som ble identifisert under rekonstruksjonen, var en tidsfeil ved innsamling av defokuserte sporingsbilder med Instamatic og et inhomogent vippetrinn. Analysene resulterte likevel i nyttige innspill til optimalisering av datainnsamlingen. Basert på resultatene fra denne studien kommer vi med konkrete forslag til fremtidig arbeid.

For å oppsummere er dette arbeidet et viktig skritt mot å implementere 3D ED datainnsamling og -behandling for strukturanalyse ved NTNU. Funnene viser at selv om en vellykket databehandlingsprosedyre er etablert, gjenstår det fortsatt ytterligere forbedringer av datainnsamlingsprosedyren for å nå det endelige målet om en robust 3D ED-metode.

Preface

This project was carried out at the Department of Physics at NTNU. The author has been working with this project related to 3D electron diffraction from August 2023 to June 2024. During autumn 2023 the work was part of a specialization project TFY4520, includeing a written report that counts for 15 study credits [1]. The current project, a TFY4905 Nanotechnology Master's Thesis of 30 credits, was run during Spring 2024 with official starting date on the 11th of January and date of delivery on the 6th of June. Major parts of the theory in this work are adapted from my project thesis written in autumn 2023. The objectives of the project were to study and optimize a specific 3D electron diffraction data processing routine and data collection parameters to establish the technique for future use at the TEM Gemini Center at NTNU. Datasets used in this project were collected by Oskar, Tina, and Emil, except for reference datasets from Stockholm University that are available from Zenodo [2]. This Master's thesis was written by myself. AI tools have been used for translation (DeepL Translate), coding assistance (ChatGPT), and briefly for language enhancement (Writefull), but all AI suggestions have been thoroughly inspected, edited, and read through manually by myself.

Aurora Teien
Trondheim, June 6th 2024

Abbreviations

3D ED 3D electron diffraction

ADP atomic displacement parameter

ADT automated diffraction tomography

BFP back focal plane

CA condenser aperture

CBED convergent beam electron diffraction

CCD charge coupled device

CCTBX Computational Crystallography Toolbox

CMOS complementary metal oxide semiconductors

COF covalent organic frameworks

cRED continuous rotation electron diffraction

CTEM conventional transmission electron microscopy

DED direct electron detector

ED electron diffraction

EDT electron diffraction tomography

FCC face centered cubic

FEG field emission gun

HRTEM high resolution transmission electron microscopy

ITA International Table of Crystallography A

MicroED microcrystal electron diffraction

MOF metal-organic frameworks

MOR mordenite

ND neutron diffraction

PED precession electron diffraction

PSF   point-spread functions

PXRD powder X-ray diffraction

RED   rotation electron diffraction

REDp Rotation Electron Diffraction processing (software)

SA   selected area

SAED selected area electron diffraction

SCED single crystal electron diffraction

SCXRD single crystal X-ray diffraction

SEM   scanning electron microscope

STEM scanning transmission electron microscopy

TEM   transmission electron microscope[*]

XDS   X-ray Diffraction Software (software)

XRD   X-ray diffraction

---

[*]TEM as an abbreviation can be used about both the technique (transmission electron microscopy) and the instrument (transmission electron microscope)

Contents

Chapter 1

Introduction

Crystals are materials containing repeating patterns of structural units [3, 4]. These repeating units, which can be atoms, molecules, or molecular groups, are arranged in a specific order that defines the properties of a crystal, referred to as its crystal structure [5]. Thus, determining the crystal structure of a crystalline material reveals its properties [6].

Central for structure determination is the phenomenon of diffraction, where the scattering of the incoming coherent radiation is used to create a reciprocal representation of the two-dimensional atomic planes of a crystal [7]. A diffraction pattern in one direction will provide only 2D information about the three-dimensional crystal. Hence, to know the full crystal structure one would need to encounter all three dimensions, which can be done by tilting the sample and collecting diffraction patterns for each tilt obtaining a stack of 2D diffraction patterns. Rotational single crystal methods exist for X-ray diffraction (XRD), neutron diffraction (ND), and electron diffraction (ED) [4, 8]. This opens up the possibility for a full 3D analysis.

Historically, the main radiation used for diffraction to study crystals has been X-rays. Specifically, single crystal X-ray diffraction (SCXRD) and powder X-ray diffraction (PXRD) are two common methods that have been used for a long time to analyze small crystalline materials. Unfortunately, SCXRD require perfectly ordered crystals of considerable size ($\sim 100 \, \mu m^3$) which might be hard to come by in cases. PXRD, as an alternative for polycrystalline and powder samples, struggle with large unit cells and multiple phases due to peak overlap [9]. Further, some specific structural features as, for example, chirality cannot easily be determined by kinematic diffraction.

Due to the limitations of X-ray diffraction, as mentioned above, recent research analysing complex submicron samples, like zeolites, metal-organic frameworks (MOF), and covalent organic frameworks (COF), have required a shift towards using radiation that can be focused to a small probe and give a strong enough diffraction signal from a much smaller volume ($1.0 \, \mu m^3$ - 10^{-4} μm^3), i.e. electrons [10]. The fact that electrons are charged particles, cause them to interact stronger with matter than X-rays do. Electrons have an advantage because of the lower wavelength (251 pm at 200 kV vs 0.1-10 Å for X-rays) and the relative ease to manipulate a charged particle beam compared to photons. Electron diffraction (ED) in electron microscopes (scanning electron microscope (SEM) and transmission electron microscope (TEM)) and dedicated setups (LEED, RHEED) is well established. However, there are challenges such as the fact that vacuum is required to propagate the beam, in addition to inelastic and dynamical effects.

3D electron diffraction (3D ED), also known as microcrystal electron diffraction (MicroED) [11] or automated diffraction tomography (ADT) [12], has experienced great technological and methodical improvements over recent years [13, 14]. Although the concept was proposed

some decades ago [4], the method really took off when new developments, including improved detector technology such as the entry of the direct electron detector (DED), in addition to automated data collection and processing routines, were introduced [10]. One major breakthrough for 3D ED was the Science paper by Lukas Palatinus [15], this led to the recognition as the Top 5 "Breakthrough of the Year" by Science in 2018 [16].

3D ED experiments can be conducted using TEM, or dedicated 3D ED setups developed by electron microscopy companies [17, 18]. TEM is a widely used scientific instrument for various imaging and analysis techniques. Hence, 3D ED may be considered a relatively cheap and accessible technique, especially compared to other methods based on synchrotron radiation or single-particle EM / CryoEM [19]. A suitable setup for 3D ED, which could be a TEM, needs a stable goniometer, tomography holders that allow high tilt, and a good detector to collect diffraction patterns efficiently with high dynamic range. In a TEM, the illuminated area that contribute to the patterns can be seen as a single crystal, hence will ED of powders in TEM resemble SCXRD. Conventional TEM characterisation typically combines imaging and diffraction techniques, like high resolution transmission electron microscopy (HRTEM) or scanning transmission electron microscopy (STEM), that have lenses and detector combined, which with data processing recreate real space images. ED in conventional TEM techniques is collected along well-defined zone axes, which can cause severe dynamical effects and beam damage* because the signal required is relatively high [20]. 3D ED, on the other hand, is performed on arbitrary off-zone axes, which reduces dynamical effects. Automation of the data collection procedure makes the method efficient; a full tilt series is obtained in seconds to minutes [12, 21].

This thesis addresses two main objectives related to 3D crystal structure analysis based on the latter approach for data collection. The first objective is to establish a robust data processing routine for 3D ED crystal structure analysis. The data processing procedure will be applied to zeolite reference datasets from Stockholm University [2]. This is done to identify the specific purposes and optimal implementation of each constituent program of the overall data processing workflow. For the method to be reproducible, a data processing manual for new users will be developed. In setting up the data processing routine, there is a feedback loop to data acquisition, and hence different acquisition parameters need to be tested. This leads to the second objective, which aims to apply this data processing procedure, to 3D ED data obtained with the experimental setup at NTNU, in order to optimize the data acquisition procedure and parameters. The general goal is to be able to collect adequate 3D ED data for a successful crystal structure analysis with the current data processing routine.

The thesis is written for the TEM Gemini Center at NTNU, which first started establishing 3D ED as an internally funded project in spring 2023, as the first in Norway using modern detector technology (i.e. DED). As an essential part of the establishment of the 3D ED procedure, a data processing routine was necessary to determine, solve, and refine the structure, which will be done in this work. The practical data collection procedure is assisted by the Python-based software *Instamatic* [22], which was modified to tailor the JEOL ARM equipped with Norway's currently most advanced DED at NORTEM Node Trondheim autumn 2023. Both Instamatic [22] and the processing route chosen here, continuous rotation electron diffraction (cRED), were developed in the same research groups [23]. The internally funded project has an active collaboration with these research groups, through Dr. Tom Wilhammar from Stockholm University. Therefore, the data processing routine explored and optimized in this work is the one described by Wilhammar et al. in Nature Protocols

---

*Depends on material and electron dose or dose rate

[23], which take use of the following programs: REDp, XDS, XPREP/EDtools, SHELXT, SHELXL, and Olex2. These programs will in turn reconstruct the reciprocal space in three dimensions, index and integrate reflections, solve, and refine the crystal structure. REDp and XDS were central the first report in this project [1]. This is not the only existing data processing protocol; in fact, there are multiple different protocols, such as PETS [14], DIALS [24], MOSFLM [25], in addition to commercial software for specific instruments such as Apex (from Bruker) and CrysAlis^{Pro} (from Rigaku) [18]. Furthermore, it is common for 3D ED data processing protocols to use well-established XRD structure solution software packages such as SHELX, as will also be done in this work [26, 27]. Although this combination is suitable for many cases, the increased dynamical effects caused by electrons can lead to more inaccurate fits between the observed and calculated models. In the field of electron crystallography, it is debated which data processing approach will crystallize out in the end and bench tests have not been reported as far as the author knows. The current status is based on the choice the developers make and further developments are expected as the 3D ED technique becomes more accessible.

This work will specifically use zeolites as test materials. Zeolites are aluminosilicate microporous materials mostly known for their absorption properties and their ability to exchange ions. These properties make them suitable as catalysts, membranes, sensors, absorbents and for drug delivery [28]. Previous electron tomography studies show that the morphology of the 3D porous network of zeolites is correlated with its catalytic and adsorption properties [29, 30]. Hence, it is essential to determine the crystal structure of the zeolites to know its properties. However, zeolites are beam sensitive materials that can gain severe amorphization damage from conventional TEM experiments and can be hard to grow into large single crystals [31]. 3D ED is considered an adequate technique for zeolite structure analysis and is used in previous published work [32, 9].

In summary, 3D ED provides a fast and simple data collection method for samples that are too small or too complex for XRD, or too beam sensitive for conventional TEM techniques [10]. Different data processing routines for 3D ED exist, although it is disputed which approach will dominate the future of crystal structure analysis. In this work, an optimized data processing and acquisition routine for 3D ED data is presented based on zeolite test materials. A manual for new users will be constructed for the full data processing procedure.

The work is presented as follows. First, a summary of the theoretical knowledge required to understand this thesis is presented. This includes information about basic crystallography, diffraction, structure solution, TEM, zeolites and the specific technique of 3D ED. Then the experimental details will be presented, more specifically tools and techniques used for both data acquisition and data processing, as well as the type of materials studied. Data collection parameters for selected datasets are presented. The experimental details of the all 63 datasets obtained at NTNU, are presented in Appendix A.1. The full structure solution and refinement procedure will be described in the results section, along with a minimum dataset requirement analysis of zeolite reference data. Next, a discussion of hardware limitations and evaluation of the data acquisition procedure and parameters for the data collected with the NTNU setup follow. Finally, the previously described processing procedure is evaluated and potential future work is presented, before the conclusions of the work follow towards the very end. The data processing manual will be presented in Appendix A.3. To clarify, the first aim of the thesis, i.e. evaluation of the data processing procedure, will be presented in the results chapter, while the second aim, optimizing the data collection procedure, will first be presented in the discussion chapter, because applying the data processing routine to the experimental data required in depth parallel discussion and evaluation.

Chapter 2

Theoretical Framework

2.1 Crystallography

Crystals denote solid materials with a repeating periodic arrangement of atoms. The crystal can be described as a basis placed on a lattice. The basis consist of a fundamental set of atoms, while the lattice can be described with translational vectors that span the whole 3D space. The lengths and angles between the translational vectors vary for different crystals, and make up the fundamental repeating unit of the lattice called the unit cell. The basic structure can conveniently be described by symmetry operations, which describe how fragments of the crystal structure are related and can be repeated.

Following, will first the 3D lattices be introduced, before going on to the symmetry in the repeating units, called the point groups, and finally the combination of these two as 3D space groups. The section is based on [6, 3].

2.1.1 Bravais lattices

The crystal lattice consist of all points that can be written as an integer linear combination of the basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{c}_3$,

$$\mathbf{r}_i = x_i\mathbf{a}_1 + y_i\mathbf{a}_2 + z_i\mathbf{a}_3, \tag{2.1}$$

where x_i, y_i and z_i each are arbitrary integers. The length of a vector $\mathbf{a}_i$ is usually denoted by the norm symbols $|\mathbf{a}_i|$, although for the basis vectors, the lengths are commonly named $|\mathbf{a}_1| = a$, $|\mathbf{a}_2| = b$ and $|\mathbf{a}_3| = c$. A generic representation of a unit cell is shown in Figure 2.1(a), where a, b and c denote the lengths of the basis vectors, while α, β and γ denote the angles between those vectors. These are the lattice parameters of a unit cell and makes it possible to determine the differences between basic crystal systems.

To properly denote the planes and directions of a crystal lattice, it is common to use Miller indices. This allows us to keep track of where certain lattice points and planes are located relative to each other. A vector or direction is denoted by square brackets $[hkl]$ visualized in Figure 2.1(b), while the set of all symmetrically equivalent directions are denoted $< hkl >$. A plane on the other hand is denoted with round brackets (hkl) visualized in Figure 2.1(c), while the set of all symmetrically equivalent planes are denoted $\{hkl\}$. The distance between equivalent parallel planes is called the d-spacing, which for a cubic system is given by,

$$d_{hkl} = \frac{a}{\sqrt{u^2 + v^2 + w^2}} \tag{2.2}$$

where a is the unit cell length.
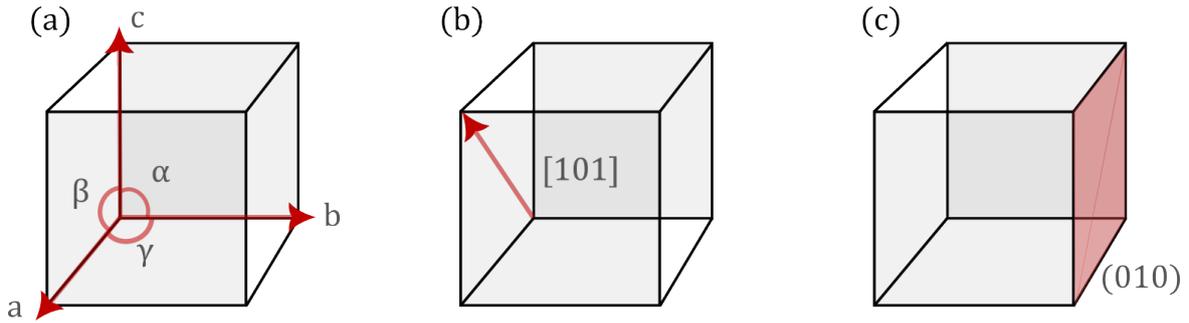
Figure 2.1: (a) Generic schematic representation of unit cell with lattice parameters a, b, c, α, β and γ. (b) Schematic drawing of direction [101]. (c) Schematic drawing of the (010) plane.

In three dimensions there are seven crystal systems; triclinic (Tr), monoclinic (M), hexagonal (H), trigonal or rhombohedral (R), orthorhombic (O), tetragonal (Te) and cubic (C). The differences between their lattice parameters can be seen in Table 2.1.

Table 2.1: The seven different crystal systems and their corresponding lattice parameters.

Name	Axial lengths	Axial angles	Centering	Characteristic (minimum)
Triclinic	a≠b≠c	$\alpha \neq \beta \neq \gamma \neq \frac{\pi}{2}$	P	1-fold
Monoclinic	a≠b≠c	$\alpha = \beta = \frac{\pi}{2} \neq \gamma$	P, C	2-fold axis along y
Orthorhombic	a≠b≠c	$\alpha = \beta = \gamma = \frac{\pi}{2}$	P, I , C, F	2-fold axis in three perpendicular directions
Rhombohedral	a=b=c	$\alpha = \beta = \gamma < \frac{2\pi}{3}$, $\neq \frac{\pi}{2}$	P	3-fold along z
Tetragonal	a=b≠c	$\alpha = \beta = \gamma = \frac{\pi}{2}$	P, I	4-fold axis along z
Hexagonal	a=b≠c	$\alpha = \beta = \frac{\pi}{2}$, and $\gamma = \frac{2\pi}{3}$	P	6-fold along z
Cubic	a=b=c	$\alpha = \beta = \gamma = \frac{\pi}{2}$	P, I, F	Four 3-fold axes parallel to body-diagonals of unit cell

The crystal systems might have multiple possibilities for defining the origin of the unit cell as all lattice sites are equivalent throughout the crystal. This is called centering of a lattice. If the unit cell contains only one atom in the unit cell, then it is a *primitive lattice*, denoted P. If the unit cell instead contains two atoms, assuming one is at the origin $[0, 0, 0]$, the other has the following possibilities of positioning:

$$\mathbf{A} = \left(0, \frac{1}{2}, \frac{1}{2}\right); \tag{2.3}$$

$$\mathbf{B} = \left(\frac{1}{2}, 0, \frac{1}{2}\right); \tag{2.4}$$

$$\mathbf{C} = \left(\frac{1}{2}, \frac{1}{2}, 0\right); \tag{2.5}$$

$$\mathbf{I} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right). \tag{2.6}$$

For non-primitive cells with four filled positions you have a **F**-centered cell with additional atoms in position **A**, **B** and **C**. **I** is called body centered and **F** is called face centered. Centering is shown for an orthorhombic lattice in Figure 2.2. The possible centering for the seven systems are given in Table 2.1. This results in 14 possible 3D lattices, the so-called Bravais lattices.
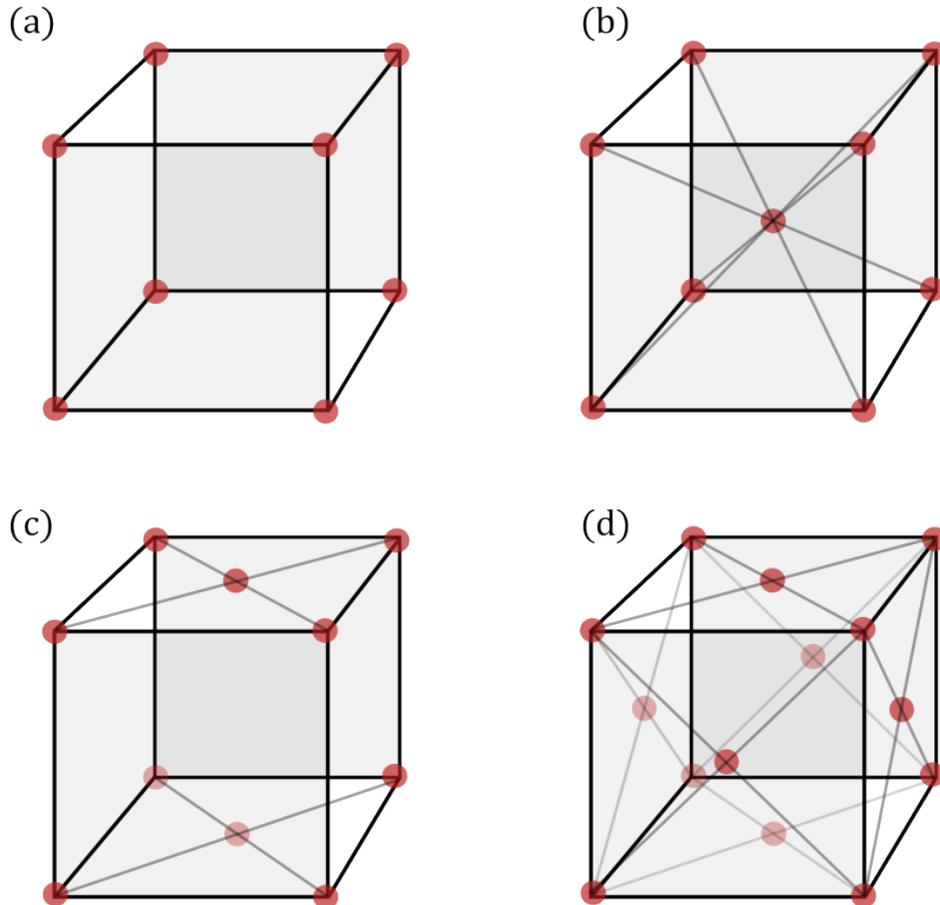
Figure 2.2: Orthorhombic unit cells with different centering. (a) **P**-centering which is primitive, (b) **I**-centering commonly called body-centered, (c) **C**-centering and (d) **F**-centering.

In addition to different centering, can a crystal unit cell have either the same atom(s) or a mix of different atoms in all lattice points. An example is the the cubic diamond structure that is common for elements like Si, Ge and C, which is structurally identical to the zincblende structure that contains two different atoms in the unit cell and is common for

GaAs, GaSb and InP. The difference between the unit cell of a diamond and a zincblende structure is shown in Figure 2.3.
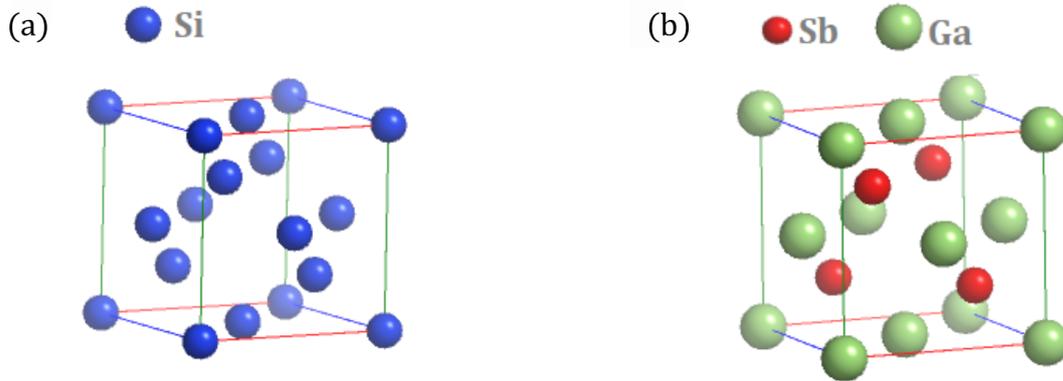
Figure 2.3: Diamond structure of Si (a) and zincblende structure of GaSb (b). Silicon (Si) atoms shown in blue, antimony (Sb) atoms shown in red and gallium (Ga) in green. Figures made with ReciPro [33].

Note that system and centering can be differently defined for the same configuration, for example, face-centered cubic can be described as a primitive rhombohedral lattice or a 2D centered rectangle can be described as a primitive diamond lattice, see Figure 2.4. The choice for a certain lattice is based on what is convenient. For the previous diamond-zincblende example, both diamond and zinblende structure are face-centered cubic, and can hence also be represented by a primitive rhombohedral unit cell. The lengths of the primitive rhombohedral unit cell, a_r, and the conventional cubic $\mathbf{F}$ cell, a_c, is connected by the formula,

$$a_r = \frac{\sqrt{2}}{2}a_c, \tag{2.7}$$

and can hence be transformed from one unit cell to the other. For computer programs and algorithms, finding the primitive unit cell is usually the best starting point for lattice constant determination, before searching for higher symmetry. The primitive cell is then later transformed into the conventional unit cell [34].
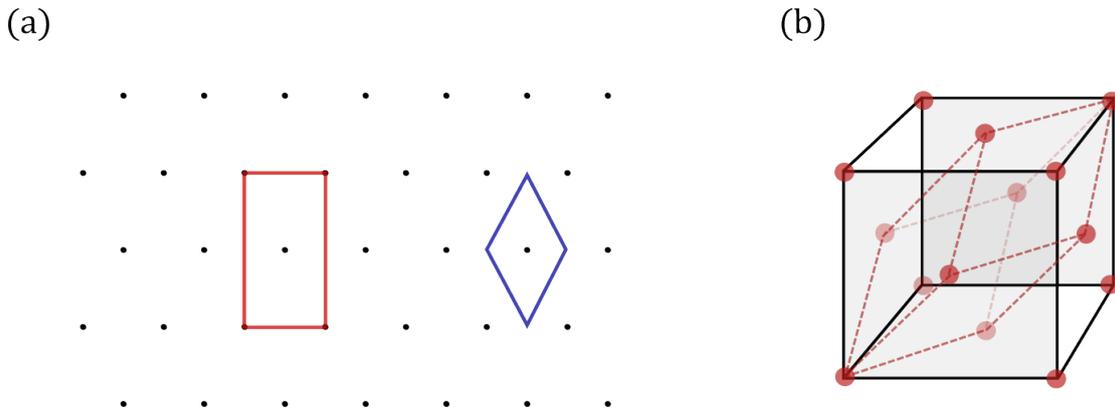
Figure 2.4: Different choices of unit cell for (a) 2D lattice and for (b) 3D cubic lattice. For the 2D lattice, we have a non-primitive centered rectangular unit cell in red and primitive diamond unit cell in blue. For the 3D lattice, we see that a face-centered cubic lattice (black) can also be described with a primitive rhombohedral unit cell here shown in dotted red.

The example materials are high symmetry materials of technological importance. Descriptions of the material will be extended in Section 2.1.4, where the full lattice description, including all possible lattice positions, will be addressed. However, first the basic crystal description will be reviewed.

2.1.2 Point groups

At a cell point there can be specific symmetries. For example rotational symmetry, as seen in Figure 2.5(a), where the point can be rotated an angle and still be invariant. It can also have reflection symmetry, seen in Figure 2.5(b) as well, where the point can be mirrored and still be invariant. Many crystals feature one or multiple of these types of symmetries, as well as combinations of them.
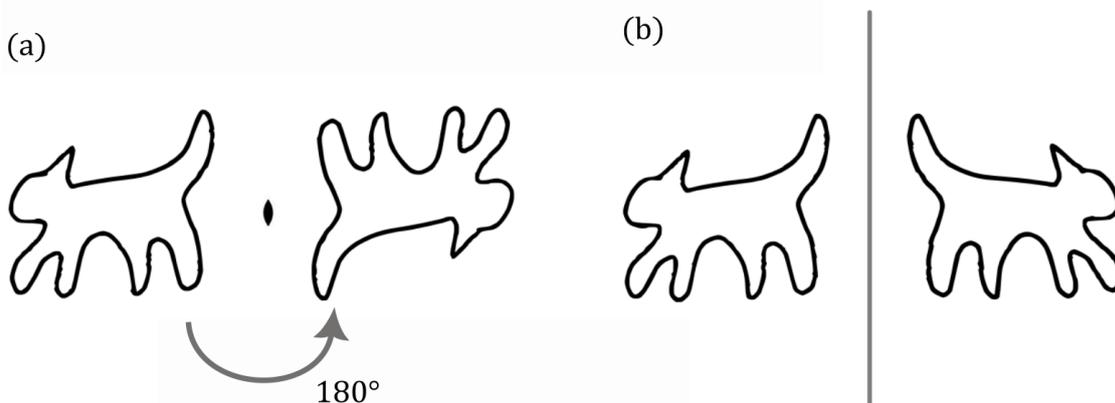
Figure 2.5: Symmetry in crystals visualized with an arbitrary pattern. Figure (a) shows rotational symmetry, here 2-fold with rotation axis marked with a lens-shaped symbol. Figure (b) shows reflection symmetry.

Point groups are mathematical groups of symmetry operations that preserves the overall structure of an object. A mathematical group is an abstract algebraic structure with the fundamental properties of closure upon multiplication, associative multiplication, existence of identity elements and inverse elements within the group. For point groups all of these conditions are satisfied. The symmetry operations involve rotations, translations, inversions and combinations of these, which are all isometric operations. All symmetries intersect each other in one point, which is the origin of the name point group. The point groups are a closed group with only 32 entries. A point group in itself is also closed, take point group 4 as an example, which has entries 1, 4^1, 4^2 and 4^3 and is hence of order 4.

There are two types of notation for the point groups, namely the Hermann-Manguin and the Schönflies notation. The Hermann-Manguin notation is based on the idea of listing the minimum amount of symmetry elements needed to describe a point group, while the Schönflies notation is based on group theory. The Hermann-Manguin notation is the international standard within crystallography and is considered more insightful for the purpose of this thesis, hence it will be the one used here.

Each point group is denoted by the type of operations that are possible and their corresponding angles or axes of rotation or reflection. In three dimensions there are a total of 32 unique point groups summarized in Table 2.2. These are combinations of the rotations of order 1, 2, 3, 4 and 6, as well as the possible rotation axes 222, 223, 224, 226, 233 and 234, with all possible combinations of inversion and mirror symmetry. The mirrors can be either parallel, denoted m, or perpendicular, denoted $1/m$. Thus even though point group 3m and

3/m both have a 3-fold axis, they still differ as they have a parallel and perpendicular mirror respectively. The specific point groups are associated with specific crystal systems, as their lattice parameters determines the type of symmetries that are possible. Symmetry can hence be used as a selective way of describing a crystal.

Table 2.2: Table of the 32 different point groups sorted by crystal structure and type of symmetry operations involved. Duplicates are coloured in red. Crystal systems are abbreviated Tr=triclinic, M=monoclinic, O=orthorhombic, R=rhombohedral, Te=tetragonal, H=hexagonal, C=cubic. Adapted from Table 2.11 in *Crystallography* by D. Schwarzenbach [35].

	Tr	M/O	R	Te	H	C
Rotation	1	2	3	4	6	23
Rotoinversion	$\bar{1}$	$\bar{2}=\mathrm{m}$	$\bar{3}$	$\bar{4}$	$\bar{6}$	
Rotation + ⊥ mirror		2/m	3/m $=\bar{6}$	4/m	6/m	m$\bar{3}$
Rotation + ⊥ 2-fold		222	32	422	622	432
Rotation + ‖ mirror		2mm	3m	4mm	6mm	
Rotoinversion + ‖ mirror		2mm	$\bar{3}$m	$\bar{4}$2m	$\bar{6}$2m	$\bar{4}$3m
Rotation + ‖ + ⊥ mirror		mmm	3/mmm $=\bar{6}$2m	4/mmm	6/mmm	m$\bar{3}$m

One can further define the different symmetry operations to be either of first or second kind. The first kind include rotations and translations, while the second one contains reflections and inversions. The symmetry operations of both kinds can be either pure, or combined to form new symmetry operations.

A pure rotation is defined by its rotation axis, denoted $[hkl]$, and rotation angle $\alpha = 2\pi/n$, that depends on the integer n which is the order of the rotation. Usually, a rotation of order n is said to be n-fold. Pure reflections around hkl-planes are denoted m for mirror and pure inversion flips points from a position r to the inverse point $-r$. A structure is called centrosymmetric if it possesses an inversion center as one of its symmetry elements, and non-centrosymmetric if not. Rotoinversion is a combination of rotation and inversion, and will hence both rotate $(2\pi/n)$ and invert $(r \rightarrow -r)$ a point through the inversion center. Rotoinversions are denoted by the order of the rotation with a bar $\bar{\mathbf{n}}$. The other combinatorial symmetry operations involve translation, which is the only of the previously mentioned pure symmetry operations that through repeated operation does not return to the initial point. This type of symmetry is further introduced in Section 2.1.3.

To summarize, together these symmetry operations lead to the 32 point groups listed in Table 2.2. Note that another convention for classifying the symmetry operations of a structure, is using Laue groups, which essentially are point groups with added symmetry. A structure's Laue group is the centrosymmetric parent point group of the actual point group of the structure. There are hence 11 Laue groups. These are commonly used for crystal structure solution algorithms, which will be described in Section 2.3.

2.1.3 Space groups

A crystal is a basis, described previously in Section 2.1.2 as point groups, placed on a lattice, described in Section 2.1.1. Space groups are symmetry operations applied to the crystal lattice that preserve the overall crystal structure. A space group is simply a convolution of the Bravais lattice and point group. In three dimensions there are in total 230 unique space groups that categorize the arrangement of atoms that are all listed in International Table of Crystallography A (ITA). Hence, this is a further refine of describing the structure of materials. Determining the space group is a crucial aspect of this work.

Along with the symmetry operations of the point groups, there are two additional translation symmetry operations that become relevant for the space groups. Pure translations are determined by the translation vector $\mathbf{T} = u\mathbf{a}_1 + v\mathbf{a}_2 + w\mathbf{a}_3$, where the integers decide the extent of the operation. The additional operations are combinations of rotation and translation, called screw axes, and combinations of mirror planes and translation, called glide planes. A screw axis is denoted by n_m where n is the order of the rotation and m is the extent of the translation. The pitch of the screw axis is defined by both n and m, $\mathbf{T} = \frac{m}{n}\mathbf{r}_{[hkl]}$, and the plane $[hkl]$ being the positive direction along the screw axis. The screw axis is said to be without hand if $m = 0$ or $m = n/2$, left-handed if $m > n/2$ and right-handed if $m < n/2$. Now a glide plane on the other hand has no rotations involved, but rather reflections combined with partial translation. This results in reflections along glide vectors sized fractions of a lattice vector parallel to mirror planes. Glide planes are denoted by letters a, b, c, n, or d that represent the corresponding glide vectors. a, b and c represent axial glide along different lattice vectors and represent the glide vectors $a_1/2$, $a_2/2$ and $a_3/2$ respectively. n represents diagonal glide, with possible glide vectors $(a_1 + a_2)/2$, $(a_2 + a_3)/2$, $(a_3 + a_1)/2$ and $(a_1 + a_2 + a_3)/2$, where the last one is only available for cubic or tetragonal systems. Lastly, we have d representing diamond glide with possible glide vectors $(a_1 \pm a_2)/4$, $(a_2 \pm a_3)/4$, $(a_3 \pm a_1)/4$ and for cubic or tetragonal systems we also have $(a_1 \pm a_2 \pm a_3)/4$.

All crystallographic symmetry elements can be represented by graphical symbols. A complete list of the graphical symbols used in the symmetry diagrams in ITA is represented in Figure 2.6.

2.1.4 Atom positions

So far geometry of repeating units are considered. However, within them there are special positions where scattering units (atoms) can sit. In describing a structure, these positions and what scattering unit sits on them need to be determined.

A unit cell can contain multiple structurally identical motifs arranged symmetrically according to the symmetry of the space group [36]. A motif can be an atom, a group of atoms, or molecules. These unit cell motifs can be interpreted as points related by symmetry elements. The coordinates of the points are expressed as fractions x, y and z of the unit cell lengths a, b and c and are commonly referred to as Wyckoff positions [4]. These coordinates can be interpreted analytically or graphically and specify where atoms are located in a crystal. In ITA, the graphical and analytical coordinates of the symmetry related points in the unit cell are listed for each space group [37].

There are two types of positions, special and general. General positions are located at arbitrary positions within the unit cell without lying on any symmetry elements. Special positions, on the other hand, are located on one or more symmetry elements. Since the special positions lie on symmetry elements, these points have a lower multiplicity than the general positions. For example, if a motif is located on a mirror-plane, there will only be this one equivalent point, but if a point is located outside the mirror-plane, there will be another
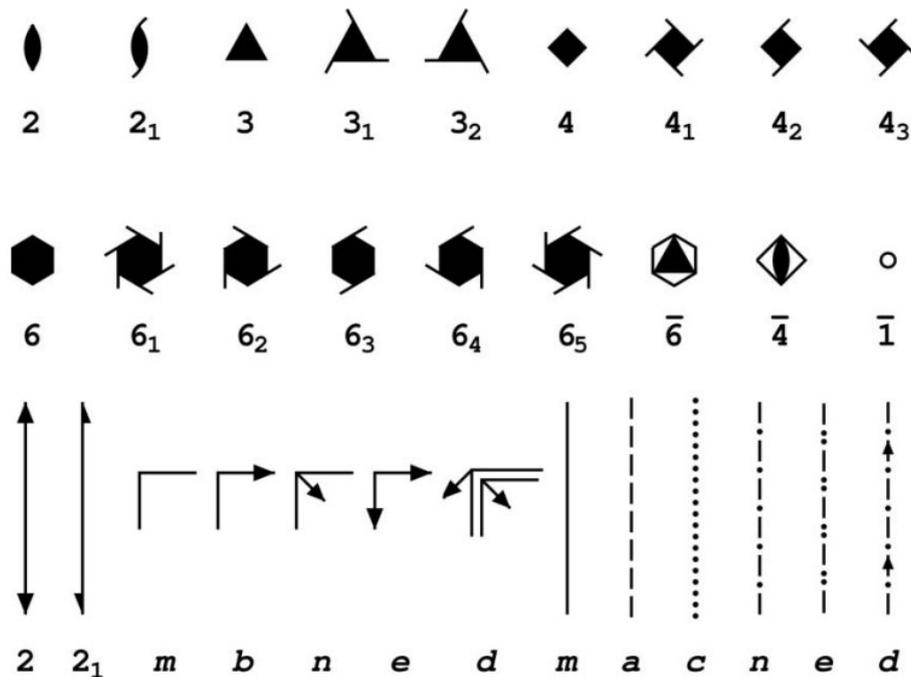
Figure 2.6: Graphical symbols of symmetry operators used in symmetry diagrams. Figure adapted from Dauter and Jaskolski [36] with permission of the International Union of Crystallography.

Table 2.3: Wyckoff positions for space group $Cmc2_1$ as given in ITA.

Multiplicity	Wyckoff letter	Site symmetry	Coordinates $(0, 0, 0)+ \quad (0, \frac{1}{2}, \frac{1}{2})+$
8	b	1	(1) x, y, z (2) -x, -y, z+$\frac{1}{2}$ (3) x, -y, z+$\frac{1}{2}$ (4) -x, y, z
4	a	m..	0, y, z 0, -y, z+$\frac{1}{2}$

equivalent point related to the point by the mirror-operator. As special positions typically are located on symmetry axes, inversion points or intersections between axes, they require that one or more of their fractional coordinates are fixed.

An example of a table of Wyckoff positions listed in ITA is shown in Figure 2.3 for space group $Cmc2_1$ (number 36). $Cmc2_1$ has two Wyckoff positions. The positions are sorted after decreasing multiplicity, hence will the general positions be listed first followed by the special positions with lower multiplicity. The Wyckoff letter labels positions alphabetically with a, b, c, up to i, not to be confused with glide type or lattice parameters. The letters are assigned from highest to lowest site symmetry, i.e. a is the position with the highest site symmetry and lowest multiplicity. The site symmetry reveals which symmetry the point passes through, using the same notation as for point groups where no symmetry is marked with a dot. This is visible in the 'Site symmetry'-column in Table 2.3, where the lower special position has site symmetry $m..$ corresponding to a mirror plane parallel to the **a**-axis. The general positions have no symmetry (only a one-fold axis).

The number of points in the 'Coordinates' column is visibly not the same as the number of atoms in the 'Multiplicity' column. This is due to the centering of the space group (it is
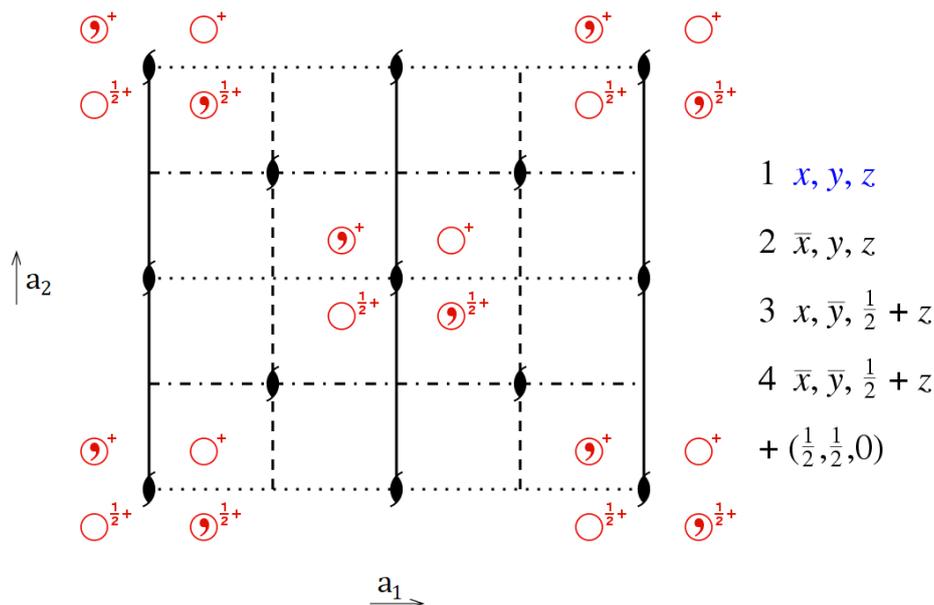
Figure 2.7: Graphical diagram showing the symmetry elements (in black) and equivalent points (represented by circles in red) for space group Cmc2$_1$ projected along in the a_1a_2-plane. The coordinates of the general points are listed to the right along with the additional centering vector. Figure adapted from *A hypertext book of crystallographic space group diagrams and tables* [38].

C-centered), hence the centering vectors $(0, 0, 0)$ and $(0, \frac{1}{2}, \frac{1}{2})$ at the top of the 'Coordinates' column need to be added to the listed points.

Figure 2.7 shows a graphical diagram of the symmetry elements in space group Cmc2$_1$ projected along in the $\mathbf{a}_1\mathbf{a}_2$-plane with positive $\mathbf{a}_3$-direction pointing towards the reader. The general points are marked with red circles. The fractional number next to the red circles correspond to the elevation of the coordinate compared to the base of the unit cell. Graphically a '+' sign indicates a positive z-coordinate, while a '-' sign indicates a negative z-coordinate. For example if $z = 0.3$, then $\frac{1}{2}-$ would mean $z = 0.5 - 0.3 = 0.2$ while $\frac{1}{2}+$ would mean $z = 0.5 + 0.3 = 0.8$. The general point coordinates previously seen in Table 2.3, are listed to the right of the projection along with the additional centering vector.

From a series of 3D diffraction patterns, one wants to find the crystal structure of a sample. If one is able to determine the space group of it, then the path is short to also finding all lattice parameters and atomic positions as there only are a few possible Wyckoff positions. The 230 space groups are summarized in ITA [37]. For each space group, the characteristics such as symmetry operations and options to place origin or assign axis are listed together with the Wyckoff positions, describing the possible positions and how many atoms could be at a certain point in that space group.

In this section, the systematic way to describe (single) crystalline materials is introduced using symmetry nomenclature and the ITA. This will be used in the present work. The next section describes how to collect crystallographic information based on the phenomena diffraction.

2.2 Diffraction

When a wave of sufficiently small wavelengths such as X-rays (0.1-10 Å) or fast electrons (2.51 pm at 200 kV) are directed at a crystal, they interact primarily interelastically with

the cloud of electrons in an atom which is placed on a crystal lattice. These interactions cause the incoming wave to scatter in phase in distinct directions and form a coherent diffraction pattern on a detector. The diffraction pattern contains direct information about the arrangement of atoms within the crystal, because the angles and intensities of the diffracted beams are directly related to the distances between atomic planes in the crystal lattice through Bragg's law. The intensity of the reflections give more detailed structural information. Crystal structure details as introduced in 2.1, will in this work be deduced from a series of patterns based on the (relative) position and intensity of reflections. In this section, the basic diffraction theory will be introduced which lies behind analyzing the experimental patterns. The section is based on [3, 7, 39, 40].

2.2.1 Bragg's law

Diffraction is a scattering phenomena where incident waves scatter coherently in distinct directions when hitting atomic planes in crystalline materials. The coherent interactions occurring in diffraction are elastic scattering events with fixed phase relations. To obtain structural information about a crystal lattice, we take use of Bragg's law, which relates the incoming waves to the desired structural information. This equation was derived in 1913 by Lawrence Bragg in his study of X-ray diffraction and describes the interference maxima of the coherent scattering event [41]. The wavelength λ of the incoming wave needs to be smaller than or roughly the size of the lattice constant of the crystal unit cell in order to be diffracted in different directions than the incident beam. If the wavelength of the incident wave is too big, we will only observe simple optical refraction. The general formulation of Bragg's law is given in Equation 2.8 and shown schematically in Figure 2.8,

$$n\lambda = 2dsin(\theta) \tag{2.8}$$

where n is an integer describing the order of the diffraction spot, d is the lattice spacing of the sample crystal and θ is the angle of which the incoming beam hits the crystal. This relation is only satisfied when $\lambda \leq 2d$, and is because of it's simplicity, extremely useful in diffraction. If you know the wavelength of the wave you apply on the sample, you can from the angle between the diffraction spots, determine the lattice spacing d of the sample crystal, which links to the lattice parameters through Equation 2.2 in Section 2.1. Evidently, Bragg's law makes diffraction a powerful tool for lattice parameter determination of crystals.

Figure 2.8 shows incoming waves being diffracted by the parallel atomic planes, functioning like a "mirror", when the Bragg-condition is fulfilled. There is a certain probability for the electrons to scatter multiple times in a sample material depending on the crossection for elastic scattering, the relative orientation of the beam compared to the crystal, and thickness. This effect, called dynamical scattering, is especially visible under electron diffraction, as electrons, being charged particles, interact stronger with matter than X-rays and neutrons. The more scattering events, the less predictability we have of what will happen to the electron, which makes the interpretation of images more challenging. The intensity of the diffracted beam will then depend non-linearly on the specimen thickness, in general favouring thinner samples as with thicker samples the likelihood for inelastic scattering increases which reduces the maxima and increases the background to the coherent diffraction pattern.

2.2.2 Reciprocal lattice

Diffraction can be described as a phenomena where waves are bent or scattered coherently when encountering obstacles or pass through narrow openings. As a mathematical tool for understanding diffraction has the reciprocal space been constructed, which is spanned out
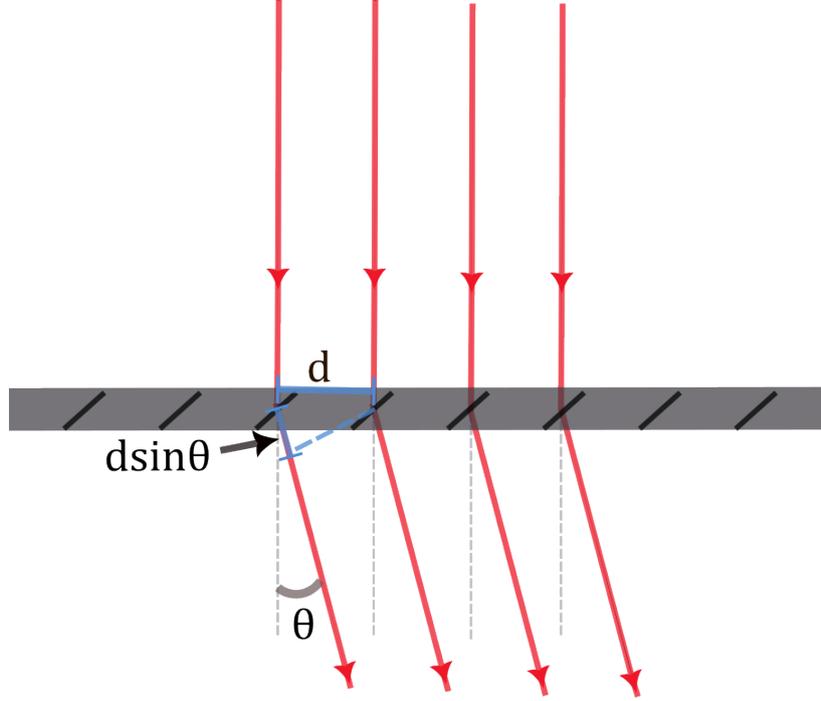
Figure 2.8: Atomic planes with spacing d in a periodic lattice scatter incoming waves of wavelength λ by an angle 2θ. This shows the relation between the parameters in Bragg's law (see Equation 2.8).

by three reciprocal lattice vectors that are perpendicular to the real space vectors. The reciprocal lattice therefore has the inverse dimensions of the direct lattice. This means that as the vectors of the direct lattice has dimensions [length], the reciprocal lattice vectors will have dimensions [1/length]. We say that the reciprocal lattice is in the Fourier space of the direct crystal.

The reciprocal vectors spanning the reciprocal lattice is given by the crossproducts of the direct lattice vectors,

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3}; \qquad \mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_2 \cdot \mathbf{a}_3 \times \mathbf{a}_1}; \qquad \mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_3 \cdot \mathbf{a}_1 \times \mathbf{a}_2}. \tag{2.9}$$

The reciprocal lattice also has a translational reciprocal lattice vector $\mathbf{G}$ expressed in Equation 2.10, in the same way that the real crystal lattice has a translational vector under which the crystal is invariant.

$$\mathbf{G} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \tag{2.10}$$

h, k and l are arbitrary integers of a point coordinate in reciprocal space, which correspond to a plane (hkl) in real space, see Section 2.1.1.

The spacing in the reciprocal lattice is d_{hkl}, which depends on the reciprocal lattice vector $\mathbf{G}$ by,

$$d_{hkl} = \frac{2\pi}{|\mathbf{G}|} \tag{2.11}$$

Bragg's law, as shown in Equation 2.8, can also be expressed in terms of properties of waves in the reciprocal lattice. We first note that waves can be described with a wavevector $\mathbf{k}$, capturing both the intensity and direction of the wave. The difference between an incoming wave $\mathbf{k}_i$ and a scattered wave $\mathbf{k}_s$, is commonly called the scattering vector $\Delta\mathbf{k}$ and measures the change in wavevector,

$$\mathbf{k}_s - \mathbf{k}_i = \Delta\mathbf{k}. \tag{2.12}$$

15

We observe constructive interference in a diffraction pattern when the scattering vector is a multiple of the reciprocal lattice vector $\mathbf{G}$. This makes up the second formulation of Bragg's law,

$$\Delta\mathbf{k} = \mathbf{G}, \tag{2.13}$$

which means there are only a certain allowed values for $\Delta\mathbf{k}$ that give Bragg diffraction.

Elastic scattering has it that the magnitudes of the wavevectors satisfy $k_i = k_s$, as well as $k_i^2 = k_s^2$. With this knowledge combined with Equation 2.12 and 2.13, we can thereby deduce another condition for Bragg diffraction,

$$2k \cdot \mathbf{G} = G^2. \tag{2.14}$$

This formulation of Bragg's law can easily be transformed back to the first direct space version Bragg's law seen in Equation 2.8, by exchanging $\mathbf{k}$ with it's equivalent $2\pi/\lambda$ and rearrange Equation 2.11 to exchange $|\mathbf{G}|$ with $2\pi/d_{hkl}$.

2.2.3 Laue equations

The Laue equations for Bragg diffraction are a rewriting of Equation 2.13. Broken down into three separate equations involving the three direct lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, the scattering vector makes up the solution for all these three equations.

$$\mathbf{a}_1 \cdot \Delta\mathbf{k} = 2\pi h \qquad \mathbf{a}_2 \cdot \Delta\mathbf{k} = 2\pi k \qquad \mathbf{a}_3 \cdot \Delta\mathbf{k} = 2\pi l \tag{2.15}$$

The Laue equations can be visually represented in the Ewald sphere. A 2D representation of the Ewald construction can be seen in Figure 2.9. When the sphere cuts through a reciprocal lattice point the Bragg condition is satisfied and the spot will appear in the diffraction pattern. The reciprocal lattice points visible in the diffraction pattern are commonly called reflections. If the incoming wave with wavevector $\mathbf{k}_i$ intercepts a reciprocal lattice point at the end of the vector, then scattering with an angle 2θ from the origin of $\mathbf{k}_i$ occurs, which also has an endpoint on the sphere only shifted by a reciprocal lattice vector $\mathbf{G}$ away from the endpoint of $\mathbf{k}_i$.
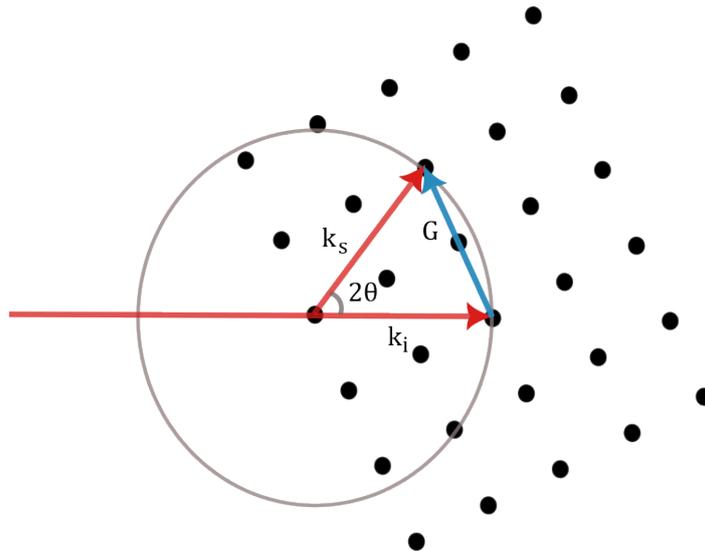

Figure 2.9: Schematic 2D construction of the Ewald sphere on an arbitrary reciprocal lattice. The Ewald sphere has a radius $k_i = 1/\lambda$ about the origin of the incoming wavevector $\mathbf{k}_i$.

The construction shown in Figure 2.9, visualizes that the sphere will be enlarged compared to the crystal lattice the smaller the incoming beam wavelength. Hence, more points will fulfill the Laue conditions and the resulting diffraction pattern will contain more solutions. As X-rays have larger wavelength than electrons, they will also have fewer points fulfilling the Laue conditions.

When using a transmission electron microscope (TEM), which is further described in Section 2.4, the crystal specimen is rotated and struck by beams of electrons with energy in the range 60-300 kV (wavelength 4.87-1.97 pm). In order to get a reciprocal lattice of perfectly defined lattice points, we would need an infinitely large crystal. Usually, the size of the crystal specimen is restricted, making the reciprocal lattice points have a certain spread. The use of thin specimens in TEM give rise to an elongation of the reciprocal lattice points parallel to the beam direction and a size inversely proportional to the thickness. The elongated reciprocal lattice points are called reciprocal lattice rods, for brevity called "relrods". When the Ewald sphere cuts through such a rod, we might observe a diffraction spot even though the Bragg condition is not satisfied. The relrod is said to relax the Laue conditions, which we can express as $\delta\mathbf{k} = \mathbf{G} + \mathbf{s}$, where s is a deviation parameter. The deviation parameter determines how far away from the actual reciprocal lattice point the Ewald sphere hits the relrod. s is negative if the actual point is outside the sphere and positive if inside the sphere.

What makes up the final diffraction pattern on our screen or camera is the projected intensities from the points in the reciprocal lattice that fulfill the Laue conditions, with the non-diffracted beam in the center of the image. If the Ewald sphere moves, the intensity of the reflections will change as relrods are crossed differently and the points fulfilling the Laue conditions change. We will soon see that the structure factor F_{hkl}, which depends on the angular-specific scattering ability of each individual atom in the crystal lattice, called the atom form factor f, is what determines the intensity of the reflections.

2.2.4 Intensities

Relating the repetition in real and reciprocal spacings or positions, give some structural information. However, for more detailed analysis the intensities in diffraction is crucial. The maxima of the electron density function are located where the atom positions in a unit cell are. To determine the electron density function is thus essential for crystal structure determination. As the diffraction pattern of a crystal sample is the Fourier transform of the direct lattice, crucial information about the direct space is available in the diffraction pattern through the Fourier transform. This applies also for the electron density. The Fourier expansion of the electron density in real space depends on the structure factors $\mathbf{F}_{hkl}$. The structure factors, with an amplitude and a phase, represent the diffracted waves from a diffraction experiment.

For a perfect crystal lattice with a reciprocal lattice vector $\mathbf{G}$ and assuming the diffraction condition in Equation 2.13 is satisfied, the scattering factor $\mathbf{F}$ is given by the following,

$$\mathbf{F} = N \int_{cell} dV \rho(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} = N\mathbf{F}_{hkl} \tag{2.16}$$

where N is the number of atoms inside the unit cell and where we define the structure factor $\mathbf{F}_{hkl}$ as the unit cell scattering factor.

The electron density $\rho(\mathbf{r}) = \rho(xyz)$ is a periodic function throughout the lattice and can through Fourier analysis be expanded to yield,

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} \mathbf{F}_{hkl} e^{-2\pi i(hx+ky+lz)}. \tag{2.17}$$

The hkl-suffix represent the atomic planes, see Section 2.1.1 and Equation 2.10, of which direction the structure factor correspond to. The structure factor can be considered a complex number or a vector with an amplitude and a phase, where the amplitude is the length of the vector and the phase correspond to the angle from the horizontal axis. Evidently, $\mathbf{F}_{hkl}$ as a complex number has a magnitude $|F_{hkl}|$ and a phase $e^{i\phi_{hkl}}$, and can be written as the product

$$\mathbf{F}_{hkl} = |F_{hkl}|e^{i\phi_{hkl}} \tag{2.18}$$

By further defining the electron density as a superposition of all electron concentration functions ρ_j associated with each specific atom j contributing on the specific position $\mathbf{r}$, we are able to determine the atomic form factor f_j as a function of $\mathbf{r} - \mathbf{r}_j$,

$$f_j = \int dV \rho_j(\mathbf{r} - \mathbf{r}_j)e^{-i\mathbf{G}\cdot(\mathbf{r}-\mathbf{r}_j)}. \tag{2.19}$$

The atomic form factor represents the individual scattering power of an atom. The structure factor can now be expressed as the sum of the atomic form factors for all atoms j in the unit cell multiplied by the phase factor,

$$\mathbf{F}_{hkl} = \sum_j f_j e^{-i\mathbf{G}\cdot\mathbf{r}_j} = \sum_j f_j e^{-2\pi i(hx_j+ky_j+lz_j)} \tag{2.20}$$

assuming the diffraction conditions are fulfilled and (x_j, y_j, z_j) being the integer coordinates of the jth atom. Here, the dot product between $\mathbf{G}$ and $\mathbf{r}_j$ is fully written out. The scattering factors $\mathbf{F}_{hkl}$, as here signal is coherent, describes the intensity I_{hkl} of the reflections in a diffraction pattern, as

$$I_{hkl} \propto |F_{hkl}|^2. \tag{2.21}$$

For the hkl-planes that makes $F_{hkl} = 0$, we have so called "extinction rules" that are kinematically forbidden reflections, hence giving no intensity in the diffraction pattern even if Bragg-diffraction conditions are fulfilled. The diffraction pattern of a crystal will look differently from different rotational angles. Some reflections will be invisible as a result of the extinction rules, and some will be much brighter than others. This also depends on what type of atoms are in the basis of a unit cell. The atomic form factor f varies depending on what type of element is present.

Let's look at the structure factor, and hence I_{hkl}, for the two different structures illustrated in Figure 2.3, diamond and zincblende structure. These structures both have eight atoms in the unit cell and can be considered a face centered cubic (FCC) inside another FCC, although diamond structure has eighth of the same kind, while zincblende has four of each type of atom. The atoms are positioned in the regular face-centered positions, as well as $(1/4, 1/4, 1/4)$, $(3/4, 3/4, 1/4)$, $(1/4, 3/4, 3/4)$ and $(3/4, 1/4, 3/4)$. The x, y and z positions of the unit cell atoms inserted in Equation 2.20, gives a simplified structure factor for the two structures,

$$F_{hkl} = \left(f_1 + f_2 e^{i\pi/2(h+k+l)}\right)\left(1 + e^{i\pi(k+l)} + e^{i\pi(l+h)} + e^{i\pi(h+k)}\right), \tag{2.22}$$

where f_1 and f_2 are the atomic form factors for the atoms in the $(0, 0, 0)$ and the $(1/4, 1/4, 1/4)$ basis respectively. For the diamond structure, $f_1 = f_2 = f$, hence can the atomic form factors be summed together. It is visible through Equation 2.22, that there are certain sums of h, k and l that give rise to specific intensities. For the diamond structure, we have,

$$F_{hkl} = 8f, \qquad \text{for } h + k + l = 4n \tag{2.23}$$

$$F_{hkl} = 4\sqrt{2}f, \qquad \text{for } h + k + l = 4n + 1 \tag{2.24}$$

$$F_{hkl} = 0, \qquad \text{otherwise.} \tag{2.25}$$

For the zincblende structure, we have,

$$F_{hkl} = 4(f_1 + f_2) \quad \text{for } h + k + l = 4n \tag{2.26}$$

$$F_{hkl} = 4(f_1 - f_2) \quad \text{for } h + k + l = 2n \tag{2.27}$$

$$F_{hkl} = 4(f_1 \pm if_2) \quad \text{for } h + k + l \text{ odd.} \tag{2.28}$$

Here we see the following: For both structures F_{hkl}=0 for hkl is mixed (i.e. not all even and odd integers, for example 001). The same kinematically forbidden are for other face-centered cubic lattices (see Section 2.1.3). In the diamond structure there are other planes, formed by additional occupied lattice positions, that give out-of-phase scattering and hence F=I=0, in this case for example hkl=002. The absence of these reflections can indicate another space group. In ITA these specific additional conditions, based on Equation 2.19, are given for the Wykhoff positions described in Section 2.1.4. Further, deviations from the ideal lattice position (in Section 2.1.4, x,y,z) will affect the intensities. Note that translation symmetry elements as glide and screw axis, see Section 2.1.3, can also give specific kinematically forbidden reflections and which is used for identification of these structural characteristics. The zinc blende example demonstrates that the intensity in a reflection depends on the atom form factor f_i, hence on the type of atom (how many electrons) located on a position. So to summarize, for a detailed structural analysis, the intensity of reflection is important.

Until now, we have used the assumptions of kinematical scattering theory, in which the incident wave is scattered only once. Dynamical scattering can cause kinematically forbidden reflections to appear in the diffraction pattern as well as altering the intensities of the resulting diffraction pattern. The parallel crystal planes might scatter the incoming waves multiple times, depending on the thickness of the material and the wavelength of the incoming wave. Hence, the reflection position is not affected, but the intensity is redistributed over the reflections in the pattern. Even kinematically forbidden reflections can have a non-zero intensity if they fit a summation of allowed reflections (eg. in diamond example above: 111+1-1-1=200). Kinematical diffraction is easier to interpret and model and sufficient for X-rays. For electron diffraction, dynamic effects need to be minimized (by thin specimens, summing of weak patterns). In the present work, we will assume that we can treat the data as kinematic and use proven X-ray based packages. After kinematic treatment, dynamic refinements can be made [42].

2.3 Crystal structure solution

With the term crystal determination, crystallographers commonly refer to the determination of unit cell parameters like a, b, c, α, β and γ and space group [4]. To further determine the atom positions and occupancies in the material, a refined structure analysis is needed where the intensities are used in a better way. The structure analysis determines the atom positions from diffraction data through finding the positions of the electron density maxima, as previously described in Section 2.2.4 and 2.1.4. The electron density is commonly visualized as a map either in two or three dimensions. The electron density map resembles a geographical contour map with peaks and valleys, where the peaks should correspond to atom positions [5].

The electron density is given in Equation 2.17 and is proportional to the structure factors seen in Equation 2.18. However, as the diffraction reflection intensities are only proportional to the magnitude of the structure factor, see Equation 2.21, information about the phase of the structure factors are lost in the experiment. The scattered radiation from

a diffraction event experience a phase shift ϕ_{hkl} caused by the spatial separation of the individual scattering objects. This problem is commonly encountered in crystallography and is referred to as the phase problem [43]. For a complete crystal structure analysis of a sample material, this issue needs to be taken into account. Hence, will a structure be solved when enough of the phases are elucidated from the intensities of the same reflections. There are different methods for doing this in structure analysis based on diffraction, described in Section 2.3.2, which are referred to as structure solution methods [4].

2.3.1 Figures of merit for structural models

In Equation 2.20, the atomic form factors f and the reciprocal lattice vector $\mathbf{G}$ are known quantities, meaning that the only remaining unknown is the atomic position vectors $\mathbf{r}$. This leads to a set of non-linear equations, which might have multiple solutions. Solving the non-linear equations therefore depend on an initial approximate solution that can be refined further until it fits experimental data.

A common criterion for evaluating the performance of the initial model, is the R index, given by

$$R = \frac{\sum_{\mathbf{G}} ||F_c| - K|F_o||}{\sum_{\mathbf{G}} |F_o|}, \tag{2.29}$$

where $|F_c|$ are the computed structure factor moduli, $|F_o|$ are the observed moduli and K is a scale factor to bring $|F_c|$ and $|F_o|$ to the same scale,

$$K = \frac{\sum_{\mathbf{G}} |F_c|}{\sum_{\mathbf{G}} |F_o|}. \tag{2.30}$$

Typical good values for the initial model would be $R \leq 0.5$ for centrosymmetric structures and $R \leq 0.4$ for non-centrosymmetric structures [4]. A structure is centrosymmetric if it possesses inversion centers, as was described in Section 2.1.2. For example in the diamond-zinc-blende example studied in Section 2.2.4, diamond structure is centrosymmetric, while zinc-blende is not.

Another parameter that can be used to evaluate the performance of the model is the normalized structure factor, commonly known as $\langle |E^2 - 1| \rangle$. The normalized structure factors converts the measured structure factor amplitudes $|F_{hkl}|$ into point atoms at rest and are derived from the probability distributions of the structure factors.

$$E_{hkl} = \frac{F_{hkl}}{\sqrt{\sum_j f_j^2}} \tag{2.31}$$

Naturally, for E_{hkl} to be normalized, $\langle |E_{hkl}|^2 \rangle$ must sum up to one, which it evidently does as $\langle |F_{hkl}|^2 \rangle = \sum_j f_j^2$. Since the probability distributions will be different for centro- and non-centrosymmetric unit cells, the normalized structure factors will differ as well. Common theoretical values for $\langle |E^2 - 1| \rangle$ are 0.968 for centrosymmetric structures and 0.736 for non-centrosymmetric structures [4]. These can be compared to experimental values to evaluate whether or not one has an inversion center.

In crystallography, correlation coefficients are commonly used to describe the consistency of the data. For example the half-set Pearson correlation coefficient, CC1/2, is a parameter that is used as a figure of merit for measuring data accuracy as well as the agreement of model and data instead of, or in addition to, R-values [44].

2.3.2 Structure solution methods

Structure solution methods are mainly divided into two categories: direct space methods and reciprocal space methods [4]. The direct space methods are the traditional methods that take use of Patterson functions and electron density modification methods. The reciprocal space methods, for example direct methods, take use of quantities defined in reciprocal space. In addition, there is an emerging new branch of structure solution methods that combines the two previous categories, called dual space methods. These methods alternately use information from both direct and reciprocal space to solve the phase problem. Recent developments of dual space methods like charge-flipping [45], intrinsic phasing [46], and molecular replacement [47] have been established as promising alternatives to the traditional direct methods.

In general, will structure solution methods require the space group and lattice parameters to have been determined in advance, which is usually done by studying the Laue symmetries, systematic absences and statistical tests [46]. Direct methods for solving small-molecule crystal structures take use of the probability relationships of strong-reflection phases. Dual-space methods, such as charge-flipping, correct the phase problem by using iterative Fourier transforms [45, 48]. One of the most promising recent methods is intrinsic phasing, which is embedded in the computer program SHELXT, that takes use of a novel dual-space algorithm to account for the phase problem for single-crystal reflection data [46]. The data is first merged according to the detected Laue group, see Section 2.1.2, and further expanded to the nominal space group P1. Space groups in the same Laue group are all tested to see which one fits best. This method has been shown to work well even with lower completeness and data quality [46].

2.3.3 Refinement of the structural model

In practice, there will be need for both a structure solution algorithm and a refinement algorithm. The structure solution and refinement can be interpreted as solving a set of nonlinear equations, where the reflection intensities are the equations [5]. This implies that it is essential to have at least as many equations as variables.

A typical flowchart of the crystal structure refinement algorithm is shown in Figure 2.10. First, will the intensities based on the initial model, $|F_c|^2$, be calculated. These values will be compared to the observed intensities, $|F_o|^2$, from indexed diffraction patterns, typically through R-values. The refinement algorithm will then make small iterative changes to the structure based on the initial model from the structure solution to obtain a better fit between the calculated, F_c, and observed, F_o, structure factors.

When the refinement process reaches convergence, the final model is finished. Usually, the refinement is considered successful when R-factors are minimized and the electron density map corresponds well to the refined model.

A commonly used algorithm for refinement is SHELXL. SHELXL uses least squares method to calculate parameter shifts between the observed and calculated structure factors for each cycle. Least squares is a common regression analysis method for parameter estimation, where the squares of the residuals are minimized [49]. SHELXL allows constraints and restraints to improve the fitting of experimental and observed data. Special positions are accounted for with automatic constraints by SHELXL [27].
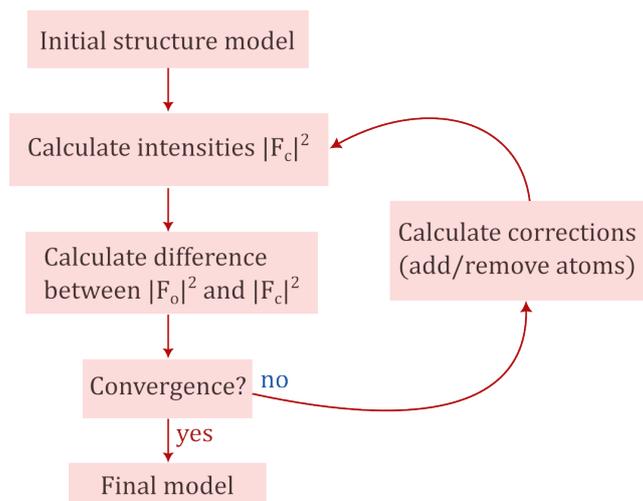
Figure 2.10: Flowchart of a refinement algorithm. The refinement should proceed until it reaches convergence.

2.4 Transmission Electron Microscopy

The resolving power of a microscope is dependent on what radiation it uses to illuminate the sample and the resolving power of the radiation is determined by its wavelength. This is expressed in the Rayleigh criteria for diffraction limited spatial resolution [50]. Microscopes using visible light, with an average wavelength of 550 nm, are therefore limited to only distinguish distances of that size. To see even smaller features one needed to use a type of radiation with even smaller wavelength, for example electrons. That is why transmission electron microscopes (TEM) uses accelerated electrons in the range 2-5 pm to achieve high spatial resolution [50]. This section will first go through some basic electron optics, before describing the components and set-up of a TEM. The section is mainly based on DeGraef, Ch. 4 [6], and Michler and Lebek Ch. 3 [50].

The 3D electron diffraction (3D ED) technique that this thesis is based on, requires the use of a TEM. The TEM uses thin samples (typically below 100 nm) and high energy electrons (in range 60-400 kV typically), which gives it an increased spatial resolution compared to other microscopy techniques. In addition, the strong interaction of electrons with matter, allows for analysis of small volumes (nm^3 to a few μm^3) compared to X-rays (μm^3-mm^3) or neutrons (mm^3). The TEM can be used for both imaging, spectroscopy and diffraction of the same small sample volume, making it a very versatile tool. Flexible lens systems makes it possible to operate the TEM in a broad range of magnifications and different camera lengths for diffraction. The two most common TEM-modes are scanning transmission electron microscopy (STEM) and conventional transmission electron microscopy (CTEM). STEM uses a convergent electron beam to scan the sample specimen and build up an image pixel by pixel. This mode differs from CTEM, which uses a coherent parallel electron beam and collect a coherent signal. In this work a coherent static broad electron beam will be used to collect the diffraction signal.

2.4.1 Electron wavelength

A fundamental principle of quantum mechanics is that particles have a wave-like nature. Electron beams can therefore be described as a plane wave. The particle and wave properties

of an electron can be described by a relation known as the de Broglie wavelength,

$$\lambda = \frac{h}{p} \tag{2.32}$$

where λ is the wavelength, h is the Planck's constant and p is the the magnitude of the particle momentum. The particle momentum is the product of its speed v and mass m,

$$p = mv. \tag{2.33}$$

If a charged particle, like the electron, passes through an area of large potential difference V, it will be accelerated to high speeds v. It might actually be accelerated so much that it approaches the vacuum speed of light c. In that case, we have to take relativistic effects into account. The mass of an electron m experiences a change in comparison to its rest mass m_0 when accelerated to high speeds, expressed as

$$m = \frac{m_0}{1 - \frac{v^2}{c^2}}. \tag{2.34}$$

The energy of the electron, which has electric charge e, also change during this transit through the potential V,

$$mc^2 = m_0 c^2 + eV. \tag{2.35}$$

If now the Equations 2.33, 2.34, and 2.35 are combined and all put into the de Broglie equation, we get

$$\lambda = h \left[2eVm_0 + \left(\frac{eV}{c} \right)^2 \right]^{-\frac{1}{2}}, \tag{2.36}$$

which is only dependent of constants and the adjustable accelerating voltage V. This means that we can directly control the wavelength of the electrons in a TEM, by adjusting the accelerating voltage. Table 2.4 shows the corresponding wavelengths of different accelerating voltages between 60 and 300 kV.

Table 2.4: Electron properties as a function of accelerating voltage

Acceleration voltage V [kV]	Wavelength of electrons λ [pm]	Ratio electron to light velocity v/c
60	4.8661	0.446
80	4.1757	0.502
100	3.7014	0.548
200	2.5079	0.695
300	1.9687	0.777

Interaction of charged electron with matter is strong, allowing small volumes can be analyzed. In this work we primarily focus on coherent elastic scattering (diffraction). Inelastic interactions as for example formation of characteristic X-rays, although useful on its own, will give a reduced signal of diffraction maxima and an increased background noise. Hence, deteriorate the diffraction pattern used here. Inelastic interaction will not be discussed further in this chapter.

The electron in a TEM is guided through the column by electromagnetic lenses. The external fields originating from the lenses conduct a total force,

$$\vec{F} = q\vec{E} + q(\vec{v} \times \vec{B}), \tag{2.37}$$

on the electron, where q is the elementary charge of the electron and v is the velocity of the electron. In absence of an electric field $\vec{E}$, we denote the remaining magnetic force as the Lorentz force,

$$\vec{F} = q(\vec{v} \times \vec{B}). \tag{2.38}$$

This force could potentially lead to rotations of the signal relative to the position of the object [51].

2.4.2 Instrument components

Figure 2.11 shows a schematic setup of a TEM. The schematic TEM setup includes an electron gun, condenser system, objective system, sample specimen in a holder that can be controlled over 5 axes, intermediate lens, projector system and detector plane. Electrons from a field-emission or thermionic source is first accelerated to high voltages by the electron gun. A field emission gun (FEG) is usually preferred, as these have higher stability, coherence, and brightness than a thermionic gun. The electron beam will further continue the path towards the sample through a series of electromagnetic lenses and deflector coils, controlling the electron beam when going through the column. The condenser system, consisting of the condenser lenses and condenser aperture (CA), together with the upper part of the objective lens allows to control the illumination (parallel or convergent, i.e. illumination angle α) and flux per area. The deflector coils are usually present between different segments of the column to put the beam on the optical axis of a segment or shift the beam. The specimen is in the objective lens as a high magnetic field is required to focus the electron beam, since the electrons have high speed and low mass. The CAs are present to limit the angular range of the electron beam, removing electrons traveling further away from the optical axis, that are more vulnerable to geometric aberrations. The lower objective lens focuses the scattered beams from the sample to form a diffraction pattern in the back focal plane (BFP). At this plane there is an objective aperture whereby the direct or a selected beam can be selected for bright-field and dark-field imaging respectively. At the image plane located under the BFP is a selected area (SA) aperture. This allows selection of a part of the specimen area illuminated that is contributing to the formed diffraction pattern. Practically, an area down to 100 nm in diameter can be selected, although as the SA aperture and specimen have the lower objective lens in between them, there could be a displacement between them due to spherical aberrations [40].

Figure 2.12 shows a more detailed view of the scattered trajectories of electrons after going through the specimen to form an image. Here, the red and green beam paths illustrate scattered electron trajectories in opposite distinct angles as a result of electron beam hitting parallel crystal planes. The electrons that are scattered from the same crystal planes, are brought to the same point in the back focal plane. The electrons originating from the same point in the sample on the other hand, are brought to the same point in the image plane.

The intermediate lenses allow magnification of the image of choice, either in diffraction or image mode. The projector lens is used to project the final wave-front onto the imaging system or viewing screen. When utilizing a TEM one can easily switch between projecting either the diffraction pattern from the back focal plane or the image from the image plane onto the viewing screen dependent on what is desired.
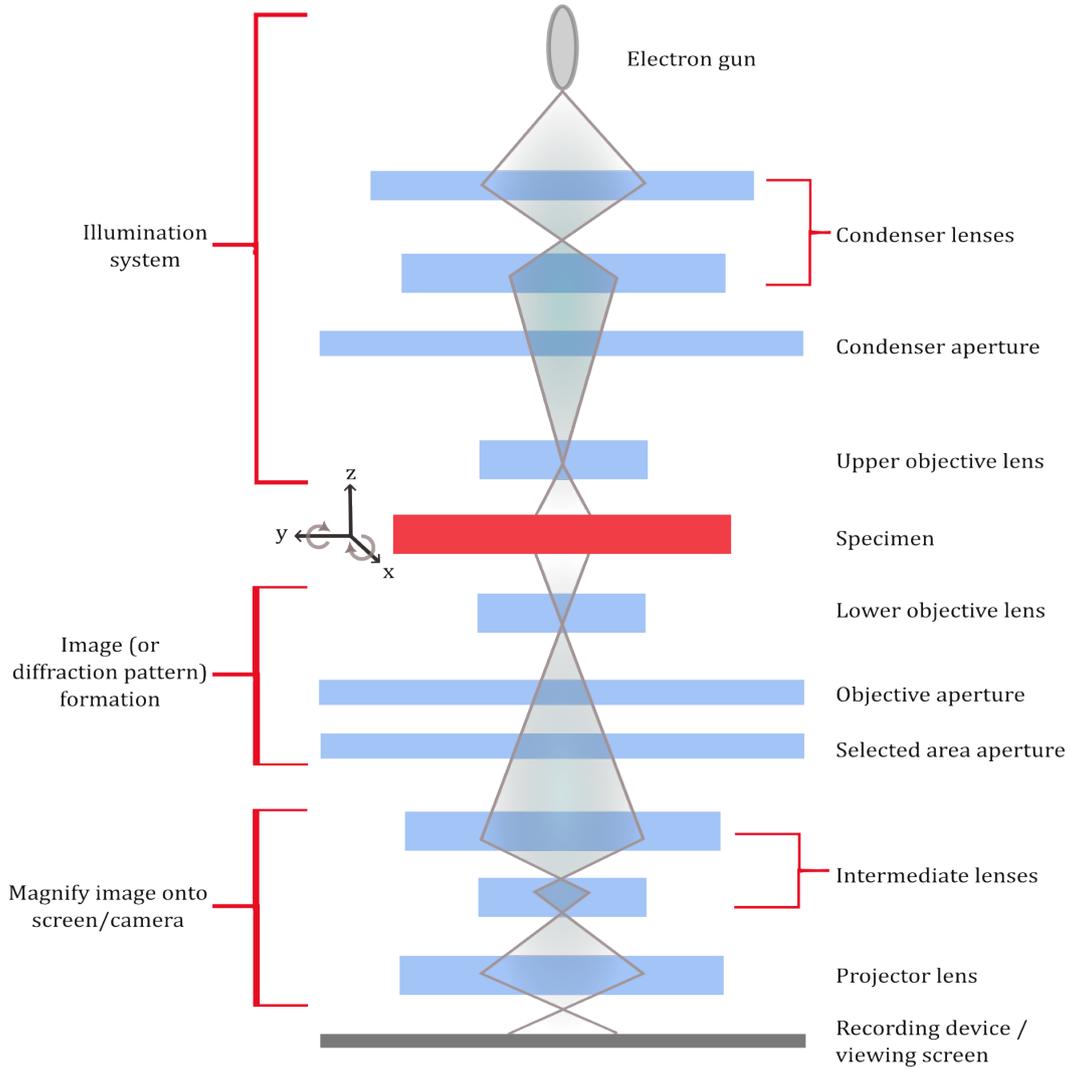
Figure 2.11: Schematic setup of a TEM. The grey lines illustrates the electron beam path through the lenses and apertures to form an image. The specimen is free to move in x-, y- and z-direction and can be tilted.

The camera length L of a TEM describes the magnification of a diffraction pattern. Although the camera length might sound like a physical distance, it is rather a calculated value. Figure 2.12(b) shows the conceptual camera length in the imaging system of a TEM, only without the lenses. The distance between the diffraction spots on the viewing screen R is related to the camera length L by,

$$\frac{R}{L} = tan2\theta \simeq 2\theta, \tag{2.39}$$

assuming θ is small. Using this assumption also for Bragg's law from Equation 2.8, where $sin\theta \simeq \theta$, we can express

$$Rd = \lambda L. \tag{2.40}$$

Hence, given the physical pixel size of the detector, the camera length (λL) can be determined from the data which is crucial for finding the lattice parameters.

(a)

(b)
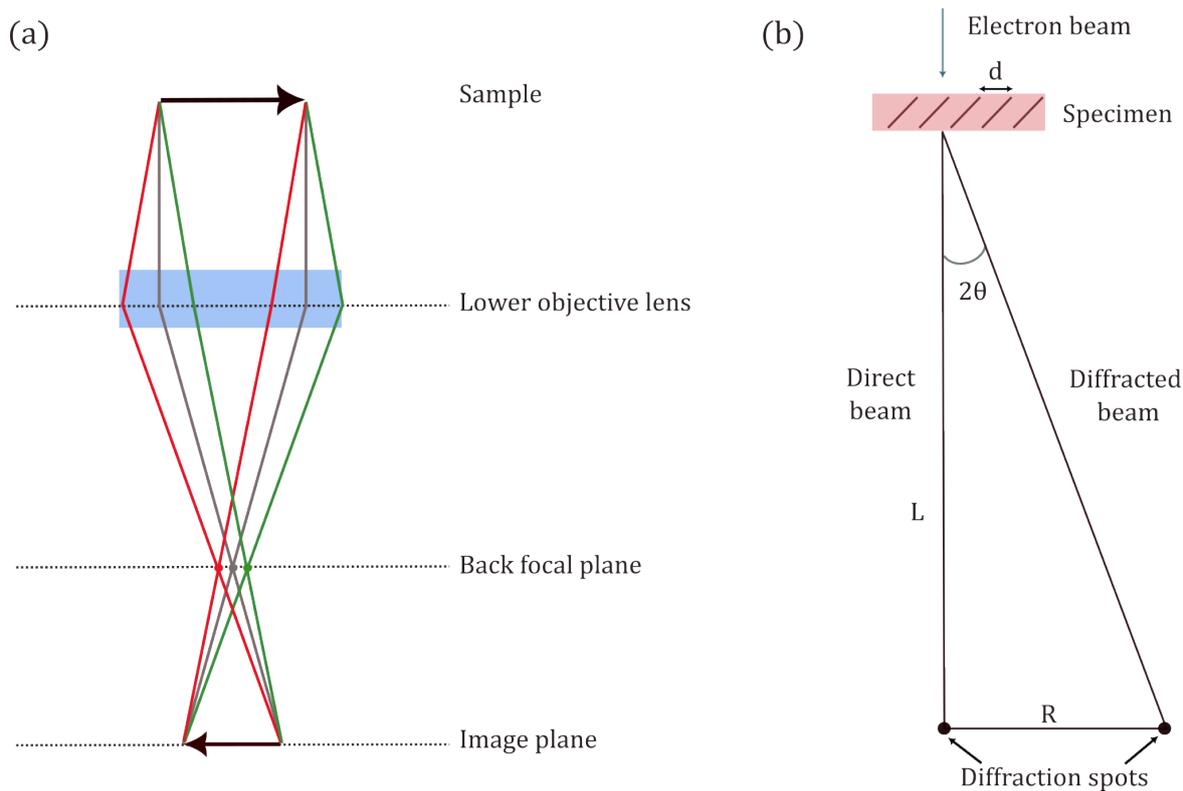
Figure 2.12: (a) Electron beam trajectories in a TEM after being scattered by the sample. The direct beam path is shown in grey, while the scattered beam paths are shown in red and green. Diffraction images are formed in the back focal plane, a flipped image of the direct space is formed in the image plane. (b) Conceptual camera length L of a TEM.

2.4.3   Electron detection

There are multiple ways of detecting electron signals in a TEM. A charge coupled device (CCD) detector record images by storing charge generated by light or electrons. The CCDs are metal-insulator-silicon devices stacked in arrays that usually contain millions of pixels. A pixel in this sense is a silicon-based photodiode, an electrically isolated capacitor as a result of the potential well created under each CCD cell [52]. The stored charge in the pixels are read out one by one and further amplified and digitized. CCDs are vulnerable to direct exposure of electrons as the electrons cause long-time fading of sensitivity of the pixels. Hence, are CCDs usually covered by a medium that converts the incoming electrons to photons, a scintillator. A possible problem with CCD detectors is blooming, that is when a pixel is oversaturated and the charges spreads to other nearby pixels.

A complementary metal oxide semiconductors (CMOS) image sensors consist of a multilayer stack of semiconductor materials and proposes an alternative type of electron detection to CCDs. CMOS are voltage-driven detectors, meaning the charge is converted into current directly in each pixel. Compared to the CCDs, are CMOS less prone to blooming and can have faster readout speeds [53].

The relatively new direct electron detector (DED) is directly exposed to the electron beam, in contrast to CCDs and CMOS that both use a scintillator to convert electrons to photons that reduce the detective quantum efficiency (DQE) of the device [51]. The main advantage for the present study is the high dynamic range, fast readout (hence short exposures) and low noise levels. A disadvantage is that pixels, each with separate readout,

are relatively large (50 μm) and as the space in TEM is limited the max number of pixels is limited compared to CCD/CMOS.

The dynamic range of CMOS or CCDs is defined as the maximum achievable signal divided by the camera noise. A higher dynamic range implies a higher ability of quantitatively measuring dimmer intensities in an image. If dynamic range is too low then one might get weak signal or saturated reflections. As intensity is important for structure solution and refinement, both should be avoided.

The image quality achievable in a TEM is mainly determined by point-spread functions (PSF), which describes the radial distribution of electron intensity originating from a single point source. In diffraction patterns of crystal structures, the PSF tails overlap, which reduces contrast unless the atomic spacing is sufficiently large for the central peak to be sharp enough to ensure visibility. The PSF are high tension dependent, thus overlap more at higher acceleration voltages. Information about the spatial frequency in the resulting image can be determined from the central peak of the PSF, whereas the tails of the PSF determines the overall contrast [54].

2.4.4 Beam damage from TEM

Knowing the full functioning of the TEM, the next step is to describe the effect the TEM can have on the sample material studied. There are different types of damage that materials can suffer from induced by a TEM: knock-on damage, charging effects, radiolysis, and thermalization. The damage is caused by interactions between the beam electrons and the atoms in the sample. The interactions can be inelastic or elastic. Inelastic collisions arise from interactions between incident electrons and atom specific electrons in the sample specimen. Elastic interactions arise from incident beam electrons interacting with the electric field of the atom nucleus. The elastic interactions cause atom displacements and may give bulk knock-on or surface sputtering damage, which is common for conducting samples, although uncommon for insulating materials such as zeolites, which will be further described in Section 2.6 [31]. The damages induced by the inelastic collisions are classified on behalf of the de-excitation and momentum conversion mechanism. There are mainly two different ways of de-excitation, radiative or non-radiative. Radiolysis (non-radiative) involves dissociation of chemical bonds as a result of exited outer-shell electrons. The effect of radiolysis on structural composition is inversely proportional to acceleration voltages, which is opposite of knock-on damage. For higher voltages (>150 kV) knock-on damage is hence more likely, but for lower voltages, radiolysis is the dominant damage type.

Other types of inelastic collisions involve thermalization (radiative) and charging effects (non-radiative). Charging effects occur when electrons are ejected as secondary or Auger electrons. Reducing the incident beam current can mitigate damage from charging effects. During thermalization, energy from excitation will be converted to phonons (lattice vibrations) that transfer energy throughout the lattice. This can potentially lead to a temperature increase in the sample that additionally enhances structural phase instability and atomic diffusion causing defect aggregation [31].

A common way to measure the effects of irradiation is using the physical quantities dose (C/nm^2), dose rate (C/nm^{2}s), or Gray units (Gy = J/kg) [55]. The characteristic dose is defined as the maximal charge density that a material can tolerate before undergoing irreversible morphological, structural, or chemical changes. As electrons have fixed charge, it is in TEM also possible to use electrons/Å^2 as a measure of irradiation. The X-ray community commonly uses grays. Conversion formulas have been developed between the different units [56, 55].

2.5 Electron diffraction for structure analysis

The most common diffraction experiments for structure analysis are conducted with X-ray methods such as single crystal X-ray diffraction (SCXRD) or powder X-ray diffraction (PXRD). X-rays have a wavelength of 10 nm to 0.01 nm, which is perfect for studying atomic spacing of crystals as they are usually within the range of a few Ångströms. Weak X-rays can go through thick samples without undergoing additional scattering events and a rather straightforward analysis can be done based on kinematic diffraction, which makes X-ray diffraction preferable for bigger samples. Pure crystals can, however, be hard to grow from scratch to large volumes, which poses a practical limitation for XRD methods. A TEM is primarily designed for imaging, rather than as a dedicated diffractometer, although recently dedicated 100 kV diffractometers based on TEM has come to the market [17]. Electrons interact stronger with matter than X-rays, because the electrons and matter experience strong Coulomb interactions. This gives a high signal-to-noise ratio also from very small volumes (nm^3) like the ones used in TEM, which makes it easier in principle to identify light atoms like hydrogen and lithium than in XRD [10]. Electron diffraction can be more beneficial than SCXRD on small volumes also because the TEM possibly can have parallel nanometer-sized electron probes, and it takes considerably shorter time to record the diffraction pattern. Unless having a synchrotron with a high brilliance, it can take minutes to hours to obtain XRD diffraction patterns, while it for electron diffraction takes only a fraction of a second for a single electron diffraction pattern in a conventional TEM set up. Other practical advantages are that one can switch between diffraction and imaging, see Section 2.4.2, and combine with electron spectroscopy. A risk is inelastic scattering that lead to higher background (can be filtered out) and damage of the illuminated crystal, although the ratio for elastic vs inelastic is more favorable for electrons than for X-rays [57].

Electrons do however, experience dynamical effects as mentioned in Section 2.2.1, where multiple scattering events occur as the electron proceeds through the sample. This may give a far more complex redistribution of the scattered intensity as function of structure and specimen thickness. Dynamical effects can be reduced by limiting thickness. Furthermore, can beam precession help reduce dynamical effects [58]. By tilting the beam away from the zone axis, one avoids the direct Bragg conditions that cause the most prominent dynamical effects while preserving the geometry of the zone-axis pattern [59]. Unlike inelastic scattering, that gives background noise, dynamical intensity variations can not be filtered out from the signal. However, dynamical effects are not all bad, in fact an advantage of dynamical effects is enhanced determination of non-centrosymmetric space groups. As far as kinematical Fourier approximation applies, the result is centrosymmetric. One of the main reasons for choosing electron diffraction over X-ray has therefore been that non-centrosymmetric groups can relatively easily be deduced compared to X-rays [14].

2.5.1 Conventional electron diffraction

Furthermore, conventional TEMs and recording media have high electron dose set-ups, where beam damage can be a limitation. Low-dose strategies and new detector technology can minimize this beam damage [53].

Note here the volume contributing to the diffraction pattern is selected by selected area electron diffraction (SAED) aperture and the Bragg spots are sharp. convergent beam electron diffraction (CBED) is another common electron diffraction technique, which gives discs with intensity distributions and features within the disc stemming from further out in reciprocal space. These can be used for accurately determining lattice parameters a, b, c and α, β, γ as well as determining point and space groups [40, 6]. However, such approaches

are manual and can require initial information to do simulations. CBED and other forms of electron diffraction based on combination of inelastic and elastic (Kikuchi) will not be used in this work for structure analysis.

2.5.2 3D electron diffraction

Diffraction patterns are, as mentioned in Section 2.2.1, two dimensional representations of reciprocal space giving information about the three dimensional structure in real space. However, a single diffraction pattern is not necessarily enough to know the full 3D crystal structure. Indeed, we need to tilt the crystal in order to get 3D information and be able to conduct a full 3D structure analysis. Discrete tilt series of on-zone diffraction patterns is a possible way of doing this. This method will however, only be able to collect a limited range of reflections and suffer from severe dynamical effects [10]. Vincent & Midgley then introduced precession electron diffraction (PED), where the tilted beam is precessed on the specimen around the TEM optical axis to reduce dynamical scattering, although data completeness still remained a problem [58]. However, Hovmöller et al. showed that structure determination with ED was possible even with dynamical effects present [60]. Using ED tilt series for structure determination gets the overall term 3D ED and became an up-and-coming method for analysing a wide range of materials, such as inorganic materials, organic molecules, proteins, zeolites, covalent organic frameworks (COF)s, metal-organic frameworks (MOF)s, and pharmaceutical compounds [32, 61, 62]. Both Kolb et al. and Hovmöller et al. developed softwares to analyze 3D ED data independently from each other in the early 2000s [63, 21]. Since then, several groups have developed protocols, software, and routines for 3D ED data processing. Evidently, 3D ED have made huge technological and methodological steps forward the last decade, and is still under constant development [42].

Through the years that 3D ED have been used, different data acquisition methods and routines have been developed, such as automated diffraction tomography (ADT)[12], electron diffraction tomography (EDT)[21], and single crystal electron diffraction (SCED) [13]. All of these three techniques sample 3D reciprocal space by rotating the sample around some arbitrary axis and collect the electron diffraction pattern at each angle, as shown schematically in Figure 2.13 [21]. This implies that the resulting 3D ED data is a stack of multiple frames of diffraction patterns. The data is later reconstructed in 3D reciprocal lattice, and intensities of the observed reflections are extracted and indexed. The first developed methods were done with stepwise rotation, although continuous rotation methods like microcrystal electron diffraction (MicroED) [11], fast electron diffraction tomography (Fast EDT)[64], rotation electron diffraction (RED), and continuous rotation electron diffraction (cRED) were developed soon after and established as the better alternative as it can be automated and reduces dose [10, 9].

The method have previously faced some problems related to the dynamical effects of electron scattering. These problems were solved by changing data collection strategies, whereas researchers went from orienting the crystal in low-index zone orientations, to orient the crystal along a random off-zone axis, the goniometer axis of the TEM stage [9, 63]. This change reduces dynamical effects, resulting in the obtained intensities from 3D ED being treated as kinematical intensities. Hence, 3D ED data can, in most cases, be assumed similar to the SCXRD data, implying that its possible to apply X-ray diffraction (XRD) routines and software to the 3D ED data. If the absolute structure is desired, a dynamical analysis should be taken into account [42]. Another strength of the 3D ED method, is that it does not have, nor need, a good spatial resolution because the spatial resolution is compromised by tilt stability. Still the analysis can be done on much smaller volumes than with X-rays and relative simple (i.e. cheap) set-ups. One of the key breakthrough papers demonstrating

(a) Sample illumination   (b) Diffraction patterns   (c) Reconstructed 3D ED
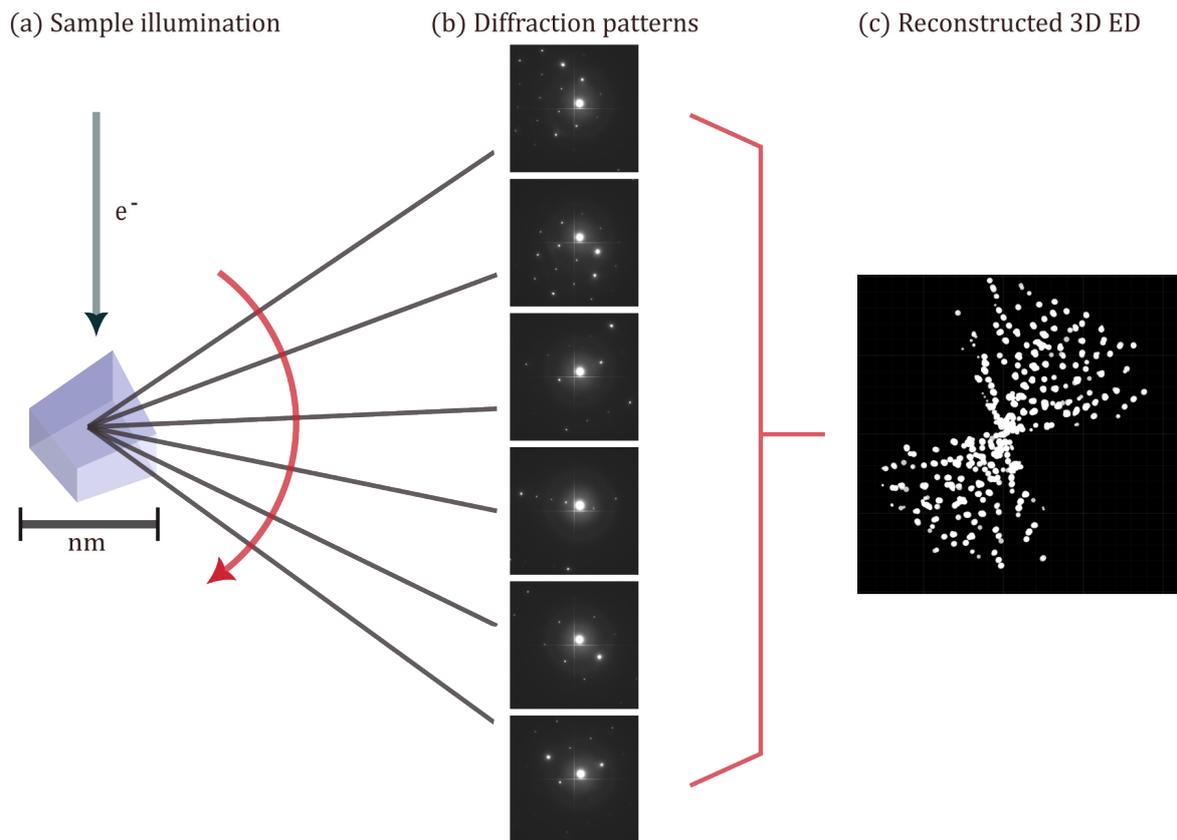
Figure 2.13: Overview of the 3D ED process. (a) A sample is illuminated with electrons in a TEM. (b) Diffraction patterns are collected for each angular range. (c) The diffraction patterns form a stack of frames which can be used for a 3D reconstruction of the sampled reciprocal space.

the power is Palantinus' Science paper in 2017 [15, 16].

Figure 2.14 shows a schematic drawing of the continuous rotation data acquisition method. By mechanically tilting the sample within the goniometer range, the detector acquires the diffraction patterns simultaneously while the beam, the brown column in the figure, is fixed. The tilt step, and the acquisition time per frame, is determined by the sum of the exposure time, the blue cones on the right side, and the readout time, the yellow cones on the right side [10]. The resulting non-sampled wedge between two consecutive patterns is a result of the latter. The TEM goniometer stability poses a mechanical limit to the quality of the data acquisition, as the sample can drift laterally and out of the illuminated area during the data collection process. Hence, goniometer and specimen stability, rather than electron optics, can form a limitation for the reconstruction. The red area in Figure 2.14, is what we often refer to as the "missing wedge". The missing wedge is an angular region of missing information about the crystal, as a result of limited tilt range of the sample holder during data acquisition. Missing wedge-problems can sometimes be solved by collecting datasets from multiple randomly oriented crystals or merge smaller tilt range datasets to obtain higher data completeness [65].

Each dataset can be collected in only a few minutes*, meaning that the data collection is very fast. 3D ED requires a simple setup and a lower dose than TEM imaging techniques. The time consuming part of the technique is to extract the structural information from the actual tilt datasets. It is common practice to use established X-ray crystallography software

---

*from this project: tilt from angle -56.78 to 55.37 takes 562.48 s (9 min 22 s) and collects 424 frames
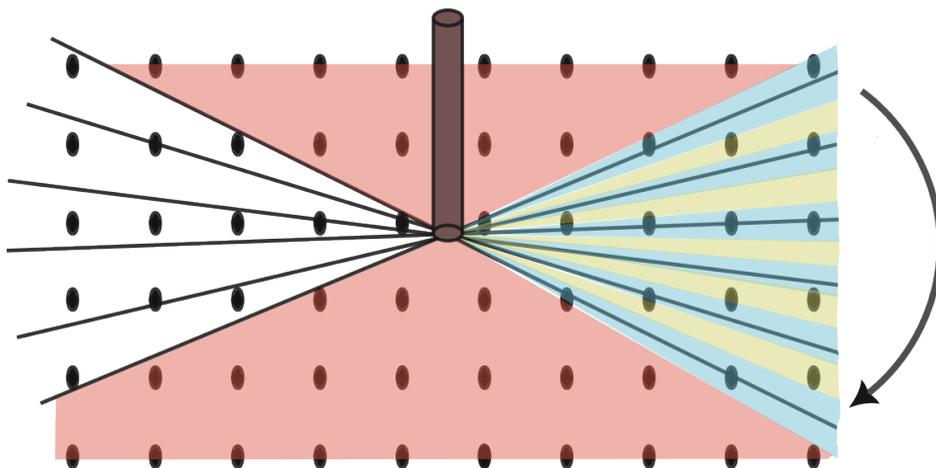
Figure 2.14: Schematic drawing of the continuous rotation method for 3D electron diffraction inspired by Figure 2D in [21]. The brown cone illustrates the near-parallel electron beam, which here is fixed, and the arrow indicates the direction of the continuous rotation. The area marked in red illustrates the missing wedge and the relrods illustrates the scatter positions. The blue cones illustrate the exposure time and the yellow cones illustrate the readout time.

also for 3D ED data for structure analysis, such as XDS [66], MOSFLM [25], or DIALS [24] for indexing, SHELXT for structure solution [46], and SHELXL for refinement [27], although specific 3D ED softwares have been developed as well, such as REDp, PETS, and ADT3D [67, 68, 14, 12]. As of now, is it still necessary to have a lot of knowledge about data collection on conventional TEM and crystallography in order to collect and process 3D ED data, although it might be beneficial to automate the process for non-experts to use 3D ED in structural analysis as well [9]. Commercial setups dedicated for 3D ED have gotten emerging attention and prevalence, for example Rigaku's XtaLab Synergy ED with ED-optimized X-ray software CrysAlis^{Pro} and WIT [18, 69], Eldico offer ED-1, a dedicated 3D ED instrumentation [17], Tescan offers Tensor which includes a 4D-STEM system with precession electron diffraction capabilities [70], and Nanomegas has the product DigiSTAR, which is a TEM ED tool for nanocrystal structure determination to perform PED[71].

3D ED has in very short time grown from research area to becoming a routine technique with different commercial suppliers of hardware and belonging software, for example Tescan, Rigaku and Eldico came with dedicated solution as expecting that the market will grow. Several open source solutions for 3D ED are also under constant development [72, 24]. However, despite that the principle of diffraction and reconstruction is conceptually established, the field is relatively young and especially on the data processing (structure solution and refinement as described in Section 2.3), developments are ongoing to fit specific data collection methods and specialize data processing for electron diffraction.

2.6 Zeolites

So far in this theory chapter, concepts, and basic techniques are reviewed, partly illustrated with simple examples (e.g. diamond and zincblende). The 3D ED technique fills a gap where volumes of single crystals are limited because large grown structures can be faulted (i.e. high spatial resolution is required, even beyond what SCXRD can give) and structures are so complex or modified that more dedicated structural analysis is needed to understand and optimize properties. 3D ED is additionally important for beam sensitive materials, where exposure to the electron beam should be limited and hence TEM lattice imaging and

CBED are not possible options. 3D ED-based structure solution as a field of research, is hence especially important for small proteins, complex minerals, and the material central in this work, zeolites.

Zeolites are microporous crystalline materials consisting of a dense network of AlO_4 and SiO_4 with pores in which different polar molecules can be located [73, 74]. The framework is built up by four oxygen atoms in a tetrahedron around the silicon or aluminum atoms, which encloses interconnected cavities [75]. The zeolite frameworks are three-dimensionally connected, but can also be two-dimensionally connected and is then typically called a zeotype material. To balance the negative charge of the tetracoordinate Al-atoms, positive ions are usually present in the pores, typically alkaline or alkaline earth metals such as Na or K. The structure of zeolites is directly correlated to its properties [29], thus the properties can be tailored for desired purpose [30]. Furthermore, these materials can not be studied with conventional SCXRD as they are difficult to grow into large single crystals [32]. Zeolites are widely used in catalysis, adsorption, separation, and ion exchange [32], for purposes such as waste water treatment, radioactive waste absorption, and CO_2-capture [76, 77, 78]. Some zeolites exist naturally, while others are produced synthetically. In this work, the naturally occurring zeolite called mordenite (MOR) is studied. MOR is widely used as a catalyst and a selective absorbent [79].

Zeolites have intricate 3D structures, indicating a need for a 3D analysis when solving the structure. The porous, but crystalline nature of zeolites is visible in Figure 2.15 of the example structure mordenite. MOR is orthorhombic, meaning all unit cell angles are 90-degrees, but each unit cell dimension is different. The porous nature of these materials makes them highly beam sensitive [31]. When being irradiated by dense beams, the microporous structure of zeolites may collapse. Furthermore, zeolittic structures are difficult to grow into larger volumes and therefore it is not possible to use SCXRD or PXRD to solve their structure [32]. Evidently, zeolites demand a low-dose 3D structural analysis that can perform well on sub-micron samples. Thus a method like 3D ED would be ideal for this purpose.
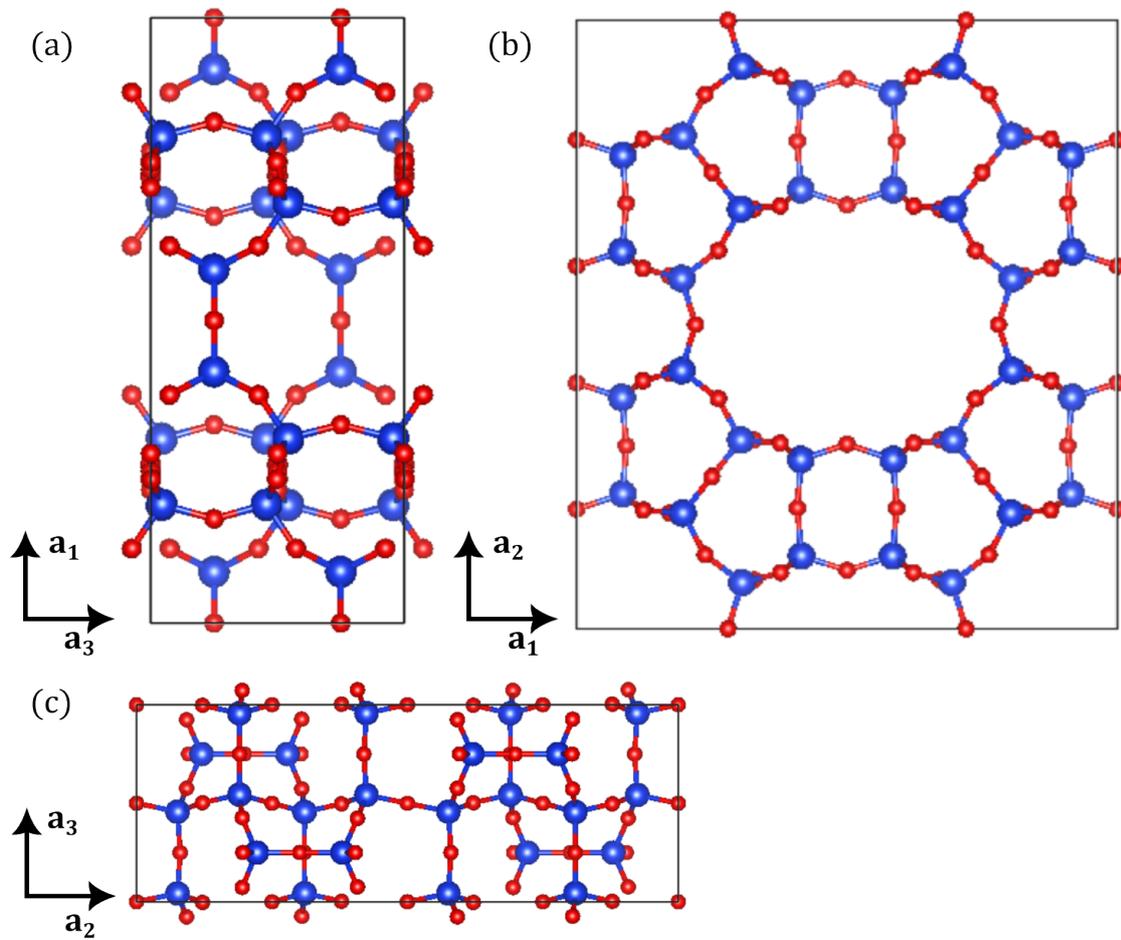
Figure 2.15: Structure of mordenite. Si-atoms are coloured blue, O-atoms coloured red. (a) shows the unit cell in [010] direction, (b) [001] and (c) [100]. Structure adapted from [65].

Chapter 3

Experimental

3.1 Materials and sample preparation

The materials studied are listed in Table 3.1 along with their crystal structural details according to literature. The different variations of MOR, i.e. Na-MOR, K-MOR, Cu-MOR-K, Cu-MOR-Na, and H-MOR-K are all samples that have been supplied by Sebastian Prodinger, Center for Materials Science and Nanotechnology (SMN), Department of Chemistry, University of Oslo. Some of these samples have been studied by X-ray diffraction in [80]. The samples of different variants of MOR were prepared by dissolving a small amount of powder in isopropanol and crushing it using a mortar. The solution was transferred to a glass container and placed in an ultrasonic bath. A drop of the solution was deposited onto a 200 mesh Cu TEM grid covered by 20 nm holey carbon support film and left to evaporate.

The Si and GaSb samples were made by scraping the surface of pure single crystal (111)Si- and GaSb-wafers with a diamond scriber of top radius 60 µm. The wafer pieces were grinded in a mortar before they were put into an organic solvent, ethanol, and placed for a few minutes in an ultrasonic bath to disperse. Then droplets of the solution were pipetted with a glass pipette onto a 100 µm thick grid. The grid consist of a single crystal silicon frame, with 50 nm thick silicon nitride (SiN) amorphous thin film on top. The grid has 9 windows of dimension 0.1 mm x 0.1 mm where the Si is back-etched away.

Pure MOR datasets were used as a reference material for this project and were obtained at Stockholm University by Cichocka et al. [2, 65]. These datasets are available for download from Zenodo*.

3.2 Microscope

Except for the MOR reference datasets [2], the TEM used for the 3D ED experiments is a Jeol JEM ARM200F, which is a double corrected coldFEG microscope. All data were obtained at room temperature with an acceleration voltage of 80 or 200 kV. The system has both probe and image spherical aberration correction. The detector used in this work is a Quantum Detectors 4R Merlin EM DED of 512x512 pixels with a width of 55 µm. Detector bit depths of 12 or 24 were used. The TEM was operated by Tina Bergh, Emil F. Christiansen, and Oskar Ryggetangen. The data were obtained over a span of five sessions and with varying data collection parameters listed in Table 3.3. There are four possible CAs, each with different sizes. These are numbered 1, 2, 3, and 4, depending on their size in µm which respectively is 150, 100, 70, and 50. SA aperture 1 was used, which is the largest

---

*https://zenodo.org/records/1321880

Table 3.1: List of materials used and their structural details according to literature.

Material	Chemical formula of unit cell	a, b, c [Å]	α, β, γ [°]	Space group
Si	Si_8	5.43 5.43 5.43 [81]	90, 90, 90	$Fd\bar{3}m$
GaSb	Ga_4Sb_4	6.14 6.14 6.14 [82]	90, 90, 90	$F4\bar{3}m$
Na-MOR	$Na_{6.64}O_{96}Si_{41.36}Al_{6.64}$	18.0689(5) 20.3498(5) 7.49318(18)[†]	90, 90, 90	Cmcm
K-MOR	$K_{5.43}O_{96}Si_{42.56}Al_{5.44}$	18.0358(5) 20.3746(5) 7.46260(15) [†]	90, 90, 90	Cmcm
K-MOR-H	$H_xK_xO_{96}Si_{48}$	18.1649(4) 20.3926(4) 7.49553(13) [†]	90, 90, 90	Cmcm
Cu-MOR-K	$Cu_xK_{5.43}O_{96}Si_{48}$	18.0358(5) 20.3746(5) 7.46260(15) [‡]	90, 90, 90	Cmcm
Cu-MOR-Na	$Cu_xNa_{6.64}O_{96}Si_{48}$	18.0689(5) 20.3498(5) 7.49318(18) [‡]	90, 90, 90	Cmcm
α-Al(Mn,Fe)Si	$Mn_2Si_{1.8}Al_9$	12.68 12.68 12.68 [83]	90, 90, 90	Pm3
MOR	$O_{96}Si_{48}$	18.11, 20.51, 7.53 [84]	90, 90, 90	Cmcm

one available with physical size 150 μm. For a selected crystal, the height was adjusted to the mechanical eucentric height, to minimize movement of the crystal during tilt. Two different tomography specimen holders were used. Gatan Model 916 High Tilt Tomography Holder, abbreviated G, has a tilt range of ±80. JEOL EM-21010 Single Tilt Holder with EM-21311HTR High Tilt Retainer, abbreviated J, has a tilt range of ±80. All datasets were acquired via automated routines with cRED Instamatic v 2.0.0 Python-based software [22]. Instamatic allows for automated collection routines of 3D ED data. Crystal tracking was done manually during continuous rotation by periodically defocusing the intermediate lens.

The reference datasets MOR1 and MOR2 are downloaded from Zenodo and do not have the setup described above [2, 65]. The experimental details for MOR1 and MOR2 are available in Table 3.2 and were obtained by Cichocka et al. at Stockholm University with a JEOL JEM-2100-LaB$_6$ equipped with a 512 × 512 Timepix hybrid pixel detector (55 × 55 μm pixel size, model QTPX-262k) from Amsterdam Scientific Instruments (ASI).

Additional datasets obtained are presented in Appendix A.1.

[†]Values from table S2 in Prodinger et al. [80].
[‡]Values for Cu-variant is assumed to be similar to the regular K- and Na-variants [80].
[††]per frame
[§]nominal camera length
[§]+ rotation from positive to negative angles, - from negative to positive

Table 3.2: Experimental details for MOR1 and MOR2.

	MOR 1	MOR 2
λ (Å)	0.02508 (200 kV)	0.02508 (200 kV)
Oscillation angle (°)	0.2314	0.2336
Tilt range (°)	-64.06 to 63.91 (127.97)	-43.90 to 58.65 (102.55)
Frames used	554	430
Defocused frames	55	43
Exposure time[††] (s)	0.5	0.5
Acquisition time[††] (s)	0.512	0.512
Total acquisition time (s)	283.0	224.7
Spot size	2	2
Camera length (mm)	250	250

Table 3.3: Overview of some of the studied 3D ED datasets with experimental parameters used.

	Na-MOR-15	Na-MOR-28	Na-MOR-31	K-MOR-31
Voltage [kV]	80	200	200	80
Camera length[§] [cm]	120	150	150	120
Bitdepth [bit]	12	12	12	16
Holder	G	G	G	J
Number of frames	118	101	89	123
Defocused frames	39	33	29	60
Exposure time[††] [s]	0.5	0.5	0.5	0.1
Acquisition time[††] [s]	0.801	0.805	0.806	0.518
Total acquisition time [s]	156.945	134.471	118.514	156.366
Image interval	5	5	5	10
Tilt speed	-	-	-	2x
Spot size	2	2	3	2
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range [°]	0.5003	0.4882	0.5276	0.3242
Flatfield applied [y/n]	n	n	n	y
Rotation range [°]	98.06	81.53	77.55	97.91
Rotation direction[§] [+/-]	+	+	-	+

3.3 Data processing

With Instamatic v. 2.0.0, collected data is automatically saved in the required formats for REDp and XDS [22]. Metadata is saved in a log-file. However, some parameters need calibration or refinement, such as camera length. These calibration calculations are described in Section 2.4.3. A Jupyter Notebook manual for what to do with data obtained without Instamatic, involving removal of defocused frames, and format conversion, is developed by Emil F. Christiansen[**].

The workflow of different software used for 3D ED data processing is shown schematically in Figure 3.1. The tools are tailored to determine specific properties of the sample. A more detailed manual for usage and installation of the programs is given in Appendix A.3. Distortion detection and correction are done when deemed relevant in `Python` before using the other programs.
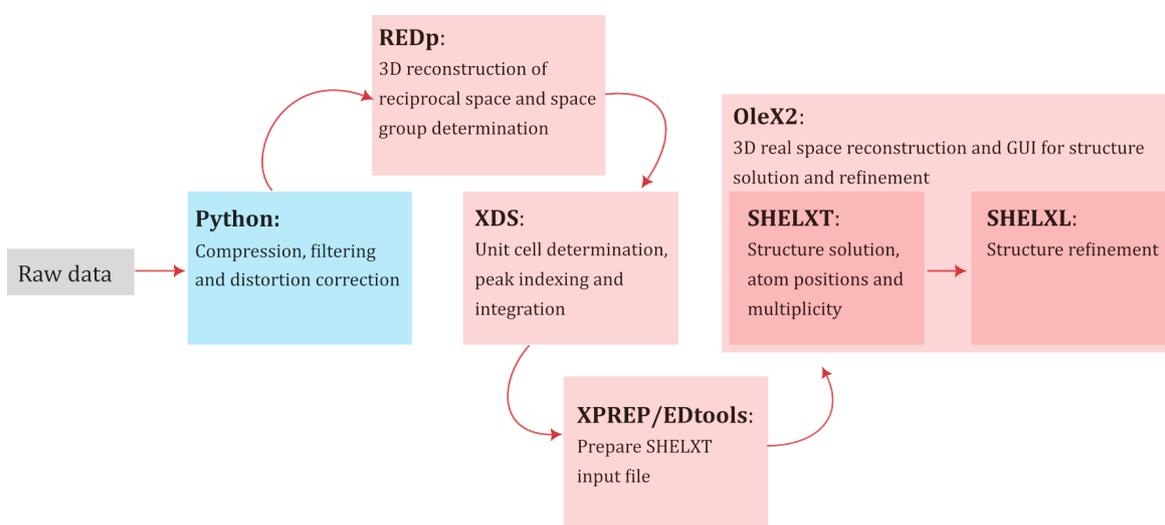
Figure 3.1: Overview of the workflow and different software used for data processing of 3D ED.

Rotation Electron Diffraction processing (software) (REDp) v1 is downloaded from [67], and is mainly used for reconstructing reciprocal space by mapping all detected two-dimensional diffraction spots in three dimensions for improved visualization. X-ray Diffraction Software (software) (XDS) version June 30, 2023 is used in this work, and is a tool for processing single-crystal diffraction data obtained by rotation methods [66]. XDSGUI is used for better visualization of the XDS process [85]. For further solution and refinement of the structure, the programs XPREP, EDtools, SHELXT, SHELXL, and Olex2 are used. XPREP version 6.10 [86] and EDtools v1.0.4 [87] is used for generating input files for SHELXT. SHELXT is used for structure solution, determination of atom positions, and multiplicity [46]. SHELXL is a refinement tool used to improve the initial structure solution found by SHELXT [27]. SHELXT and SHELXL can be run from Olex2. Olex2-1.5 is used for visualization of the determined crystal structure [88]. All programs, except XPREP which is sold by Bruker, are free for download and open source. The manual developed in this work include more info about installation of the programs, see Appendix A.3.

[**]https://github.com/TEM-Gemini-Centre/cRED/blob/main/cRED_conversion.ipynb

Chapter 4

Results

In previous research, REDp and XDS structure determination tools have been evaluated to determine best practice usage and find the ideal software parameters to use [1]. In this chapter, the data processing routine for structure solution and refinement will be presented, using the programs XPREP, EDtools, SHELXT, SHELXL, and Olex2. First, in 4.1, the specific parameters and function of each solution and refinement program will be reviewed. After the analysis of the complete data processing routine, it will be applied to 3D ED data from mordenite example datasets, MOR1 and MOR2. Additionally, the minimum requirements for datasets following this data processing routine will be addressed, as well as the effect of merging data. For clearness, results using the data collected at NTNU will be presented in the next chapter, integrated with the discussion on collection parameters affecting the structure analysis.

4.1 Structure solution procedure

The pre-study, presented in the previous project thesis [1], covers the use of REDp and XDS as tools for structure determination. This experiment focuses on further structure solution and assumes that quantities like lattice parameters and space group are known after the initial structure determination in REDp and XDS. The structure solution procedure consists of four main tools; XPREP, SHELXT, SHELXL, and Olex2. XPREP [86], or alternatively EDtools [87], prepares an instructional input file for SHELXT. SHELXT aims to determine the atom positions of the material, based on the principles described in Section 2.3. Through the algorithm called intrinsic phasing, SHELXT will suggest an approximate solution to a structural model [46]. This solution can be further refined to reach global minimum with SHELXL [27]. Both SHELXT and SHELXL can run in the terminal or in Olex2 [88]. Olex2 provides a user-friendly interface for using both tools, as well as visualizing the molecular structure in 3D. Additional tools for merging, XSCALE, and batch data processing, EDtools (again), are also investigated. The whole crystal structure analysis procedure for 3D ED is described in the manual in Appendix A.3.

The lattice parameters found in the XDS analysis are listed in Table 4.1 and were used for the structure solution and refinement analysis.

4.1.1 Convert files from XDS to SHELX

XDS produces several output files, among others a file called XDS_ASCII.HKL that contains information about the structure factors (intensities) belonging to specific reflections. This file will need to be converted to SHELX format in order for SHELXT to use it, which can

Table 4.1: Lattice parameters found by XDS for dataset MOR1 and MOR2 [1].

Material	Space group	a	b	c	α	β	γ
MOR1	Cmcm	18.691	20.518	7.710	90.0	90.0	90.0
MOR2	Cmcm	18.567	20.836	7.776	90.0	90.0	90.0

be done with XDSCONV, a tool embedded in XDS visble as a tab in XDSGUI. XDSCONV is by default set to convert XDS_ASCII.HKL to another format, called 'CCP4_I+F', hence it is necessary to change the text in the input XDSCONV file from 'CCP4_I+F' to 'SHELX' before running XDSCONV. The correct `.hkl`-file in SHELX format will contain five columns of the format $h \; k \; l \; F_o^2 \; \sigma(F_o^2)$.

An example input file for XDSCONV, XDSCONV.INP, is given Listing 4.1.

```
1  INPUT_FILE=XDS_ASCII.HKL
2  OUTPUT_FILE=MOR1.hkl SHELX
3  FRIEDEL'S_LAW=TRUE
```

Listing 4.1: Example input file XDSCONV.INP.

4.1.2 Preparing an .ins file

SHELXT needs two input files to work, a `.hkl`-file and a `.ins`-file. The first required input file, the `.hkl`-file, was made in the previous step, although the `.ins`-file is made either by using the licensed tool XPREP from Bruker [86] or the open source Python software EDtools [87]. Both tools were tested for `.ins`-file preparation during this work. The tools take information about the lattice parameter, space group, and tentative chemical composition as input.

Using the licensed tool XPREP

XPREP is terminal-based and can be started with the command `xprep NAME` where `NAME` is the name of the `.hkl` file. The program will then guide the user through a series of questions in order to specify the space group, merge reflections, and specify elemental composition. The choices made along the way are recorded in an output file called NAME.pcf along with the output `.ins` file for use in SHELXT. Some of the datasets used in this experiment were not able to use the "Determine space group" function in XPREP. In these cases, "Input space group" was used instead. XPREP needs to be run in the folder it is located, unless its file location is added to the system PATH.

XPREP is able to find higher-symmetry space groups, which can come in handy if the space group deduced from XDS/REDp is wrong. The suggestions for higher symmetries for MOR2 from XPREP is shown in Table 4.2, along with statistics describing the fit of each space group. The *lower* combined figure-of-merit (CFOM) is preferable and works as a compass for choosing the right space group for the dataset in question. This parameter takes all info from the other columns into account (these are not described further, but included in the table for completeness).

Using the open source Python library EDtools

EDtools is a Python library for handling electron diffraction data that were used as an alternative to the Bruker-licensed XPREP [87].

Table 4.2: Suggestions for higher symmetries of MOR2 dataset from XPREP.

Space Group	No.	Type	Axes	CSD	R(sym)	N(eq)	Syst. Abs.	CFOM
C222(1)	20	chiral	1	155	0.081	1617	0.0 / 3.6	10.78
C222	21	chiral	1	19	0.081	1617	0.0 / 3.6	15.14
Cmmm	65	centro	1	7	0.081	1617	0.0 / 3.6	15.42
Cmm2	35	non-cen	1	1	0.081	1617	0.0 / 3.6	60.14
Amm2	38	non-cen	4	1	0.081	1617	0.0 / 3.6	60.14
Amm2	38	non-cen	5	1	0.081	1617	0.0 / 3.6	60.14
Cmc2(1)	36	non-cen	1	102	0.081	1617	0.0 / 3.6	11.11
Cmcm	63	centro	1	100	0.081	1617	0.0 / 3.6	*3.91*
Ama2	40	non-cen	4	21	0.081	1617	0.0 / 3.6	14.69

EDtools requires installation of the program `sginfo`, that is available for download online*. This program needs to be build from scratch, which can be done through a "Developer Command Prompt"-terminal and the `clang` command. The full instructions are available in the README of the previously mentioned GitHub. The result is a fully built program called `sginfo.exe` that should be added to PATH. The following command makes the desired `.ins`-file:

`edtools.make_shelx -c (lattice parameters) -s (space group) -m (composition)`

For mordenite, the previous command would look like the following using tabulated literature values [84]:

`edtools.make_shelx -c 18.11 20.53 7.528 90 90 90 -s"63" -m O96 Si48`

However, none of the methods are complete in the sense that the .ins-file needs to be corrected in both cases. An example file is shown in Figure 4.1. The wavelength, marked with a yellow square in Figure 4.1, should be modified before applying SHELXL (but can be applied before using SHELXT also). These are by default set to X-ray values (for XPREP) and 200 kV electrons (for EDtools). For most datasets used later in this study, 80 kV electrons were used; thus, these values need to be updated.

Additionally, for XPREP, the input parameter SFAC would need to be corrected. SFAC should contain information about the electron scattering factors for each element present in the unit cell. The information about element-specific scattering factors can be found directly in SHELX-format online† [26]. The LATT-parameter in the INS-file, describes the lattice type of the material. Positive numbers are centrosymmetric, while negative numbers are non-centrosymmetric materials. The magnitude of the LATT-number describes the centering where P=1, I=2, 3 is rhombohedral obverse on hexagonal axes, F=4, A=5, B=6 and C=7. The SYMM-parameter is the symmetry operations, given as the coordinates of the general positions as given in the ITA [37].

---

*https://github.com/rwgk/sginfo
†https://srv.mbi.ucla.edu/faes/

```
TITL filename in Cmcm

        wavelength    lattice parameters
CELL 0.04176 18.11 20.53 7.528 90 90 90
number of formulas per unit cell
            standard deviation of unit cell
ZERR 1 0.0036 0.0041 0.0015 0 0 0
LATT 7  Lattice type, negative if non-centrosymmetric.
SYMM -X,-Y,0.5+Z
SYMM -X,+Y,0.5-Z      Symmetry operations
SYMM +X,-Y,-Z                            SFAC: element symbol and electron scattering factors
SFAC  O 0.143 0.305 0.51 2.268 0.937 8.262 0.392 25.665 0 0 0 0 0.63 15.999
SFAC  Si 0.363 0.428 0.974 3.557 2.721 19.39 1.766 64.333 0 0 0 0 1.16 28.086
UNIT 96 48  Number of atoms of each element in unit cell
TREF        Direct method used
HKLF 4      Standard SHELX reflection data file used
END
```

Figure 4.1: Example of an .ins-file for use in SHELXT. Different parameters are explained separately, whereas the parameters in yellow boxes need to be modified after XPREP/EDtools preparation.

4.1.3 Structure solution with SHELXT

When both the `.ins` and `.hkl` files are complete, the SHELXT structure solution program can be applied and refined with SHELXL. There are two ways of running SHELXT and SHELXL:

1. The terminal way: Put `shelxt.exe` and `shelxl.exe` in the same folder as your `.ins` and `.hkl` file and run `shelxt NAME`, where `NAME` is the filename for both `.ins` and the `.hkl` file. SHELXT saves the result to a `.res` file that can be further used by SHELXL for refinement. Change the `.res` file to a `.ins`-fileformat and run `shelxl NAME`.

2. The Olex2 way: Copy `shelxt.exe` and `shelxl.exe` to the Olex2 program folder. SHELXT will now show up in the Olex2 GUI under Work → Solve and click the arrow pointing down, where all the different structure solution tools are available. Similarly, SHELXL will show up in the Olex2 GUI under Work → Refine and click the arrow pointing down, where all the different structure refinement tools are available.

The Olex2 way was applied in this experiment as this provides a more intuitive GUI and visualization of the solved structure. SHELXT is run by clicking "Solve" or typing `solve` in the integrated Olex2-terminal.

SHELXT will run until it converges to an initial solution, see Section 2.3. If it crashes, something is probably wrong with the `.ins`-file or with the `.hkl`-file from XDS. If there are multiple solutions, these will all be listed from lowest to highest R-value. Lower R-value indicates a better fit, as described in Section 2.3. However, if one suspects a higher R-value space group fits better one can freely choose from the list of solutions. If the solution is wrong on the other hand, one can force a solution in the right space group by typing `-s"SPACEGROUP"` in the command line panel under "Solution Settings Extra", where `SPACEGROUP` is the name of the space group (for example Cmcm). The process SHELXT goes through to determine the initial model and its findings are reported in an output `.lxt`-file.

4.1.4 Refinement with SHELXL

After the initial structure solution, refinement will be needed in order to improve the fit of the initial model. Refinement in Olex2 is done by clicking "Refine", `ctrl + R` or typing `refine`. In the following paragraphs, the different methods for validating, modifying, and limiting the calculated model in order to fit the experimental data will be described.

Inspecting the structure

The atoms predicted by the initial model will be visualized in 3D in Olex2 after structure solution, see Figure 4.2(a). The asymmetric unit of the structure will be shown by default, meaning that the number of atoms shown depend on the number of symmetry elements and constituents in the unit cell. The structure can be extended, as is done in Figure 4.2(b), which can be preferential to visualize the periodicity of the crystal structure. This is done with the command `grow` or alternatively `grow -w` or `grow -s`. To display the unit cell, use the 'View' tab and click the square box. To fill the unit cell with atoms, type `pack cell` in the terminal window, see Figure 4.2(c). All subfigures in Figure 4.2 are viewed in plane (001).
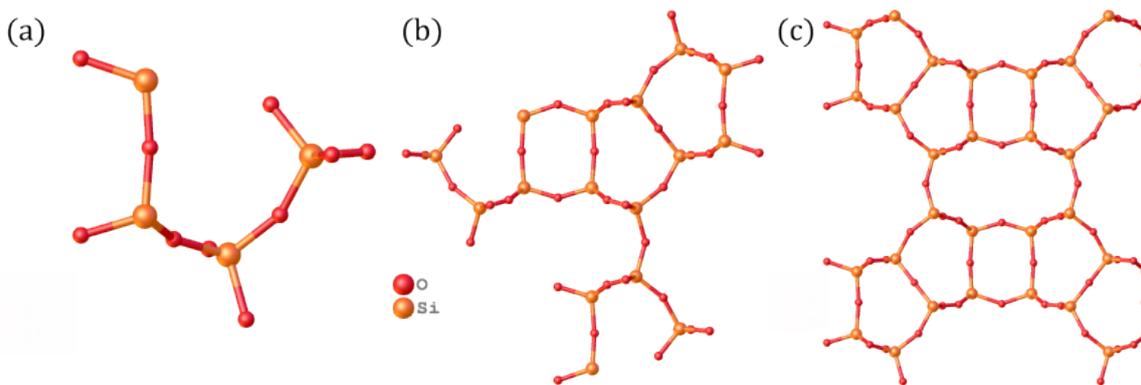
Figure 4.2: Olex2-view in plane (001) of the (a) asymmetric unit, (b) the grown structure and (c) the packed unit cell structure of mordenite MOR1.

Statistics

To express quantitatively goodness of model, certain criteria are defined based on the figures of merit described in Section 2.3.1. The softwares can have specific implementations and therefore they are given here rather than in general theory or overview given in methods. Olex2 calculates and displays different R-factors calculated by SHELXT and SHELXL, similar to the general one defined in Equation 2.29. These are important for measuring different aspects of the input data. In Olex2, R_{int} is defined as follows,

$$R_{int} = \frac{\sum |F_o^2 - F_o^2(mean)|}{\sum F_o^2} \tag{4.1}$$

while R_1 and wR_2 are defined as

$$R_1 = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \tag{4.2}$$

$$wR_2 = \sqrt{\frac{\sum w(F_o^2 - F_c^2)^2}{\sum (wF_o^2)^2}} \tag{4.3}$$

43

where w is the weight, F_o is the observed structure factors, and F_c is the calculated structure factors. Another quality parameter to observe is the Goodness of Fit,

$$GooF = S = \sqrt{\frac{\sum w(F_o^2 - F_c^2)^2}{n - p}} \tag{4.4}$$

where n is the number of reflections and p is the total number of parameters. Preferably should R_{int}, R_1 and wR_2 all be low and $GooF$ close to 1. The characteristics are colorcoded to be easier to interpret in Olex2. However, it is expected that the R-values will be higher for electron diffraction (ED) than for XRD data, for example a R_1 of 10-30% is sufficiently good for ED [65].

As mentioned in Section 2.3, refinement will require as many variables as equations. However, in practice a higher number of equations (intensities) than variables are required to assure the intensities are of good quality for further obtaining a good solution. The ratio between the number of variables to the number of intensities is presented as a battery in the upper right corner of Olex2 and is color coded to display whether the ratio is good or bad.

The refinement produces 'Q-peaks' that are suggestions for additional atoms, see Figure 4.3(a). The number of additional Q-peaks can be adjusted in the refinement menu under 'Peaks'. The Q-peaks are visible in the 3D view of the structure as brown spheres. By clicking `Ctrl + Q` the Q-peaks will first form bonds, and an additional click will remove the Q-peaks from view.

Editing the model

Usually, crystal structure refinement requires detailed chemical knowledge of the structure in question. If the initially visualized molecule seems wrong, for example if any of the bonds or atoms have prohibited or unphysical properties (such as too many bonds for an atom of a certain element), then the bonds may be removed or the atom exchanged to another. Removal of bond or atom can be done by clicking it (the atom will turn green), and then type `kill`. An atom can be switched to another element by right clicking it and choose "Type". This molecular 'editing' will be logged automatically in the `.ins`-file for further refinement, and is essential to do in order to obtain a correct model. Refine between each edit and observe if the R-values change. The R-values can be used as a guiding compass to determine whether the edit was right or not, but should not be blindly trusted. A drop in the R-value indicates a right move, because it indicates that the experimental data fits the calculated model better.

Restraints and constraints

Other features to inspect with the structural model are the angles between atoms and lengths of bonds. By clicking 'Report' under 'Work', one can see a full overview of all the angles and bond lengths of the model. These values should be inspected and be compared to tabulated values from literature. If there are larger deviations from the values in the literature, consider including restraints or constraints. Restraints are suggestions for the model to adjust parameters. Constraints are strict instructions that limit the model. These are included in the same way as other Olex2 or SHELXL commands, typed into the terminal of Olex2. Common for structures with Si and O, for example is to restrain the Si-O bond length to be about 1.6 Å and the O-Si-O angle to be around 109°.

To restrain angles, choose the three atoms concerned and use the `TRIA` command followed by the literature angle. This command is related to `DANG` and `DFIX`. To restrain bond length,

choose the two atoms concerned and use `DFIX` followed by the length found in literature. Other parameters that might be of interest are `RIGU` [89], `DELU`, and `SIMU`. These were not used in this experiment and will hence not be described any further. Hence, from an initial model, the structure can be refined within certain boundaries.

Electrostatic potential map

Atom positions are located where the electrostatic potential map has its maxima, as explained in Section 2.1.4. Hence, a good way to measure the fit of a structural model is to compare the model to the three-dimensional electrostatic potential map, which was briefly mentioned in Section 2.3. The map is visible by typing `Ctrl+M`, see Figure 4.3(b). In the default view of the electrostatic potential map in Olex2, which is a 3D wire model, the maxima of the model will be visible as green blobs and the minima will be coloured red. The default map type is the diff (difference, $F_o - F_c$) map. This map can be useful for discovering atoms that are not present in the calculated model yet. If the position of a suggested Q-peak correspond to an electrostatic potential map maximum as shown in Figure 4.3, place an atom here by changing the atom type of the Q-peak by right-clicking it. When refining after this move, the electrostatic potential map should change to fit the model better.
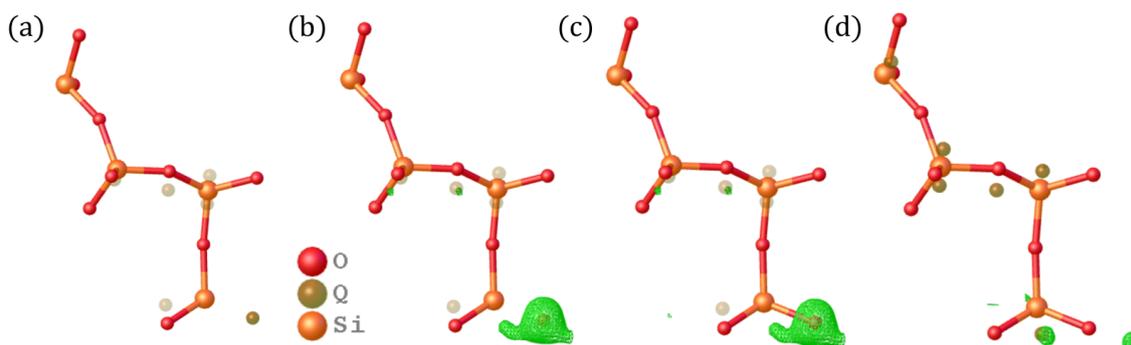
Figure 4.3: (a) Molecule with visible Q-peaks. (b) Molecule with electrostatic potential map and Q-peaks visualized simultaneously. (c) Atom type 'O' assigned to Q-peak. (d) Electrostatic potential map and Q-peak positions changed after new refinement.

There are different types of electrostatic potential maps in Olex2 and different ways to view it. Another common map to examine is the 2D contour map, visible in Figure 4.4. Here, three types of maps are viewed. Figure 4.4(a) shows the observed electrostatic potential, F_o. Figure 4.4(b) shows the calculated electrostatic potential, F_c, of the current refined model and Figure 4.4 (c) shows the electrostatic potential difference, , $F_o - F_c$. All three contour maps are collected from the same plane, i.e. $0.348\mathbf{a}_1$ $0.724\mathbf{a}_2$ $0.604\mathbf{a}_3$.

The level of detail in the electrostatic potential map can be adjusted under 'Map Settings → Level /eÅ^3'. The level, also commonly known as the σ-level, is the number of standard deviations above the mean electrostatic potential value [90]. It is common practice to study difference maps at the 3σ-level, 4σ-level, or 5σ-level. Lower levels than this should be avoided in order to not draw conclusions from experimental noise. Higher levels will show only the most prominent features.
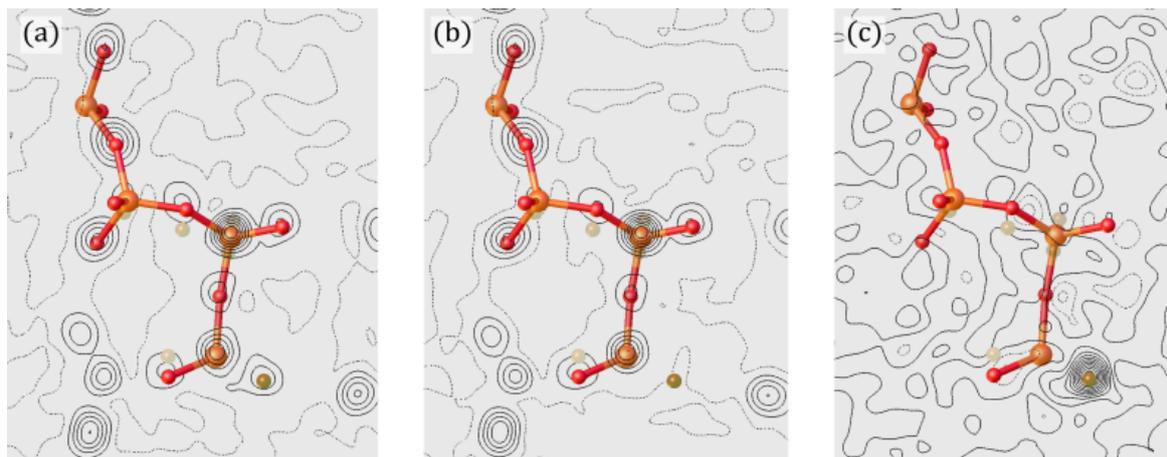
Figure 4.4: Contour map of (a) F_o, (b) F_c and (c) $F_o - F_c$ (difference map).

Atomic displacement parameters

During refinement, a common parameter to calculate is the atomic displacement parameter (ADP) for each atomic position xyz as described in Section 2.1.4. ADPs are a measure of the mean-square displacement from the equilibrium atom position in the crystal. The ADPs can also be referred to as anisotropic displacement parameters, usually denoted by U_{ij}, which describe the extent of an ellipsoid (meaning the atom vibrates differently in different directions). However, if atoms are assumed isotropic then the ADPs might be referred to as isotropic displacement parameters, U_{iso}. The difference between an isotropically and anisotopically refined model is visible in Figure 4.5. Figure 4.5(b) shows each atom displayed as an ellipsoid. Oxygen atoms (in red) have visibly more elongated ellipsoids than the silicon atoms (in orange), indicating larger ADPs, i.e. more vibration in certain directions.
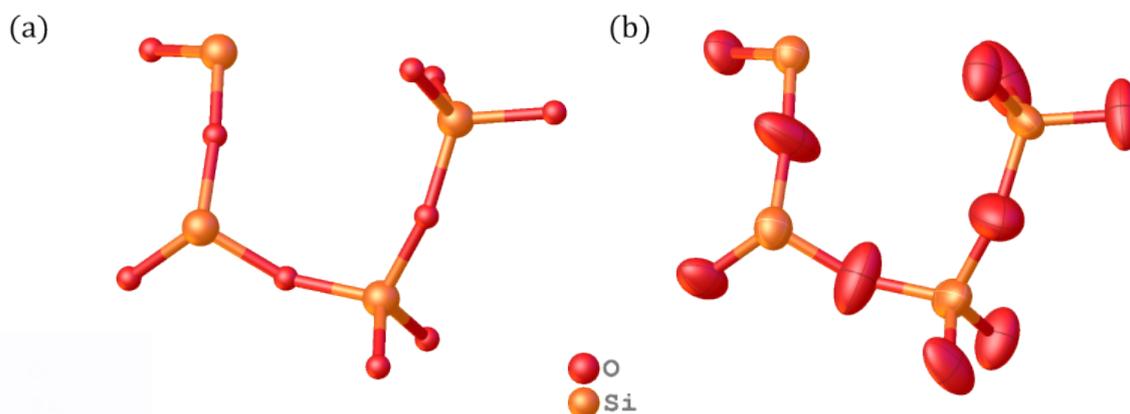
Figure 4.5: Olex2-view of asymmetric unit in mordenite. Refinement of the structure can either be done under isotropic assumptions (a) or anisotropic assumptions (b).

The resulting solved structures of MOR1 and MOR2 from this structure solution and refinement process are presented in Figure 4.6, with corresponding refinement characteristics in Table 4.3, 4.4, and 4.5. In these tables "space group" is shortened to sg. Figure 4.6(a)

shows the full grown MOR structure in [001] direction common for both MOR1 and MOR2, however MOR1 and MOR2 were found to have differently defined origins of the unit cells for the same space group, see Figure 4.6(b)(I) and (c)(I). Hence, their asymmetric units were also different, see 4.6(b)(II) for MOR1 and (c)(II) for MOR2. Both unit cells fully present the 3D MOR-structure by having the same symmetry elements, however, it is more common when describing porous structures to use unit cells that includes the whole pore in the unit cell, like for MOR2.

Table 4.3: XDS statistics for datasets MOR1 and MOR2.

|	Completeness	CC1/2	Number of reflections	Independent reflections	2Θ range
MOR1	99.2%	98.3%	6772	1663	0.158 to 1.842
MOR2	93.6%	98.7%	5510	1589	0.234 to 3.062

Table 4.4: Structure solution statistics for datasets MOR1 and MOR2.

	XPREP sg[‡]	Input sg	Solved sg[§]	Forced solution	R_{int}
MOR1	Cmca	Cmcm	Amm2	yes	17.81%
MOR2	Cmcm	Cmcm	Cmcm	no	8.76%

Table 4.5: Refinement statistics for datasets MOR1 and MOR2.

	Constraints	Parameters	GooF	R_1[*]	wR_2[*]
MOR1	4	94	1.669	0.2592	0.5641
MOR2	0	96	1.610	0.1523	0.4263

MOR1 solved the structure in a different space group, i.e. Amm2, but were forced solved in Cmcm using SHELXT command `-s"Cmcm"`. The last few frames of MOR2 were excluded after inspecting the SCALE plot in XDS, because the intensities were altered due to copper grid shadowing. Furthermore, MOR1 needed more restraints than MOR2 and had higher R-values than MOR2. However, the goodness of fit and number of refinement parameters were about the same. The restraints used were `TRIA` and `DFIX`. These statistics correspond well with results from previous papers, except fewer restraints were used to refine MOR1 in the current work (Cichocka et. al [65], used `RIGU` restraints on all framework atoms).

Refinement is a crucial part of the 3D ED data processing procedure, as it validates the absolute structure of the crystal in question. The refinement process is iterative and may require detailed chemical knowledge of the material in question. This section described some of the most important features of refinement. Further on, some additional tools for enhanced data processing will be presented, i.e. XSCALE for merging data and EDtools for batch processing.

[‡]With lowest CFOM

[§]Initial solve

[*][I $\geq 2\sigma$ (I)]

(a)

(c) MOR2

(I)

(II)
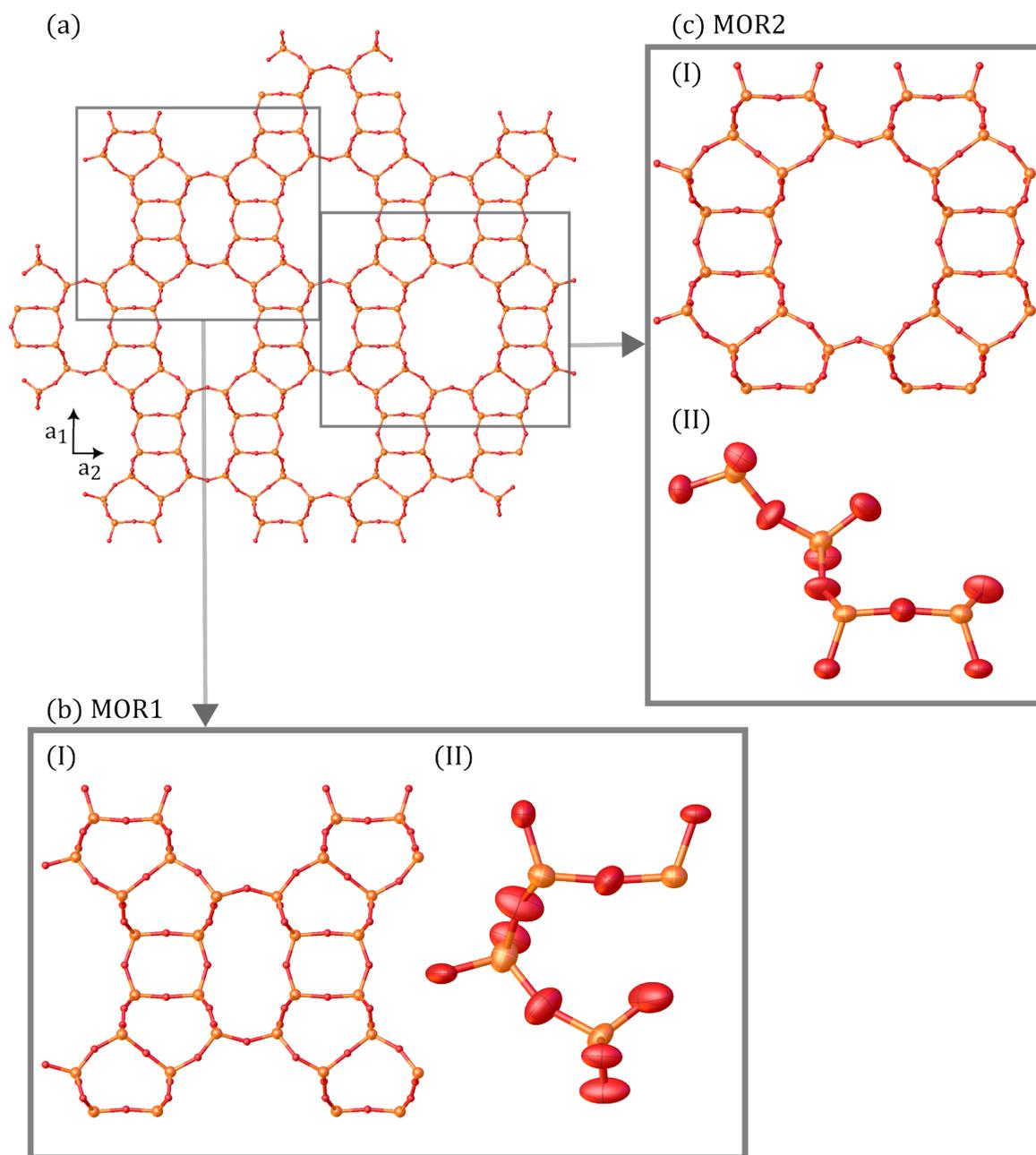
(b) MOR1

(I)   (II)

Figure 4.6: (a) Fully grown MOR structure common for MOR1 and MOR2. (b) and (c) show structures from MOR1 and MOR2 respectively. Figure (I) shows the unit cell and Figure (II) the asymmetric unit solved anisotropically. All figures shown in [001] direction.

4.1.5 Merging datasets

If the datasets obtained have a low rotation range, few reflections, or for any other reason low data completeness, using XSCALE can be a useful tool. XSCALE merges output `.hkl` files from XDS and scales the intensities to be weighted equally.

In XDSGUI, XSCALE is visible as one of the last top tabs. XSCALE takes an input file XSCALE.INP, that is automatically generated by XDS, but can also be made manually. An example XSCALE.INP is shown in Listing 4.2.

```
1 SPACE_GROUP_NUMBER= 63            !Shared space group
2 UNIT_CELL_CONSTANTS= 18.65 20.69 7.74 90.0 90.0 90.0  !Average unit cell
    constants
3
4 OUTPUT_FILE=merged.ahkl
5
6 ! List of all inputfiles to be merged
7 INPUT_FILE=../mordenite_cRED_1/SMV/XDS_ASCII.HKL
8 INPUT_FILE=../mordenite_cRED_2/SMV/XDS_ASCII.HKL
```

Listing 4.2: Example input file XSCALE.INP.

From XDS, XSCALE uses the output-file XDS_ASCII.HKL as input. Insert the path to the files you want merged by using the parameter "INPUT_FILE". The files can alternatively be copied into one common folder, then only the name of the file would be necessary. The average of the lattice parameters found in the files to be merged should be used as input for "UNIT_CELL_CONSTANTS". All files to be merged should be of the same material. Furthermore, all the files should have been run through XDS with the same input space group separately before executing this step. Data collection parameters can vary, i.e. do not have to be equal for all input files.

The resulting file is of the same format as XDS_ASCII.HKL, and would hence need to be converted to SHELX format with XDSCONV for further processing.

4.1.6 Automation and batch data processing

For processing multiple datasets simultaneously and optionally choose which to merge, EDtools can be used again [87]. However, for batch processing, EDtools needs to have access to XDS, as it uses this program to index reflections. Hence, one needs to have XDS installed on the same computer as EDtools runs on. EDtools contain multiple options for batch processing. It is in this experiment used for autoindexing of multiple datasets and information extractions from batch indexing.

EDtools simply runs XDS for multiple datasets at once, by searching 'XDS.INP' files in all sub-directories. This procedure is carried out by the following command,

    edtools.autoindex

All projects should hence be copied to one common folder and the Python-terminal should be navigated to the path of the common folder. The results of the autoindexing can be summarized by using,

    edtools.extract_xds_info

where the content of the output files from XDS, CORRECT.LP, are found and read in all sub-directories. Only the most important information from these files are extracted. Hence, this command can also be useful for analyzing smaller batches of data as well because navigating in the output files from XDS can be tedious work. Statistics about found lattice parameters and space group, in addition to data characteristics like number of reflections, resolution, completeness, and a few more are written to the terminal output and an excel file. The data with sufficient completeness ($>10\%$) and CC1/2 ($>90\%$) will be saved to a text file.

In this section, different parts of the data processing routine for structure analysis of 3D ED data have been reviewed, specifically for the structure solution and refinement packages in the workflow. The following section will investigate the effect of using the procedure described above, by studying the effect of merging, minimum dataset requirements for usage of the processing routine, and the effect of different input space group.

4.2 Minimum dataset requirements for structure solution and refinement

Now that the full data processing workflow is known, necessary dataset properties and requirements will be investigated. All data obtained at NTNU were observed to not work ideally in this workflow at some point in the process, which will be discussed further in Chapter 5. Hence, a more thorough analysis were done on the dataset requirements to get well-functioning 3D ED data. The amount of excluded frames is first investigated by analysing the *minimum number of frames* needed for reconstruction, the *maximum density of excluded frames*, and *maximal gap between frames* possible for the mordenite datasets MOR1 and MOR2. Further, the effect of *merging datasets* with XSCALE and *different input space groups* in SHELXT was analyzed.

4.2.1 Exclusion of frames

To identify errors in data acquisition or data structure of the data obtained at NTNU, different characteristics of the NTNU-data were compared to the reference MOR-data. One of the observed differences was the ratio between the number of frames and the number of excluded defocus images. Frames can additionally be excluded e.g. if the grid edge shadow appears, as this affects the intensities in the frame. The ratio varied for the NTNU-data, but the fraction of excluded frames usually ranged between 0.5 and 0.15. The MOR-data only removed one in every ten frames (excluded fraction 0.1). Exclusion fraction thus refers to the fraction of the total number of frames that are excluded.

The two MOR-datasets, MOR1 and MOR2, were hence cropped to different sizes, with exclusion fractions ranging from 0.1 to 0.8. Exclusion fraction of 0.1 is the default for these datasets. This was done by using the EXCLUDE_FRAMES parameter in XDS, where an increasing interval of frames were removed. All datasets were run with EDtools batch processing. The resulting completeness, CC1/2 (as described in Section 2.3.1), and number of indexed reflections are plotted in Figure 4.7(a). The critical completeness, of 10%, and critical CC1/2, of 90%, is visible as dotted straight lines. The dataset should be above critical level for both figures of merit. By curvefitting the mean of the two mordenite completeness functions, the critical exclusion fraction becomes 0.76. When observing the CC1/2 plot, the acceptable exclusion fraction is observably lower. Curvefitting the CC1/2 plot, gives a critical exclusion fraction of 0.54. Hence, the highest acceptable fraction of excluded frames is 0.54. The number of reflections decrease with the higher fraction of excluded frames.

Note that dataset MOR1 has a longer rotation range and a higher number of total frames, which can be seen from the experimental table in Appendix A.1.

By further cropping the MOR1 and MOR2 dataset, one can observe what the critical lower amount of frames is. The completeness, CC1/2, and number of indexed reflections is again plotted against number of frames in Figure 4.7(b). The minimum number of frames is found by curvefitting the mean of the two mordenite datasets. For a critical completeness, i.e. over 10%, this analysis shows one would need minimum 21 frames. This amount of
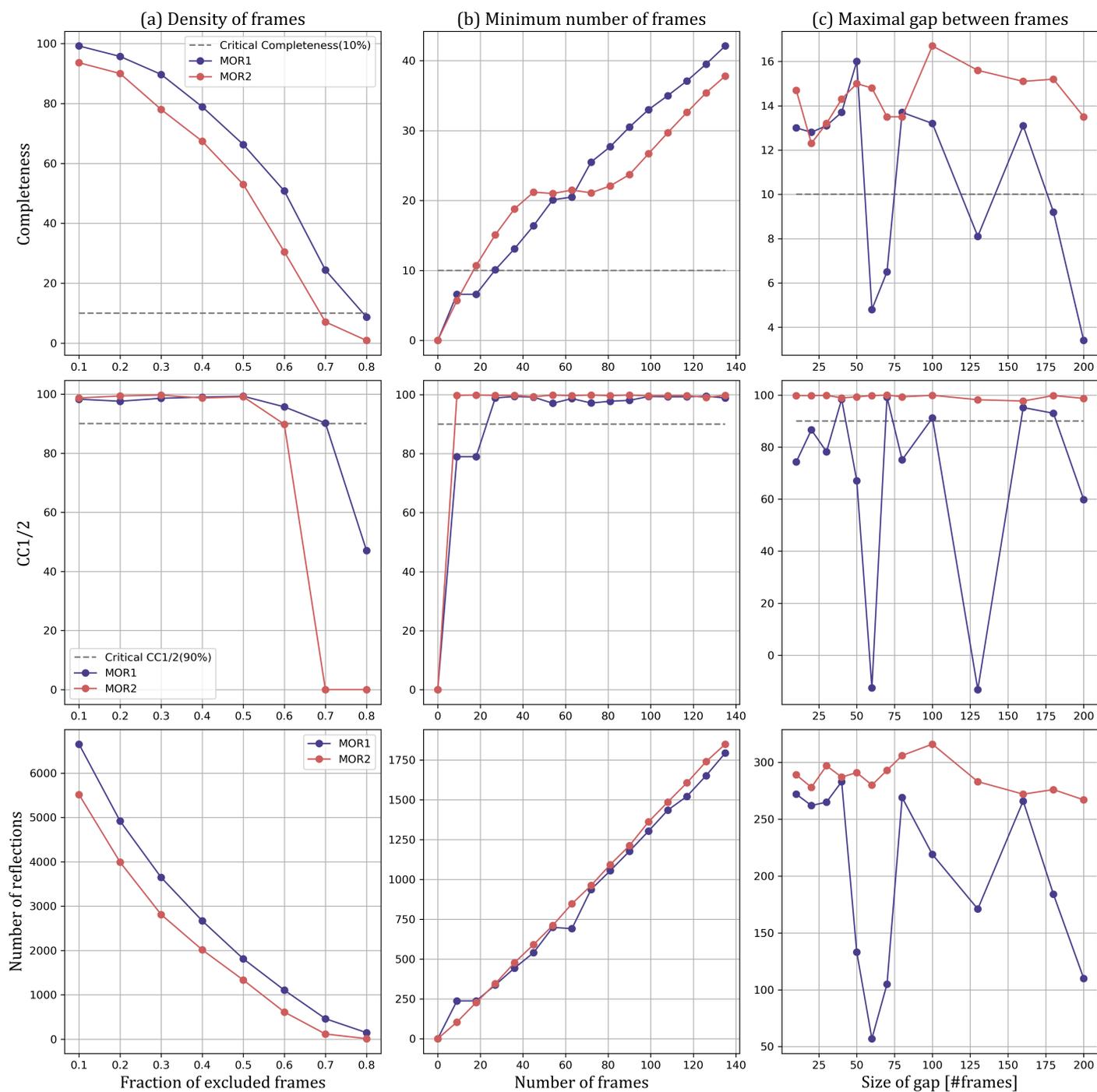
Figure 4.7: XDS characteristics on y-axis, i.e. completeness, CC1/2, and number of reflections, plotted as a function of (a) density of excluded frames, (b) number of frames, and (c) size of gap between each set of frames.

frames corresponds to a rotation range of 4.87°, using the mean oscillation range, 0.2325°, of the oscillation ranges of MOR1, 0.2314°, and MOR2, 0.2336°.

Now, both the critical lower number of frames and highest fraction of excluded frames is known. The next property to be investigated is the maximal gap between frames. This was done by keeping the number of frames constant, i.e. 27 frames which is above the critical number of frames found previously, and increasing the size of the gap between groups of 9 frames by gradually excluding a larger amount of frames. XDS characteristics, i.e. completeness, CC1/2, and number of reflections, of an increasing gap between frames for the two MOR-datasets are plotted in Figure 4.7(c). Note that MOR2 behave much more smooth than MOR1 in this study, as the investigation was performed at a critically low amount of frames.

4.2.2 Effect of merging

XSCALE was used for merging datasets together. To illustrate the effect of this tool, MOR1 and MOR2 were merged. Specific resulting characteristics is presented in Table 4.6, i.e. the number of reflections, overall completeness, and different R-values. R_o and R_e are gathered from XDS, while R_1 and wR_2 are from refinement (SHELXL).

Table 4.6: Dataset statistics for MOR1 and MOR2 separately and merged.

Dataset	Number of Reflections	Overall Completeness	R_o	R_e	R_1	wR_2
MOR1	6774	99.2%	29.5%	33.7%	35.25%	62.94%
MOR2	5512	93.6%	12.0%	11.9%	21.61%	48.11%
Merged	12277	99.8%	21.9%	24.3%	26.58%	49.66%

From Table 4.6 one can see that the completeness of the merged dataset is higher than both datasets separately. The four R-values all show a similar trend for the merged R-value to be in between the lowest and highest R-value of the separate datasets. However, wR_2 for the merged dataset is only 1.55% higher than the lowest separate wR_2 value. Note that some reflections are rejected in XSCALE so that the total number of reflections after merging do not equal the sum of the number of reflections of the separate datasets.

4.2.3 Effect of input space group

As described in Section 2.1.1, there are often multiple ways of describing a crystal structure, depending on preference of unit cell. During data processing, discrepancies between the expected space group and the one found by SHELXT during crystal structure solution were observed, hence it would be useful to know how the program reacts to different inputs and whether one can use the R-values to determine the right structure.

EDtools were used to produce input files (`.ins`-files) with different space groups, in order to study if there are possibilities for solving structures in other crystal systems or if SHELXT finds the solution either way. Eight arbitrary space groups were chosen with different symmetry elements present. Many of these were also present in the higher symmetry table from XPREP, see Table 4.2. Higher symmetry space groups were observed to not fit well. SHELXT would then produce error messages such as "Inconsistent cell and Laue group". The space group found by SHELXT, the corresponding R_{int}, and R_1 after one initial refinement for each input space group for MOR1 and MOR2 are presented in Table 4.7. MOR1 has observably higher R-values. MOR2 is solved in the correct space group (Cmcm (no. 63) according to [84]) for half of the input space groups used in this test, i.e.

Table 4.7: Table of input space groups, solved space group, R_{int}, and R_1 after initial refinement for MOR1 and MOR2.

	MOR2			MOR1		
Input space group	Solved space group	R_{int}	R_1	Solved space group	R_{int}	R_1
2	2	6.18%	22.26%	1	13.85%	28.78%
15	15	6.87%	20.52%	9	16.76%	30.65%
16	60	8.24%	Refinement unstable	28	17.08%	27.43%
20	63	8.24%	21.64%	38	17.08%	30.80%
36	63	8.36%	21.36%	38	17.97%	30.76%
52	60	9.10%	Refinement unstable	28	17.45%	32.96%
64	63	9%	21.64%	38	16.34%	30.72%
66	63	8.87%	21.64%	38	17.83%	30.85%

no. 20, 36, 64, and 66. MOR1 was for the same set of space groups solved to have a symmetry corresponding to space group no. 38 (Amm2). The R_{int} values are seemingly smaller for the lower space groups, and may hence be deemed insufficient to judge whether the space group is correct or not.

Evidently, when different input space groups are used in structure solution, many groups of similar structure were observed, see Figure 4.8. Here, the same *hkl*-reflections are solved with different input space group in SHELXT and initial refinement in SHELXL, many found by SHELXT visible in Table 4.7, i.e. (a) and (d) Cmcm (no. 63), (b) and (e) Pbcn (no. 60), (c) and (f) Amm2 (no. 38), (d) and (g) C2/c (no. 15). Note that for Amm2, the axes are flipped to $\mathbf{a_1}$'$=\mathbf{a_3}$, $\mathbf{a_2}$'$=\mathbf{a_1}$, $\mathbf{a_3}$'$=\mathbf{a_2}$. First row shows the asymmetric unit viewed in (011), which for (c) corresponds to (101), and the second row shows the packed cell viewed in (001), which for (f) corresponds to (100). The structures look very different when observing the asymmetric unit, although the packed cells are almost identical. This shows that the structure can indeed be described by different crystal systems than the space group from literature. However, all other space groups (than Cmcm) shown here have lower symmetry. Cmcm is then the highest symmetric space group the structure can be solved in, but not the only space group it can be solved in.
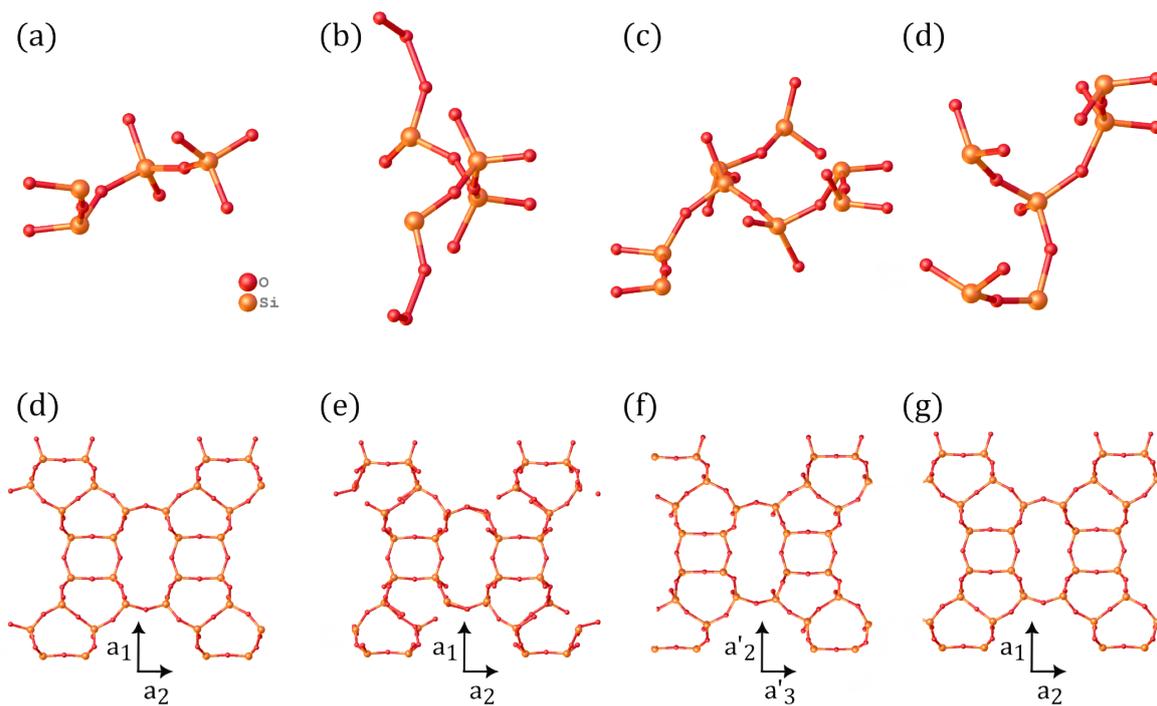
Figure 4.8: Different input space group in ins-file for structure solution. Experimental data solved in four separate space groups; Cmcm, Pbcn, Amm2, and C2/c. First row shows the asymmetric unit for space group (a) Cmcm (no. 63) viewed in (011) (b) Pbcn (no. 60) viewed in (011) (c) Amm2 (no. 38) viewed in (101) (d) C2/c (no. 15) viewed in (011). Below, the corresponding packed cell follows for (d) Cmcm (e) Pbcn (f) Amm2 (g) C2/c.

Chapter 5

Discussion

This section will discuss how the data collection parameters and routines can affect the data performance in the data processing protocol. For example, problems such as beam damage and saturation of pixels were observed. Most of these analyzes will be done on the data obtained at NTNU, which were not covered in Chapter 4. Suggestions for measures to take to correct for these issues will be presented and evaluated. How well suited the experimental setup is for 3D ED data collection will additionally be discussed, e.g. hardware performance and choice of test material. Further difficulties with the data processing protocol described in Chapter 4 will be assessed to map out pain-points and bottlenecks of the method. Potential work-arounds and improvements for the current method are evaluated, along with further progress to dynamical analyses and other workflows.

5.1 Sample material and preparation

In this section, the test material performance for 3D ED usage for the obtained NTNU datasets will be explored. A brief assessment of the desired TEM specimen grids will follow.

5.1.1 Evaluation of test material performance

In previous work [1], other materials were used as known model crystals to test and set up the method, such as Si and GaSb. The initial idea of using Si and GaSb as test materials was to test high-symmetry materials of low beam sensitivity. In this way, it would be possible to illuminate the materials for longer time intervals, i.e. a longer tilt range, while avoiding devastating beam damage. High-symmetry materials scatter radiation well, causing strong, but few reflections in the diffraction patterns which were expected to require smaller tilt ranges for completeness [6]. This last aspect is visible in the amount of reflections saved to the .hkl file in XDS. Almost all reflections detected in XDS originate from the same group of planes. Hence, only a few (<100) reflections will be saved to the final hkl-file. Si and GaSb as test materials, seemingly only require a few and very inaccurate data points to solve the structure. Other structures are rarely as symmetric, therefore are pure Si and GaSb not among the usual materials analyzed by 3D ED. This one of the main reasons why mordenite was introduced for the present study. Zeolites are commonly characterized by 3D ED due to its peculiar 3D microporous structure [21]. Mordenite for example, will give much more reflections in XDS (1000-10 000). The Si and GaSb unit cell dimensions are additionally small (lengths of 5/6 nm), resulting in diffraction patterns with large distances between spots, compared to many microporous materials (longest cell lengths of 15-25 nm). The difference between the test materials is shown in Figure 5.1, where (a) is from dataset

Si5_10, (b) is from dataset GaSb1_1, and (c) is from dataset Cu-MOR-K_31. All datasets are collected at an acceleration voltage of 80 kV and a nominal camera length of 120 cm. More experimental details about these datasets are available in Appendix A.1, Table A.1, A.2, and A.13. For both Si and GaSb, all reflections are not captured inside the frame, as the materials have small unit cell dimensions and scatter strongly. The diffraction spots are visibly more tightly packed in Cu-MOR-K (because of the larger unit cell) and all of the spots are within the frame.
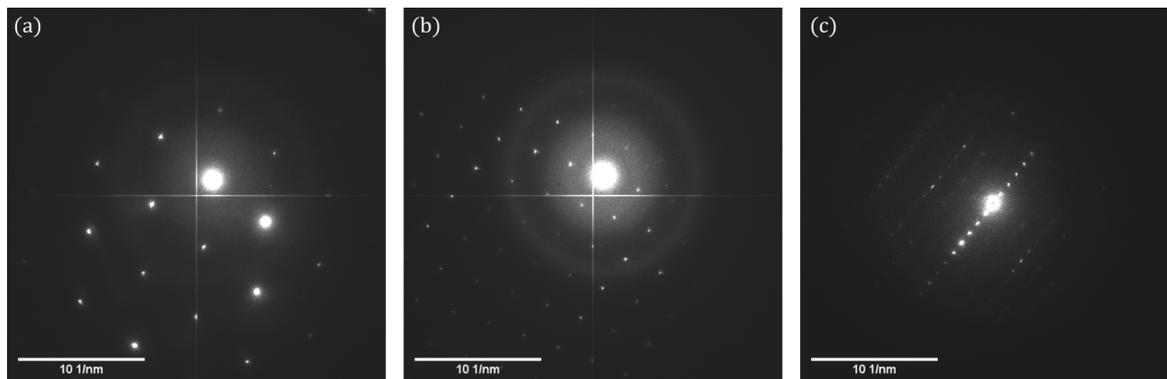
Figure 5.1: Arbitrary frames from material (a) Si, (b) GaSb, and (c) Cu-MOR-K, obtained at 80 kV and camera length 120 cm.

The loss of reflections are visible also in the 3D reconstruction of reciprocal space, see Figure 5.2(a), (b), and (c) for material Si, GaSb, and Cu-MOR-K respectively. Both for Si and GaSb, the 3D reconstruction look "square" around the edges, compared to the round sphere of Cu-MOR-K, because of reflections falling out of the frame. The maximum resolution is reduced. This issue can commonly be accounted for by reducing the camera length for small unit cell materials, which will be discussed further in Section 5.2.8.
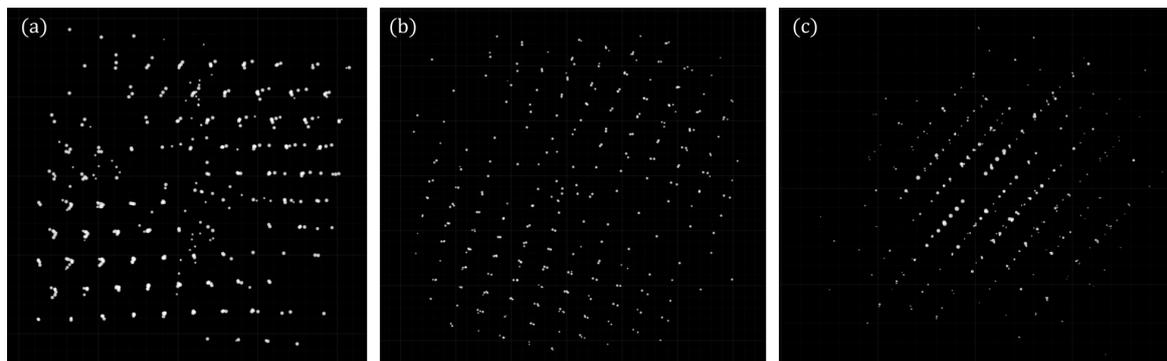
Figure 5.2: Slices of 3D reciprocal space reconstruction for material (a) Si, (b) GaSb, and (c) Cu-MOR-K.

Note that this Cu-MOR-K dataset has a short rotation range (and few number of frames) and does not have as big of 3D reconstruction as it potentially can. The maximum resolution is however, sufficient, i.e. over 1 Å^{-1}. As there were some problems with the initial data obtained at NTNU, the two example datasets, MOR1 and MOR2, obtained at Stockholm University by Cichocka et al. were used for many of the analyzes in this work [2, 65]. Mordenite made it possible to inspect a 3D material of anisotropic nature. For refinement, it is possible to simplify the structure, assuming that Si and Al are strong scatters (with atomic number 13 and 14) and that there are no cations in the pores.

To summarize, Si and GaSb were deemed nonideal for 3D ED studies, because of their strong scattering power and small unit cells. MOR variants are used in this work because of the structure and the reference data available. The material is deemed a good choice to set up, explore, and improve the 3D ED data processing method.

5.1.2   Choice of TEM specimen grid

A common problem for TEM tomography techniques is that the edges of the grid can shadow the signal at high tilts [91]. However, for 3D ED, this can become an even larger problem as the tilt angles are high. For many of the beam sensitive zeolites, this grid edge shadowing were not a major problem during the tests. It has to be said that these samples were destroyed by the beam before reaching maximal tilt angles. However, for Si and GaSb on a 100 um thick Si frame with 100 wide windows [1], edge grid shadowing was the main limiting factor for further expansion of the rotation range.

In order to avoid risk of grid edge shadowing, one would like to have large mesh sizes (few windows per plate). Large mesh sizes give more space for the particles to spread out, possibly further away from grid edges. In this experiment, different grids were used. However, only the Cu-MOR-Na material datasets used such a tomography TEM grid (with a 100 M mesh size). All other samples used "normal" grids with mesh size 200 or 300. For future experiments, it is suggested to continue using the mesh size 100 tomography grid.

In this experiment, different substrates were used for the different materials analyzed. Different substrates can cause different background and noise levels. The Si and GaSb datasets were dispersed on a Si frame with a silicon nitride supporting thin film, while the different MOR-datasets obtained in this experiment, i.e. Na-MOR and K-MOR etc., used a Cu grid with a 20 nm thick amorphous holey carbon cover film. The holey carbon cover film is used for more uniformity in background, which is favorable. In this work, no further investigations of the differences in substrate were done.

A smaller mesh size in the grid is suggested for future experiments to avoid edge-shadowing effects. Holey carbon films can enhance uniformity in background.

5.2   Data collection

In this section, limitations of different parts of the experimental setup are investigated, including the hardware, such as detector and microscope, goniometer and holder. Furthermore, the effect of different data collection parameters will be investigated, by studying a series of Na-MOR datasets, collected at varying data collection parameters, i.e. CA, spot size, bitdepth, acceleration voltage, flatfield correction, and camera length. The experimental details are available in Appendix A.1, Table A.3 and A.4, A.5, and A.6. The aim of systematically varying these parameters was to find out what parameters work best for 3D ED structure analysis of zeolites and how they can be changed to correct for potential undesired effects. Moreover, parameters that potentially could solve the observed issues of overexposed pixels and beam sensitivity will be identified and tested.

5.2.1   Hardware limitations

The JEOL ARM 200F is a double corrected instrument with a coldFEG electron source. The image corrector is essential for other TEM techniques, however redundant for 3D ED. In fact, it can distort the diffraction signal [92]. In a previous study on the effect of optical distortions on the accuracy of 3D ED lattice parameter estimations, the effect of the distortions from

a potential image corrector was not included [93]. Hence, the effect of image corrector distortion might be interesting to investigate further.

To have a small (few nanometers) parallel electron probe is important in 3D ED, because it allows collecting diffraction data from smaller sample volumes [10]. However, the JEOL ARM200F microscope used here, is not optimized for low magnification parallel beam. The detector or the positioning of the detector used in the experiments, Quantum Detectors 4R Merlin EM DED, can also be part of the reason why it is difficult to use low magnification parallel beam with the current setup. This detector is a hybrid DED. Hybrid DEDs usually have very sharp PSF at lower kVs, but suffer from limitations of being physically large, having a small number of pixels, and poor PSF at high kVs [53]. However, they are well suited for diffraction. Note that other types of DEDs exist as well, e.g. Gatan K2-summit, which have lot of pixels and are optimized for low dose, but hence not ideal for diffraction [94]. The Stockholm setup uses a 512x512 Timepix hybrid pixel detector which is related to the Medipix chip used in the present work.

5.2.2 Goniometer stability

The goniometer stability of the TEM is important for 3D ED to work properly [9]. In the reconstruction step it is expected that the tilt steps are uniform and optimized gonios are one of the main features of dedicated 3D ED setups [17]. Due to issues with the obtained data, one hypothesis is suggested that there was something wrong with the stability of the goniometer. The linearity in stage rotation assumed by Instamatic may not be in agreement with the physical setup. If so, this could lead to an incorrect oscillation range and further incorrect reconstruction. To calculate the oscillation range, Instamatic reads the start angle, end angle, and number of frames. Evidently, the angle is not read out for each frame but rather calculated from the above read out parameters. A wrong oscillation range might therefore be detrimental for the reconstruction.

Technically, tilt is driven by voltage to a motor to make rotating gears turn. If the voltage is not high enough, i.e. above a certain threshold, the gears may not be able to jump to the next notch and thus jump back to the previous tilt step. More voltage is then required to tilt at higher angles. This could possibly be tested by plotting the gonio response as a function of input tilt and observing whether it is linear or not; however, this was not done in the current experiment. Another possibility for instabilities in the goniometer is that the goniometer tilt around 0 degrees is worn out. This can cause instability in the tilt region because this is the range that is used the most for other experiments (not 3D ED). For the MOR reference datasets, a minimum rotation range necessary for the structure to be solved was found to be only 4.87°, see Section 4.2.1. A long rotation range is therefore not necessarily needed, which indicates that using smaller wedges of only a specific tilt range, e.g. around 0 degrees, could be possible.

If the problem is not the goniometer itself, but in the signal control it could possibly be an inconsistency in the Instamatic-microscope coupling. Instamatic is developed to perfectly fit the setup of the inventors, i.e. Stockholm University, and might therefore encounter issues when handling other hardware. In addition, user mistakes during data collection can also cause errors. For example, sometimes the tilt might stop before the data acquisition is stopped. This can lead to an incorrect oscillation range as Instamatic will read out more frames compared to the final tilt. The total oscillation range will then be decreased. This practice should hence be strictly avoided, although if it occurs, datasets obtained like this should be flagged with a warning. In this experiment, errors like this were reported to occur during data collection, but the datasets were not flagged, thus the effect of this error was not possible to investigate further. Although the principle behind 3D ED is simple, there is

room for error and it can be challenging to perform the experiments on equipment that is specialized for other types of experiments (for ARM it is mainly high resolution transmission electron microscopy (HRTEM) and EELS that are the applications).

Choice of holder

Two different holders were used in this experiment, the Gatan 916 High Tilt Tomography Holder and the JEOL tomo tip with JEOL single tilt holder. There exists an additional tomography holder at NTNU, Fischione Instruments Model 2040 Dual-Axis Advanced Tomography Holder with tilt range $\pm 70°$, that is possible to use. This holder can both be tilted (movement around x- and y-axis) and rotated (movement around z-axis). The maximum tilt-range of these holders are in reality limited by the positioning pole piece (the JEOL ARM200F has a high resolution pole piece (HR)). A lower pole piece gap is advantageous for higher resolution, but can limit the tilt range of high-tilt tomography holders, as the ones used here. Anyway, the goinio design is limited to $\pm 81°$ independent of the holder used.

During an initial investigation of the tilt linearity of the JEOL tomo tip with JEOL single tilt holder, it was discovered that it occasionally suffers from asymmetric tilting. Possible solutions to this problem involve routinely checking the linearity of the holder before every tilt series collection by running a full tilt series in imaging mode to reveal whether or not the asymmetry problem is present. If asymmetric tilting is observed, the tilt series should only be acquired from a rotation range that is functioning well (linear) and rather merge datasets to obtain acceptable rotation range completeness. Alternatively, the holder can be moved to a stable position inside the TEM sample holder chamber. Then the asymmetry check can be redone to ensure linearity.

Dataset K-MOR-17 was cropped into four subsets of 100 frames, corresponding to an angular range of $19.11°$, each to observe whether the linearity of the single tilt holder were the reason for the dataset problems (low completeness and CC1/2). No differences were observed except for beam damage between the four subsets, see Section 5.2.5 and Figure 5.3. This was not possible to do for datasets that were too poor to be analyzed both as complete and as sliced datasets. More studies focusing on this specific issues could preferably be done in the future.

5.2.3 Timing in cRED

The goniometer rotates constantly throughout the data acquisition, including the intervals for crystal tracking where defocused frames are collected [65]. Thus, control of timing is important. The defocus cycle needs to cover an integer multiple of the rotation range step in order for the oscillation range to be correct for the whole stack of frames. This is essential for the consistency of the dataset. XDS and DIALS will not run properly if the oscillation range is incorrect. It is also important to note that changes in the electron beam are not instantaneous because of hysteresis among other things. The larger the defocus, the longer it takes for the beam to return to its refocused state. As seen in the tabulated data in Appendix A.1, average acquisition time is tracked for each dataset. The acquisition time is defined as the sum of the readout time, the exposure time (usually 500 ms or 300 ms), and the overhead (allocation and arranging of memory, i.e. 3-4 ms) [65]. With the Merlin readout system, readout times are 1.64 ms at 24 bits and 820 µs at 12 bits (standard mode) [95]. The exposure time of the defocused frames is usually 10 ms. For other setups, automatic crystal tracking and descan takes about 200 ms, which depends mainly on the response of the computer-to-deflector interface [96]. For the data obtained at the NTNU setup, it was observed that the acquisition time typically was much higher (200-400 ms) than the

exposure time. This is very high compared to the Stockholm University setup, that reported a difference of 12 ms between the acquisition time and the exposure time [65], despite their detector having a higher readout time (8 ms for Timepix). In the NTNU ARM200F setup, the Merlin-software is installed on a different machine than Instamatic, which might be the reason for part of this delay. Initial tests showed a delay of 60 ms when transferring data between the programs on the two computers. This extra overhead may be caused by the coupling between Instamatic and the MerlinEM detector and must be reduced and improved before trustworthy datasets can be acquired routinely.

The data obtained at the NTNU setup, was observed to have gaps of 2-3 frames for every captured defocused pattern. Even with a preset image interval of 5 or 10, meaning a single defocused frame is supposed to be captured every fifth or tenth frame, more frames seem to be removed each time. The folder with defocused images contain the correct frame numbers, e.g. 5, 10, 15 etc., and total number of collected defocused frames. However, files are missing from the folder with "normal" diffraction pattern frames, i.e. the frames supposed to be collected right after the defocused frame (typically frame 6, 11, 16 etc.). This issue could potentially be related to timing, either due to the long overhead mentioned above, or the hysteresis of the defocus. In Section 4.2.1, the effect of removing gaps of frames and increasing the density of excluded frames, was investigated, which will be further discussed in Section 5.3.1.

To summarize, there is a possibility that the timing of the defocus might not be in agreement with the acquisition time and the setup at this point might need recalibration or support from the Instamatic developers who recently (mid 2023) could be used on a JEOL-Merlin setup.

5.2.4 Rotation axis

The rotation axis is found by inspecting a series of diffraction patterns, in for example REDp [97]. Another way of determining the rotation axis is by using EDtools' `find_rotation_axis`. This tool performs analyses based on the XDS output file SPOT.XDS. Alternatively, it can be used as a supplementary tool for confirming the previously found rotation axis. The direction of tilt is by default (in Instamatic, REDp and XDS software) assumed to be from negative degrees to positive degrees, e.g. from -45.2 to +32.5 degrees. If data is obtained the other way, e.g. from +45.2 to -32.5 degrees, the rotation axis in XDS will therefore be flipped compared to the "normal" rotation axis or the one found in REDp. This effect was confirmed by observations done by researchers in Stockholm as well. Thus, the data collector should be aware of either always tilting by the default method (negative to positive) or always beware of flipping the rotation axis in data processing.

5.2.5 Measures for beam sensitive materials

Zeolites belong to the category of beam sensitive materials, as mentioned in Section 2.6 [31], and in this work a powerful TEM is used (with high brightness). The incident electrons from the beam can cause irreversible change both physically and chemically to the sample material, known as electron beam damage [31]. Beam damage from electron microscopes is described in Section 2.4.4. For zeolites, the damage caused by the electron beam causes amorphization as the intricate aluminosilicate framework can collapse [98, 99].

A possible example of beam damage from the data in this work is shown in Figure 5.3. Here, K-MOR-17 were cropped into four subsets of a bit less than 100 frames each, and the detected diffraction spots for each range of frames are visualized in XDSGUI. Experimental details for K-MOR-17 are available in Table A.9. Figure (a) to (d) show gradual degradation
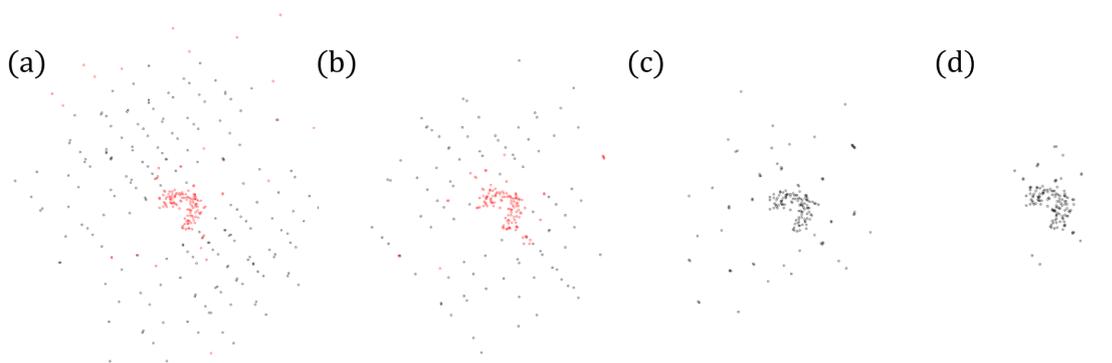
Figure 5.3: Beam damage in zeolite dataset K-MOR-17. Datasets were sliced into four parts that contained ∼100 frames.

of the material, as fewer reflections are visible later in the rotation range. If the beam damage is severe for a material, merging data can be important as demonstrated in Section 4.2.2, whereas only the useful range of frames, i.e. before degradation, is used.

The ratio of Si and Al present in the material framework is one of the most limiting factors for the beam sensitivity of zeolites [100]. According to Csencsits et al., a total Si Y zeolite will tolerate three and five times higher dose than high and low percentage Si, respectively. As this work considers zeolites with a considerable amount of Al with a Si/Al ratio of $\approx$ 7 [80], beam sensitivity is expected to be considerable. Furthermore, the Na^+, Cu^{2+}, and K^+ ions are susceptible to being reduced by the electron beam. The ions might then become free to migrate inside or even out of the zeolite [101]. Some materials can have preferred orientations, meaning they tolerate lower dose in certain directions.

The expected tolerated dose may additionally depend on the size of the analyzed zeolite. This can be due to thermalization, as described in Section 2.4.4. For smaller samples, this heat can accumulate in the illuminated area [31].

A possible solution for beam sensitive or vacuum sensitive materials, is using other sample preparation methods, such as plunge freezing. The particles will then be surrounded by an amorphous mass (ice) which protects them from beam/vacuum damage [102]. This amorphous mass may additionally require different measures to be taken when finding suited particles during data collection. The sample will look more smudged in imaging mode.

In general, it is preferable to lower the beam voltage for less charging effects. However, lower beam voltages may cause more radiolysis. One could additionally limit the amount of electrons hitting the sample, i.e. the dose, by using narrower spot size and CA. These data collection parameters, i.e. acceleration voltage, CA, and spot size, that affect the dose and hence also the observed intensities, will be investigated in the following subsections. Using a smaller SA aperture, however, will not reduce the dose experienced by the crystal as it is placed below the specimen. Hence, the effect of changing the SA will not be investigated in the present work. Instead it can decrease background and amount of noise, by hindering undesired local information and inelastic scattering signals [65].

Condenser aperture and spot size

One of the main technical parameters that affects the dose is the condenser aperture (CA), as the choice of CA will affect the fraction of the electron beam hitting the specimen. Hence, will it also control the intensity of illumination i.e. how many electrons in a given exposure time will hit the specimen. This is tested to learn how to lower the dose experienced by the studied material.

Two tilt series on the same Na-MOR particle were taken with identical data collection parameters except varying CA. Na-MOR-10 and Na-MOR-11 datasets were set to use a CA of size 150 μm (CA1) and 100 μm (CA2) respectively. These two datasets were acquired with an acceleration voltage of 80 kV and camera length 80 kV. Two additional datasets taken at 200 kV with camera length 150 cm, i.e. Na-MOR-22 with CA1 and Na-MOR-28 with CA2, were also analyzed. Though these datasets do not originate from the same particle, they do share most other parameters. The average intensities with corresponding standard deviations of the four datasets are plotted in a barplot Figure 5.4(a). Note that the average intensity can be affected by e.g. the size of the particle (not just the thickness), as well as how much of the particle is kept inside the SA during the tilt. The in total four datasets showed conflicting results. For 80 kV, the dataset with CA1 had the highest intensities, while for 200 kV, CA2 had the highest intensities. This may be due to differences like which zone-axes are crossed during tilt etc.
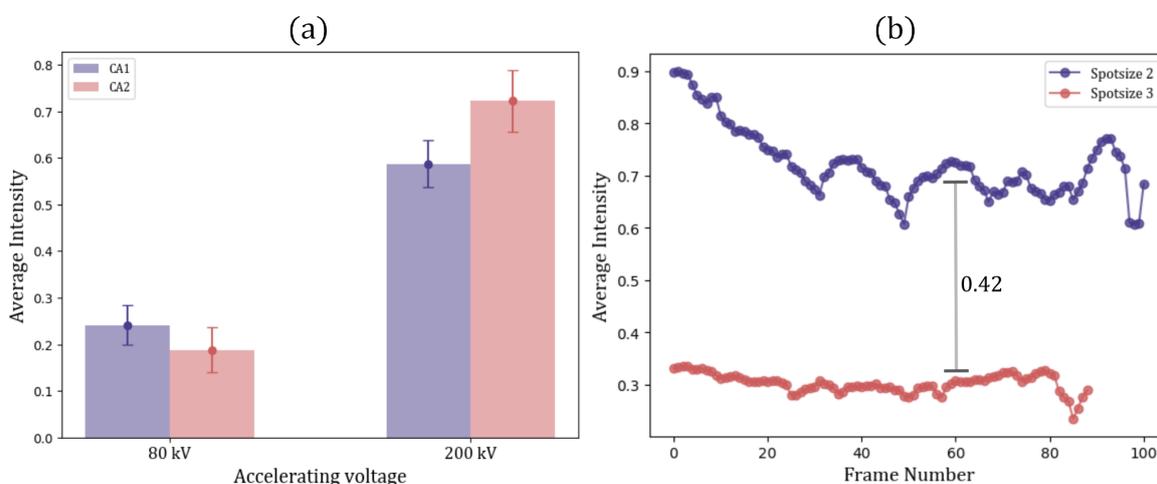
Figure 5.4: (a) Barplot showing the effect of changing CA on average intensity through frames, tested for two 80 kV dataset and two 200 kV datasets. (b) Average intensity through frames of datasets with different spot size. The average difference between dataset intensities is 0.42, marked with a gray bar.

The spot size of a TEM further adjusts the number of electrons passing through the CA. As the first condenser lens is stronger, the focus length is reduced and a smaller fraction of the total spread is captured by the second condenser lens [103]. The higher spot size number, the fewer electrons pass through (smaller radius). The effect on the resulting diffraction patterns is presented in Figure 5.4(b) and 5.5.

Two tilt series on the same Na-MOR particle were taken with identical data collection parameters except for varying spot size. Dataset Na-MOR-28 and Na-MOR-31 were set to have spot size 2 and 3 respectively. Figure 5.4(b) shows the average intensities through all frames in the two datasets. As may be seen, the dataset with the higher spot size has lower average intensities. The difference of changing spot size is seemingly much larger than for changing the CA. To further visualize the difference in intensity for the two cases, the intensities of the inner and outer main background noise rings of both datasets were analyzed. Figure 5.5(a) and (c) shows the intensity of the inner background noise ring, and Figure 5.5(b) and (d) shows the intensity of the outer background noise ring, for spot size 3 and 2 respectively. for spot size 3 and 2 respectively. The intensities at arbitrary points on the background noise ring circumferences, point (342, 214) in the inner background noise ring and point (385, 213) for the outer background noise ring, are measured with the intensity

tool in REDp. The radius of the direct beam is larger for the smaller spot size. Although the background noise rings are present in both cases, (a) have intensities in the background noise rings higher than those in (b).
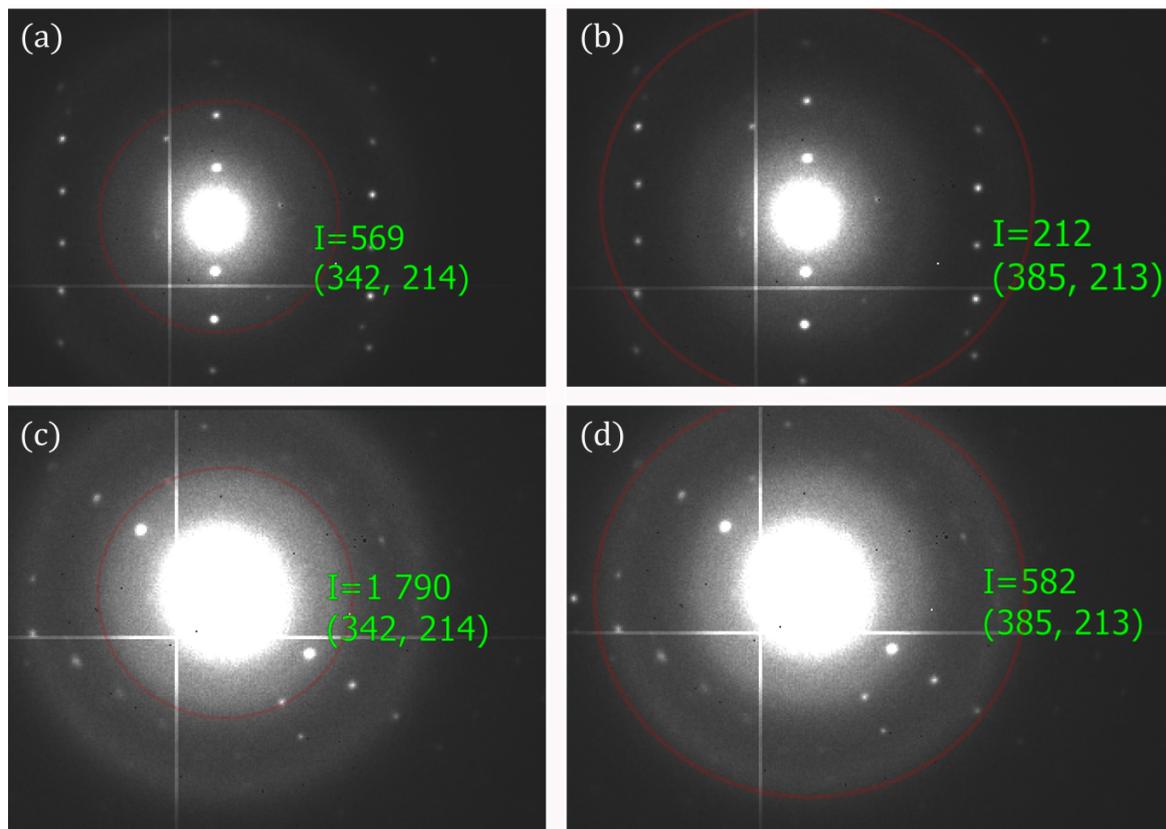
Figure 5.5: The effect of different spot sizes on diffraction patterns. The subfigures (a) and (b) have spot size 3 and the subfigures (c) and (d) have spot size 2. The background noise rings are drawn in red. The intensities of the background noise rings are shown in bright green. Figures (a) and (c) show the inner background noise ring, while figure (b) and (d) show the outer background noise ring.

Note that the analysis on spot size was done only on one pair of datasets, versus the CA analysis was done on two. The CA analysis was deemed inconclusive, whereas the spot size analysis was deemed to have a considerable effect on the observed intensities. Preferably, more datasets should have been analyzed in order to investigate this issue with certainty. More importantly, such analysis should be done quantitatively, i.e. measure the dose for the different settings, CA and spot size. Then the findings can be linked to the critical dose for a given material [55] and linked to the dynamic range of the detector, see next Section 5.2.6.

Choosing optimal acceleration voltage

The first 3D ED experiments were obtained with an acceleration voltage of 80 kV. However, previous papers on 3D ED usually use acceleration voltage of 200 kV [21, 9, 10]. Hence, this voltage was tested on a tilt series of the Na-MOR material. The changed voltage required a different optimal camera length. Generally higher intensities, broadened PSF, and more background were observed for the higher voltage. Note that all datasets in the Na-MOR series have slightly varying data collection parameters, such as camera length, spot size, and CA, see Appendix A.1 for more details. Hence, the standard deviation takes larger values.

A line plot of a row of reflections was performed of frame 73 from Na-MOR-18 and frame 47 from Na-MOR-19. The frames, with the row of reflections analyzed highlighted with a yellow square, are shown in Figure 5.6(a) and (b), respectively, along with the corresponding intensity profiles. The width of the diffraction spot intensity peaks as well as the width of the direct beam are shown in the intensity plot. Visibly, both datasets suffer from saturation of pixels at intensity 4095 visible as a gray dashed line. At 80 kV, the diffraction spots have
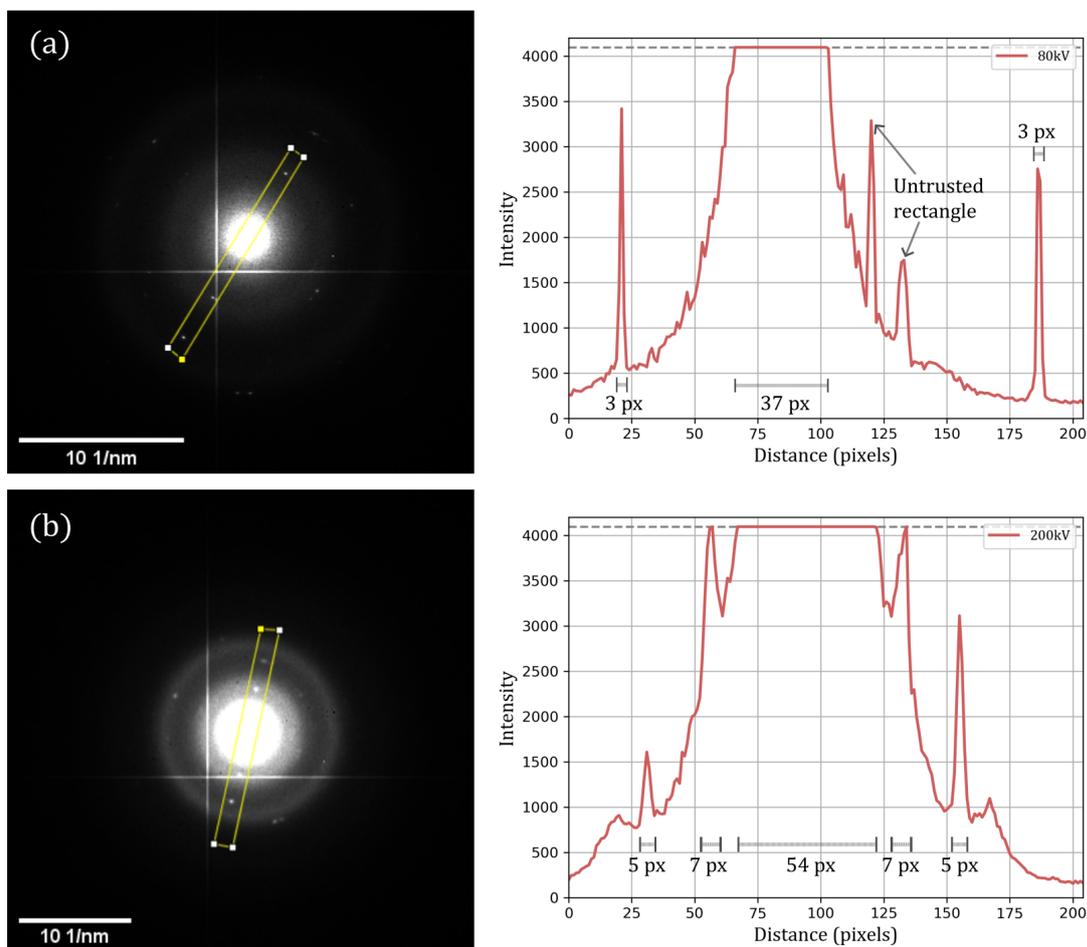
Figure 5.6: Line scan with corresponding intensity profile of (a) frame 73 from 80 kV dataset Na-MOR-18, and (b) frame 47 from 200 kV dataset Na-MOR-19.

a spread (width) of 3 pixels for both 1st and 2nd order maxima, whereas the 200 kV dataset have spreads of 7 and 5 pixels for its 1st and 2nd order maxima respectively. This, indicates that a higher acceleration voltage causes higher PSFs, so that the individual diffraction spots are spread out over more pixels. The diameter of the direct beam is observed to be larger for the higher-voltage dataset. The experimental difference between Na-MOR-18 and Na-MOR-19, additionally to the acceleration voltage, is the CA used. Na-MOR-18 uses CA2 and Na-MOR-19 uses CA1 which have a radius of 100 and 150 µm respectively. This may have affected the dose and the observed intensities. However, since the same spot size was used, which previously was observed to have a larger effect on the observed intensities than the effect of the changed CA, the effect of the changed CA is in this case assumed negligible.

Figure 5.7 shows the difference in the number of bright pixels for the 80 kV dataset vs. the 200 kV dataset. Figure 5.7(a) shows the number of bright pixels in the previously analyzed two datasets Na-MOR-18 (80 kV) and Na-MOR-19 (200 kV) with and without the

untrusted rectangle. These intensity profiles correspond well with the general trend for the 80 kV and 200 kV datasets, see Figure 5.7(b) and (c), that show the number of bright pixels for the 80 kV and 200 kV datasets respectively. The mean of all datasets is shown with a dashed black line. The mean number of bright pixels of the 200 kV datasets is 4051, which is about three times as high as the mean of the 80 kV datasets, 1404. For reference, the total number of pixels in a frame is 262 144.
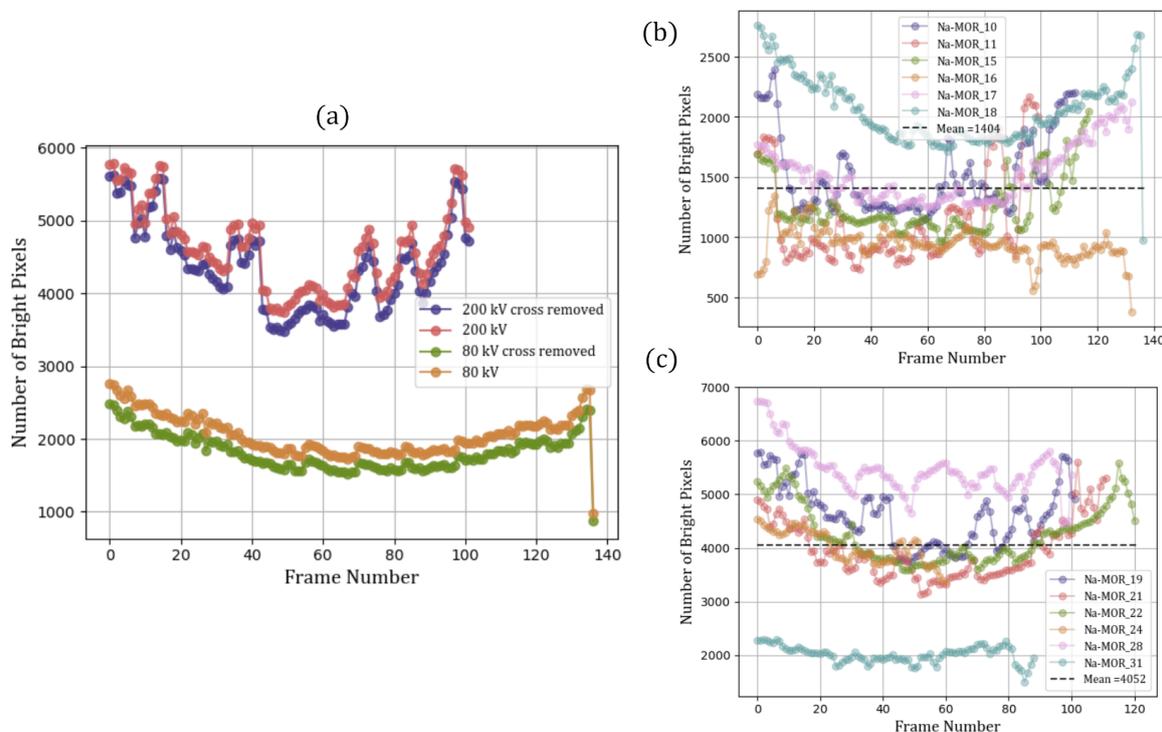
Figure 5.7: Number of bright pixels per frame, plotted for all frames in (a) 80 kV dataset Na-MOR-18 and 200 kV dataset Na-MOR-19. The equivalent datasets with the untrusted rectangle removed are plotted for reference. (b) all 80 kV Na-MOR datasets. (c) all 200 kV Na-MOR datasets.

Despite larger differences in data collection parameters between the 80 kV and 200 kV datasets obtained for Na-MOR, the mean of the number of bright pixels in each frame of the datasets seems to indicate a higher average counts in the 200 kV dataset than in the 80 kV datasets. At 200 kV, a single electron in the electron beam hitting the detector will generate many more electron-hole pairs in the detector, which means that counts are generated in several nearby pixels simultaneously (poorer PSF). At 80 kV, fewer electron-hole pairs are generated, which means that it will generally only be the pixel that was hit that will be "triggered" and activate a count (perhaps also some nearest neighbors).

A potential drawback with higher voltage is more beam damage, i.e. knock-on damage, rather than radiolysis, as mentioned in Section 5.2.5 [104, 40].

5.2.6 Overloading bits: Changing the bit depth

During the earlier data acquisitions, a lower default bit depth was used. The datasets obtained with this bit depth, 12 bit, showed intensity saturation in both the direct beam and occasionally in the diffraction spots close to the direct beam. A line plot of a row of reflections through the direct beam in selected frames was done for two samples with similar data collection parameters and different bit depth. The chosen frames with the line trace

marked with a yellow square are shown together with the resulting intensity plots in Figure 5.8. Frame 22 from the Na-MOR-15 dataset is used in Figure 5.8(a), while frame 93 from the K-MOR-31 dataset is used in Figure 5.8(b). Figure 5.8(a) has the default bit depth of 12 bit, while Figure 5.8(b) has bit depth 24 bit. Note that the scales of the intensity plots, i.e. the y- and x-axis, are different and not directly comparable. However, observing the direct beam (the highest peak) of both intensity plots, it is clear that the peak of Figure 5.8(a), with the lowest bit depth, is saturated at the top. This indicates that the highest intensities have been cut off at the peak intensity of 4095 (12 bit $= 2^{12}= 4096$ possible counts). The dotted black line shows the intensity at which saturation was observed, i.e. 4095. The higher bit depth seemingly avoids saturation of intensity at this level and thus allows for a broader range of intensities. 24 bit data can count intensities up to $2^{24}=16\,777\,216$ which should be more than enough for most experiments.
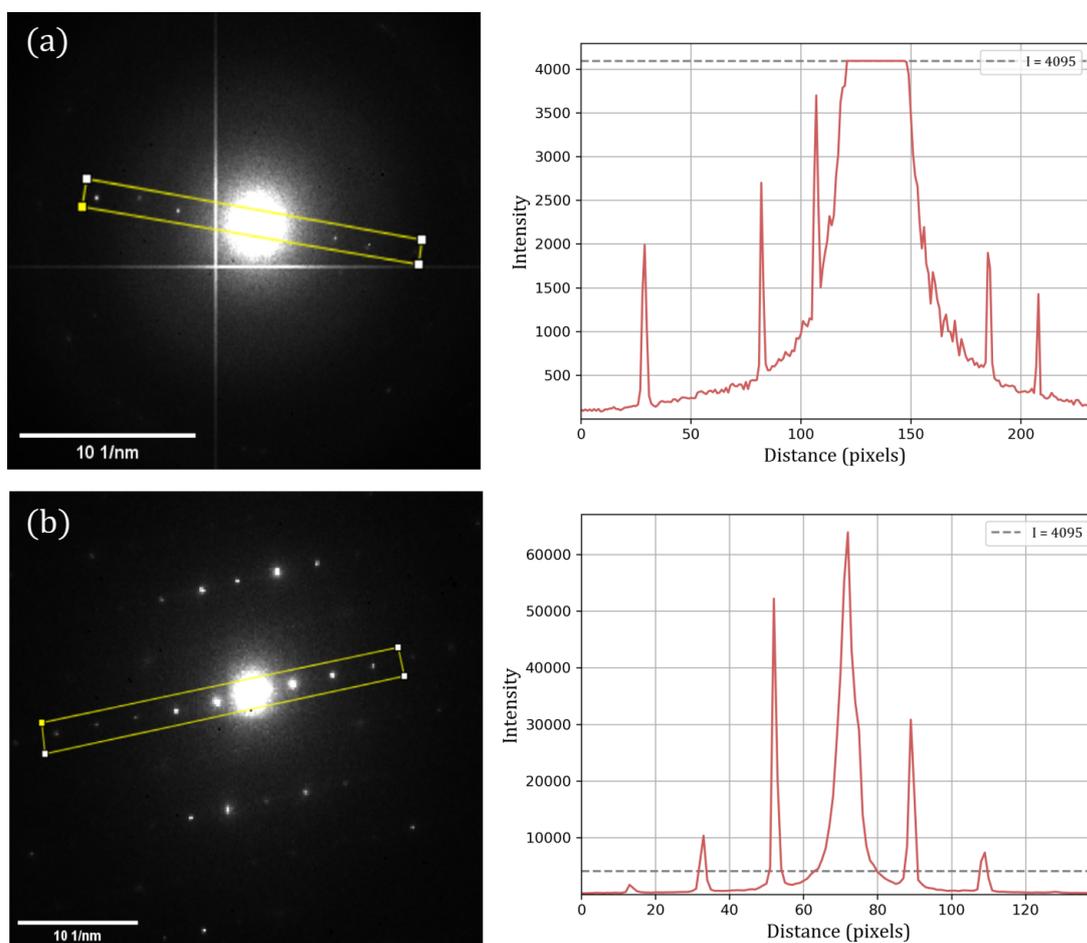
Figure 5.8: Intensity plot from line segment through a row of diffraction spots for datasets of different bit depth. The dataset in Figure (a), Na-MOR-15, is saved with 12 bit (saturation level 4095), while the dataset in Figure (b), K-MOR-31, has bit depth 24 bit (saturation level 16 777 215).

Note that the diffraction pattern in Figure 5.8(a) has a bright white "untrusted" rectangle with high intensity that has been wrongfully interpreted as a peak, visible as the peak left to the direct beam. Removal of this feature is described in Section 5.3.2.

These results shows that increasing the bit depth will not hinder overexposure, but rather make the bits able to save higher intensities than a certain cut-off, here 4095. For the naked eye, the 12-bit and 16-bit diffraction patterns will look almost the same. Only inspecting

the intensities, which can be done in REDp or XDSGUI, will reveal the difference between the two. A common effect of higher bit depth is an increase in the size of the datasets, i.e. a single frame will have a larger file size. This is not the case for our data, probably due to the data compression performed by Instamatic. The detector collects data in 2, 4, 12 or 24 bit, but the data needs to be saved as 2, 8, 16, or 32 bit. Note additionally that the readout time for the detector doubles when changing from 12 to 24 bit. One other thing is that for 24 bit, Merlin *must* spend time reading out data before it can take a new image. In the other (lower) bit modes, Merlin can start capturing a new image as soon as an image is finished, as long as the frame-time is lower than the time it takes to read out an image. Note that the latter is an option that must be activated in the software (continuous mode).

The saturation of detector pixels is, in fact, not a problem for most of the data processing procedure. REDp and XDS will have no problem finding and indexing peaks, and SHELXT will be able to solve the structure just fine. First, when refining, the magnitude of the structure factors (intensities) will be important. In the refinement step, as described in Section 2.3, the calculated and experimental structure factors will be compared to each other. Saturated intensities will cause a constant mismatch between the calculated and experimental intensity values. R-values will hence be high, it may be hard to converge to a minima, and the refinement will therefore fail or be unstable.

5.2.7 Flatfield correction

Using the built-in flatfield-correction in Instamatic[†] was observed to have considerable effect on image quality in this experiment. Flatfield correction is mainly done in order to correct for non-uniform illumination as well as sensitivity in detector and imaging system. The correction hence reduces systematic artifacts and improves data quality.

Flatfield consist of a dark reference (dark current) image and a gain reference image (the pixel gain). The dark reference image is created by blocking the direct beam, i.e. no illumination, meaning the camera records a black image. The dark current originate from thermally generated electrons [51]. The gain reference, on the other hand, is captured with the brightest possible uniform illumination so that the camera records a white image. If one assumes that each pixel in the imaging system has a linear response function to the recorded electron count, then the dark current makes up the additive component and the gain makes up the multiplicative component of flatfield. Further, the dark current will be subtracted from the gain and then divided by the mean and inverted [105, 106]. An example of a flatfield image from this experiment is shown in Figure 5.9(a), that were obtained along with the Na-MOR-datasets. The flatfield image contains a white cross, known as the untrusted rectangle, and numerous dead pixels, some white and some black. These are typical defects that are corrected by flatfield. The flatfield correction does not account for contrast variations caused by changes in imaging conditions [105].

The method, as used in this work, had great implications on the resulting image quality, as illustrated by the difference between Figure 5.9(b) and (c). Figure 5.9(b) and (c) show frame 43 from Na-MOR-15 and frame 78 from K-MOR-31 respectively. The experimental details of these datasets are available in Table A.4 and A.10. Less background noise, no untrusted rectangle, and reduced intensity background noise rings were observed. In addition were certain permanent dead pixels removed. However, a new phenomenon that emerged, potentially as a result of this correction was a handful of random low-intensity spots inside the direct beam, see Figure 5.9(d). The number, magnitude, and position of the low-intensity-

---

[†]https://github.com/instamatic-dev/instamatic/blob/main/src/instamatic/processing/flatfield.py
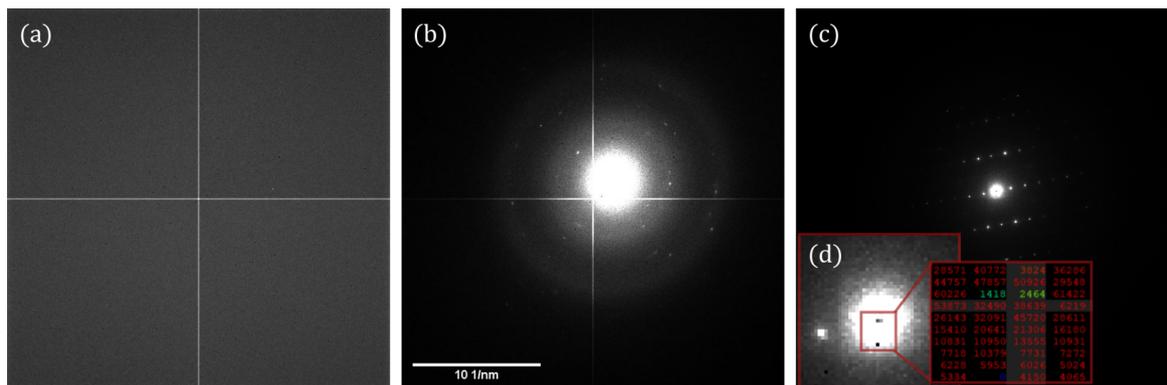
Figure 5.9: (a) flatfield image. (b) frame 43 from Na-MOR-15 without flatfield correction. (c) frame 78 from K-MOR-31 with flatfield correction. (d) cut-out of the direct beam in frame (c), showing the corresponding magnitude of intensities in the direct beam.

spots varied for each frame in each dataset. Note that the same datasets were analyzed in Figure 5.8 for the effect of changing bit depth. Both flatfield and changed bit depth can possibly have affected properties such as extent of direct beam and amount of background.

5.2.8   Does one size fit all? Finding the ideal camera length

The differences between camera lengths have previously been evaluated in the project thesis [1]. However, this analysis were based on fewer camera lengths than were analyzed in the current experiment and only Si was used as test material. The ideal camera length will naturally depend on the material used, i.e. the scattering capabilities, symmetry, and unit cell dimensions, see Section 5.1.1.

The general measures to take for choosing camera length are summarized here. First of all, one wants to capture all the relevant reflections. It is therefore important not to exaggerate the camera length, as this may cause reflections to fall out of the frame. This was an issue for Si [1] as the unit cell is relatively small, while the scattering power is large. High index reflections would in these cases fall out of the frame. Note that the physical location of the detector affects the camera length, and can be a limiting factor for finding the ideal camera length. On the other hand, reducing the camera length too much introduces the risk of not being able to separate each spot. This is more important for lower-symmetry materials with larger unit cell dimensions. Figure 5.10 shows frames from datasets obtained at 80 kV for material Na-MOR, i.e. Na-MOR-11, Na-MOR-18, and Na-MOR-15, at different nominal camera lengths (a) 80 cm, (b) 100 cm, and (c) 120 cm. The closest spots of each dataset is magnified to the lower right of each frame to see the number of pixels that separate the spots. According to [23], the minimum number of pixels that separate two diffraction points should be at least 6. For nominal camera lengths of 80 cm and 100 cm, the number of pixels separating the spots is below this threshold. Camera length 120 cm was thus chosen for the other datasets.

This section has investigated the effect different features of the data collection procedure has on the resulting 3D ED data. For further experiments, a medium small CA (size 100 µm), a fairly high spot size (spot size 2), lower acceleration voltage (80 kV) and a higher bit depth (24 bit) were used. These parameters provided sharp intensity peaks with minimal noise while hindering saturation of bits and, to some extent, beam damage. Flatfield removed permanent dead pixels and counts in the untrusted rectangle. The issues regarding the goniometer and
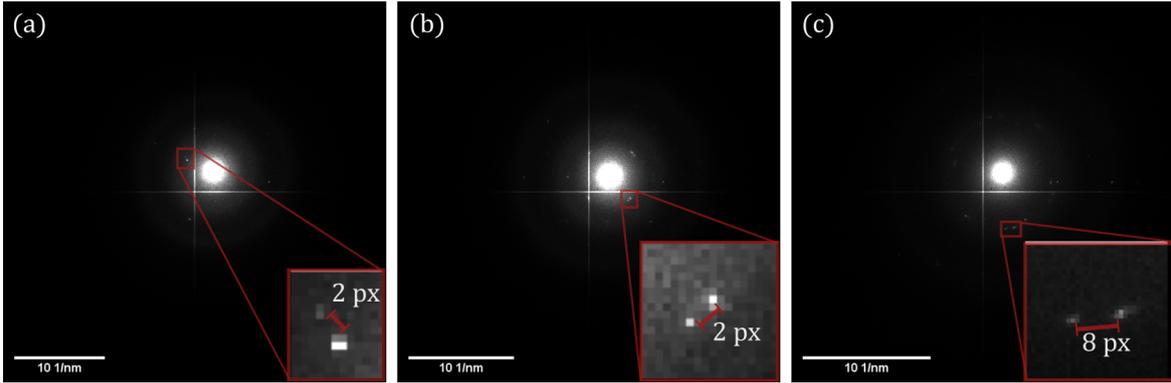
Figure 5.10: Diffraction patterns with nominal camera length (a) 80 cm, (b) 100 cm, and (c) 120 cm.

holder stability and timing of defocused frames should be further investigated. In the next section, the data processing procedure and how it can be used to reveal necessary data collection changes will be evaluated.

5.3 Data processing

This section will assess the performance of the data processing tools used in this work, according to the workflow described in Chapter 4. Python preprocessing for distortion correction and removal of the untrusted rectangle is covered in detail. Other features, such as the minimum requirements for input datasets, involving rotation range and frame density as seen in Section 4.2.1, will also be discussed.

5.3.1 Minimum datasets requirements

In Section 4.2, a series of tests were performed on the reference MOR-datasets MOR1 and MOR2, in order to determine the minimum requirements for an input dataset in XDS. The results were presented in Figure 4.7. The plots describing completeness and number of reflections commonly follow similar trends. A significant detail about the CC1/2-plots (the mid row), is that the CC1/2 values usually take almost maximal values ($\sim$100%) or very low values ($\sim$0%). Hence, CC1/2 may be a difficult parameter to track and observe trends from.

In Figure 4.7 (a) characteristics for a decreasing density of frames are plotted. This study was done on the whole range of frames. Dataset MOR1 has higher completeness, CC1/2, and number of frames for the whole range of fractions of excluded frames. This might be due to MOR1 having a larger total amount of frames than MOR2. From Figure 4.7(a) and (b) it is possible to observe trends that suggest the fraction of excluded frames should not exceed 0.5, and the minimum amount of frames needed is about 21, to ensure enough reflections are available for reciprocal space reconstruction. This minimum number of frames correspond to a rotation range of 4.85°. Figure 4.7(c) show for MOR2 a seemingly flat curve, and for MOR1, arbitrarily spread data points just above or just below the critical thresholds for CC1/2 and completeness. The graphs indicate that the data is not affected by a large gap between a group of frames. Assuming the gap between frames has no effect, then the MOR1 curves in Figure 4.7(c) actually illustrate that the differences in the number of reflections found in frames can vary greatly within a dataset. Large internal differences in number of reflections detected in frames can cause the curves in Figure 4.7(a) and (b) to look different, if based on other frames in the datasets. Hence, these results might not be directly applicable to the data obtained at the NTNU setup.

The results suggest that small rotation ranges and few number of frames are required for a 3D ED analysis to work. The data should have a critical density of frames, i.e. not too many frames excluded at regular intervals (every other frame or more). However, larger gaps between longer ranges of frames in the data do not affect data quality.

5.3.2 Python preprocessing

Preprocessing of the data stacks was shown to be useful for elliptical distortion correction and removal of the untrusted rectangle. Methods for these two correction techniques are described in the following section. Python was used for this purpose, with libraries such as `Pyxem`, `Numpy`, and `HyperSpy` [107].

Au distortion correction test

Distortion correction tests are commonly done on a well-known reference material. In this work, a sample of Au was used. A series of 80 kV SAED data for different nominal camera lengths ranging from 30 cm to 200 cm, were obtained at JEOL ARM 200F during the same session as the K-MOR-datasets. The main type of distortion we want to correct for is elliptical distortions. Different ellipticity correction tools are available. XDS has specific input parameters in XDS.INP called X-GEO_CORR and Y-GEO_CORR that apply an affine transformation to each image [108]. The parameters require geometrical correction input files of `.cbf` format, which can be generated by Instamatic, but this was not done in the current experiment [65]. Another way of correcting for ellipticity is by preprocessing the data in `Python`. Packages like `Pyxem` are specialized for this purpose and were hence used in this experiment. The script used is available from GitHub*.

The process of correcting for elliptical distortions follows certain steps. First, one wants to determine the lens distortions and scale in real and reciprocal space. Usually, the output of this is an affine matrix with parameters for correcting distortions and a number for the calibration scale of the diffraction pattern. The calibrated scales and eccentricities for each camera length are tabulated in Table 5.1. The matrix and the scale can be applied to the other 3D ED datasets obtained. The matrices were observed to be quite similar for all camera lengths, which is reflected by the similar eccentricities. An elliptical distortion with an eccentricity of 0.195 was observed for the most used camera length 120 cm. The mean of the eccentricities of all camera lengths is 0.200±0.003. A similar eccentricity has been observed for the setup that collected the reference MOR-datasets [65]. The effect of

Table 5.1: Manually calibrated scales and eccentricities from Au test data.

Camera length [cm]	Calibration scale [Å^{-1}]	Eccentricity
30	0.0223±0.0003	0.205
40	0.0176±0.0001	0.200
50	0.0137±0.0002	0.203
60	0.0119±0.0001	0.197
80	0.00823±0.00002	0.199
100	0.00707±0.00004	0.202
120	0.00584±0.00006	0.195
150	0.00490±0.00001	0.201
200	0.00397±0.00001	0.202

---

*https://github.com/aurorateien/master-thesis/blob/main/AuCalibration.ipynb

distortion correction is shown in Figure 5.11. Figure 5.11(a) shows the raw 80 kV SAED ring pattern of Au obtained with nominal camera length 200 cm. A line scan of the pattern, illustrated by a green line, is plotted in Figure 5.11(b) showing the intensity as a function of reciprocal distance. Figure 5.11(c) shows the distortion corrected SAED pattern, with the corresponding line trace plotted in Figure 5.11(d). The planes corresponding to the ring patterns are marked with indices hkl for each intensity peak to which they give rise. Their radial position is marked on the dashed line corresponding to the peak. These positions are visibly different for the corrected and uncorrected image. The theoretical values for the Au reflections 111, 002, 022, and 113 are 0.425 Å^{-1}, 0.490 Å^{-1}, 0.693 Å^{-1}, and 0.813 Å^{-1} respectively. The peaks in the uncorrected image deviate, on average, 0.034 Å^{-1} from the theoretical values, while the corrected values only deviate with 0.007 Å^{-1}. Note that the scale is calibrated on the corrected pattern. The upper right and lower left quadrant are the ones most affected by elliptical distortions. Hence the line trace analysis is done for the upper right quadrant.
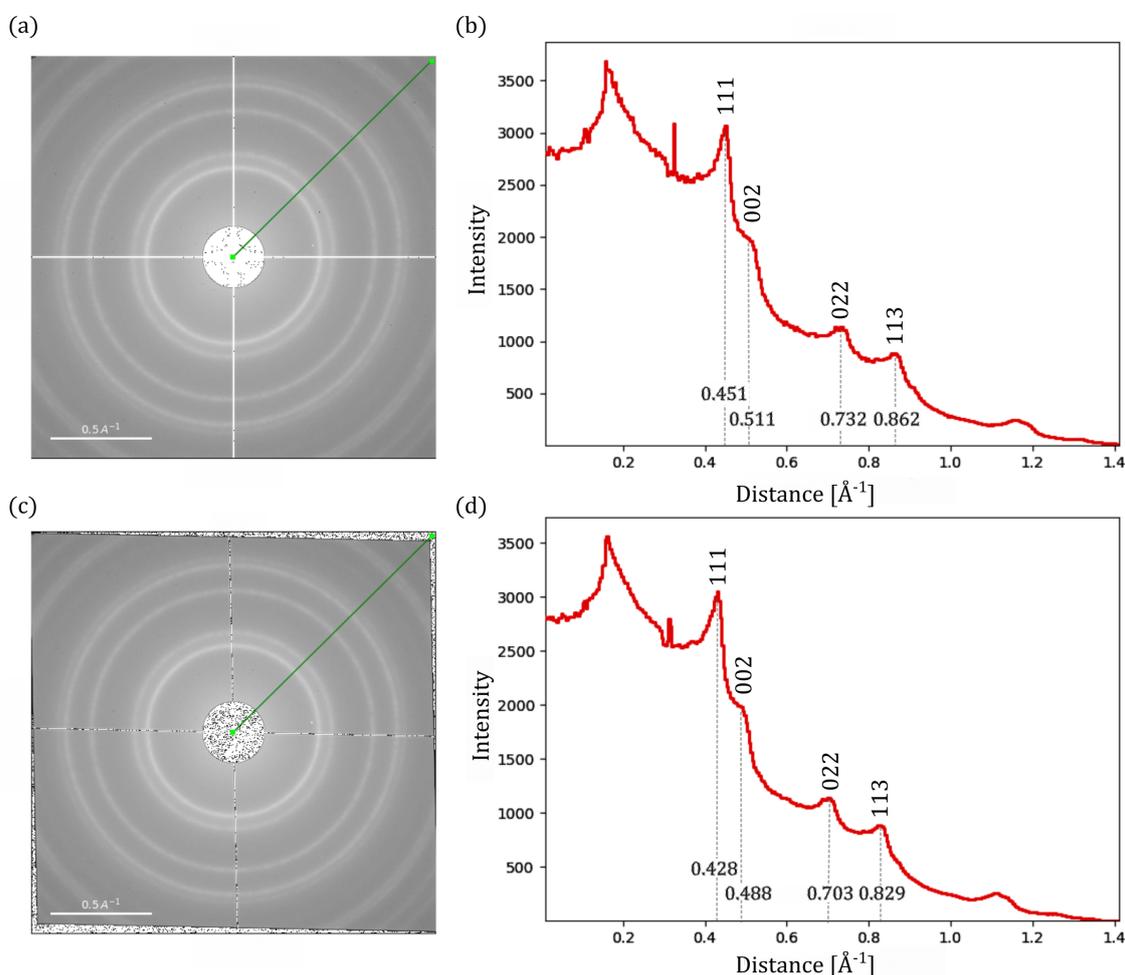
Figure 5.11: Au SAED pattern collected at 80 kV and camera length 200 cm. (a) Uncorrected raw image. (c) Elliptical distortion corrected image. Line traces are shown as green lines with corresponding intensity plots shown in (b) for raw image and (d) for corrected image.

The ellipticity in diffraction images can vary for each session they are collected. As many of the datasets used in this experiment were obtained in separate sessions, all images cannot be directly compared to each other unless the ellipticity in the image is negligible.

Removal of the untrusted rectangle

The untrusted rectangle, visible in Figure 5.12(a) as a bright white two-pixel wide cross separating the four quadrants of the detector, is removed by flatfield correction if applied. This rectangle usually has a considerable intensity, see Figure 5.7(a), which affects the total number of detected spots. For the earlier obtained datasets in this work, flatfield was not applied, hence it became necessary to find a way to remove this feature by other means. XDS already has built-in functions for ignoring this cross. REDp, however, does not. The number of detected peaks in the peak search step is directly affected by the removal of the cross, visible in 5.12(c) and (d). A simple preprocessing solution for removing the untrusted rectangle is presented in Listing 5.1. Here, two options are suggested for removing the bright cross, one where the intensity of the cross is set to 0 (black) and one where the the intensity is set to be the mean of the surrounding pixels.

```python
import numpy as np

# We want to make the white untrusted rectangle black.
def mark_cross_black(img, mask_value=0):
    img[255:257, :] = mask_value
    img[:, 255:257] = mask_value
    return img

# We want to make the untrusted rectangle a mean of the neighboring pixels
def mark_cross(img):
    img[255:257, :] = np.array(( 0.5*(img[254, :] + img[257, :]), 0.5*(img[254, :] + img[257, :])))
    img[:, 255:257] = np.array((0.5*(img[:,254] + img[:,257]), 0.5*(img[:,254] + img[:,257]))).T
    return img
```

Listing 5.1: Simple code for removal of untrusted rectangle.

The full script is available from Github[†]. Here, a full stack of frames is processed, based on the input folder containing all tiff-files. The stack of frames can further be converted to .mrc-format for REDp usage by using the REDpConverter in the cRED conversion script[‡] or the ContRotDataConverter [**]. Additionally, as illustrated with the red rings in the four frames of Figure 5.12, removal of the cross will help detecting peaks that are ignored because of the bright cross. The diffraction spot in the red circle is visible in all four frames, but is in Figure 5.12(c) wrongfully detected by the peak finder in REDp. The spot is ignored and the bright cross is rather mistaken for being a spot.

Note that the untrusted rectangle for our detector, Merlin 4R DED from Quantum Detectors, differs from the untrusted rectangle for other setups such as the one in Stockholm [65]. The pixels connecting the four modules in the Timepix detector are larger than the other pixels, hence the images are converted to a 516×516 array for further processing [109]. For the Merlin 4R detector, there are three pixels connecting the four modules that have the same size as the others, i.e. $55\,\mu m$.

5.3.3   Crystal structure determination tools

The tools for crystal structure determination, i.e. REDp and XDS, were more thoroughly analyzed in [1]. Following from this work, it was found that the peak search parameters in REDp depended on data collection parameters. Combining this knowledge with what

---

[†]https://github.com/aurorateien/master-thesis/blob/main/removeRectangle.ipynb
[‡]https://github.com/TEM-Gemini-Centre/cRED/blob/main/cRED_conversion.ipynb
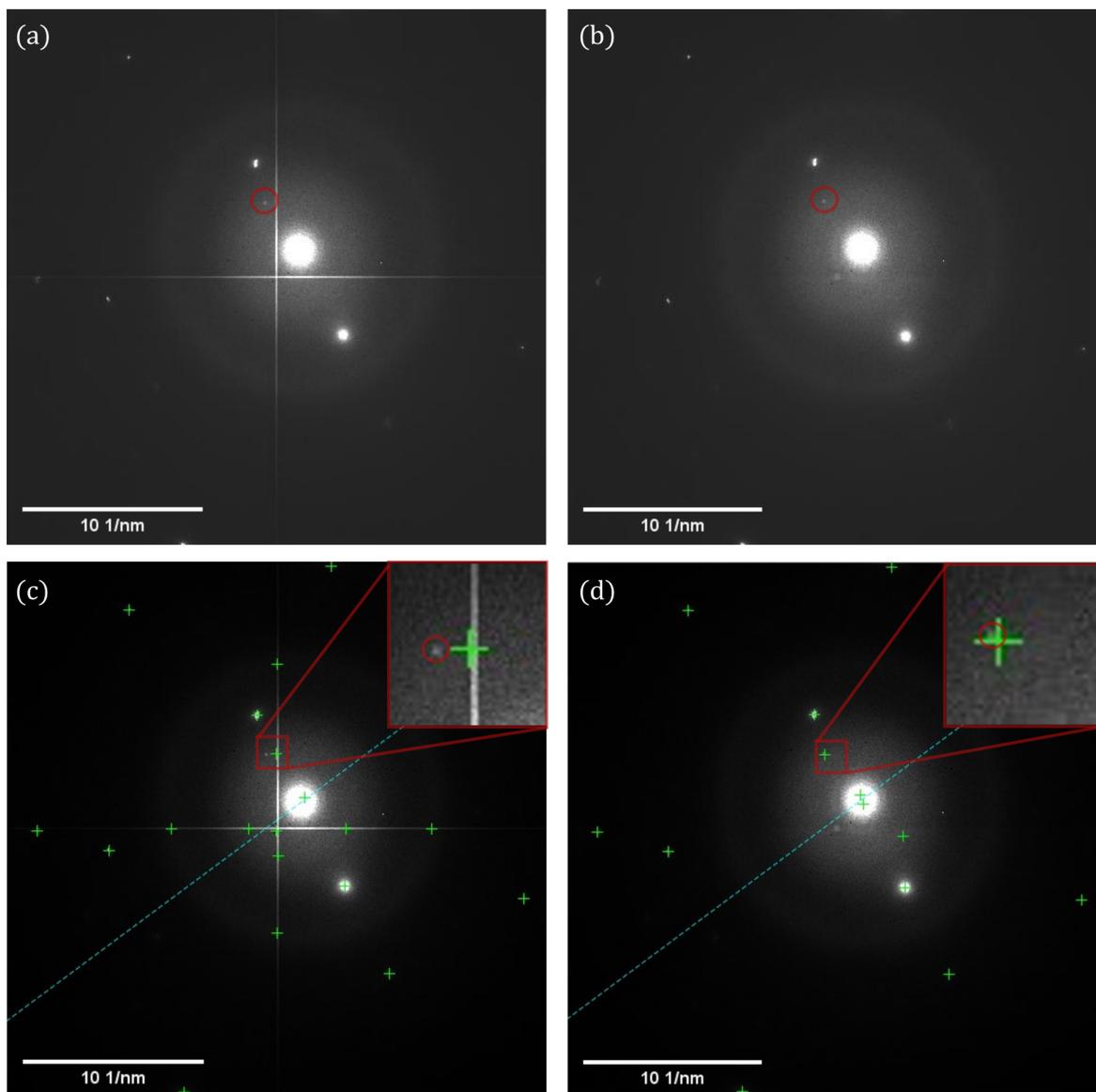[**]https://github.com/aurorateien/master-thesis/tree/main/ContRotDataConverter

Figure 5.12: Frame 13 of dataset Si5_10 (a) before and (b) after the untrusted rectangle was removed during preprocessing. The lower figures show the results of peak search in REDp (c) before and (d) after removal of the cross. The rotation axis is marked with a blue dashed line and found peaks with a bright green cross.

was found in this work, higher voltage would result in broader PSFs and more background, suggesting higher parameter values (more pixels) in the peak search step of REDp.

Going straight from raw/preprocessed data to XDS, i.e. skipping REDp, should in theory be possible. This would indeed result in one less program in the pipeline, which could be favorable to avoid confusion for new users. Instamatic already produces all necessary input files for XDS, independent of REDp. The information which is extracted from REDp, i.e. rotation axis, position of direct beam, and tentative space group, could be found by alternative approaches instead. Calibration of detector distance would still have to be done manually. The rotation axis can be found using the EDtools function `find_rotation_axis`. The position of the direct beam can be found by inspecting the frames in XDSGUI. The space group might be known from before, or some candidates may be known prior to the experiment. Anyway, the space group will roughly be found by XDS under the IDXREF tab (at least the class and centering) along with suggestions for lattice parameters after the initial run. The issue then becomes differentiating between the different space groups of similar symmetry without inspecting systematic absences of reflections. A possible approach would be looking at the output file FRAME.cbf to observe whether the space group chosen predicts the observed diffraction spots. Another possible method could be inserting the suggested lattice parameters and space group from IDXREF into the Computational Crystallography Toolbox (CCTBX) website[§] commonly used by REDp to find the space group of found lattice parameters. Redundant or not, REDp is still a useful tool for inspecting frames, finding systematic absences, in addition to providing easy and fast visualization of the 3D reconstructed reciprocal space.

Upon analysis in XDS, the data acquired at NTNU was deemed to be of poor quality. The program produced several different error messages depending on the input parameters used. Eventually, when setting certain parameters high, i.e. MAX_ERROR_IN_SPOT_ POSITION and MAX_ERROR_IN_SPINDLE_POSITION, the program ran properly, but produced poor resulting characteristics. The output `hkl`-file was deemed invalid. However, defying the poor characteristics, XPREP and SHELXT were still able to determine the right space group for some of the datasets (i.e. Si4_9 and GaSb1_1) using the invalid hkl-files.. Further refinement could produce acceptable R-values and a correct chemical structure, but would suffer from low equations (intensities) to variables ratio, as mentioned in Section 4.1.4, and large jumps in statistical characteristics between each refinement cycle.

5.3.4 XPREP vs EDtools

EDtools was shown to perform well as a generator for .ins-files. It generates SFAC values for electrons by default, which XPREP does not. XPREP differs from EDtools, in that it is based on the intensities given in the .hkl-file. EDtools runs completely on its own, separately from the hkl-file. The standard deviation of the lattice parameters are hence set to zero for EDtools-generated input-files. Another feature XPREP has, that EDtools do not, is merging. The effects of merging data with XPREP, however, were negligible. In some cases, not merging would lead to a more inaccurate initial model solved from SHELXT, but this model can be corrected in refinement. EDtools is less time consuming, as it only involves typing a one-liner of code. XPREP, on the other hand goes through several processing steps. It should be noted that XPREP has many additional features outside the scope of the present work.

In this experiment, an issue occurred when running EDtools for space group Fd3̄m (no. 227) for Si. SHELXT did not recognize the given symmetry elements in the .ins-file and

---

[§]https://cci.lbl.gov/cctbx/lattice_symmetry.html

crashed. The issue was simply solved by inputting space group "227:2" instead. This other setting ("227:2") is needed, as the first setting (i.e. "227") does not have the origin on a center of symmetry which is required for SHELXT. If in doubt of the space group, XPREP can come in handy as it can search for higher symmetries, as shown in Table 4.2. Some of the datasets used in this experiment were not able to use the "Determine space group" function in XPREP. In these cases, "Input space group" was used instead. However, this is not necessarily a common problem as our data did not function properly in XDS and thus provided invalid .hkl files.

EDtools is considered a good open-source alternative to the licensed XPREP.

5.3.5 Structure solution with SHELXT

SHELXT was found to be a powerful tool for crystal structure solution. The program posed no problems during the procedure, and was highly autonomous. If needed, it is possible, as mentioned in Section 2.3, to choose from many other algorithms for structure solution as well, such as SHELXS, which take use of other statistical methods for correcting for the phase problem. This was not included in the current experiment.

In Section 4.2.3, the effect of different input space groups in SHELXT is analyzed. From these results, it is suggested that multiple input space groups give the same solution. R_{int} was expected to act as a compass as to whether the space group is correct or not for the data, e.g. the lower R_{int} the better space group match, although this was shown not to be the case. R_{int} tended to just be lower for the lower symmetry space groups. This test proved that SHELXT will not blindly solve the space group in the input space group, but rather iterate through many different space groups (within the inputted Laue group unless stated otherwise) to find the best match for the data.

5.3.6 Refinement procedure

In Chapter 4, refinement observably made up a large part of the work needed for a final structure to be valid. A major problem with the refinement process described in this work, is that it requires both experience and detailed chemical knowledge of the structure in question. The refinement steps described in Section 4.1.4, cover some of the most important features of the refinement process, but not all. For example, many other commands for constraining and restraining the model, which were not necessary in the present study, do exist. The number of constraints and modifications needed for a properly refined structure also depends on the data quality. Too many restraints can cause validation bias, where the structural model is forced to become what is desired. In this work, getting a correct chemical structure was prioritized over achieving the lowest possible R-values.

In general, it can be confusing with all the different R-values presented, as the different software produces their own R-values. The general definition of an R-value was introduced in Section 2.3.1 Equation 2.29. The R-values from XDS for example, should not be directly compared to the ones produced by SHELXT and SHELXL because they have different definitions. Regarding the statistical characteristics produced during refinement, that are shown in colored boxes in the upper right panel in Olex2, it should be noted that for ED data it is usually not realizable for all boxes to turn green. For example, certain issues related to the very definition of the R-values make them take on grand values for ED. Some of the R-factors are based on the structure factor squared F^2, i.e., wR_2 and R_{int} see the definitions in Section 4.1.4, which can be a disadvantage as these take relatively larger values than the R-factors based on F. Since R_1 is the only of the quality parameters listed above that scale with F_o and not with the latter squared, this is the parameter that can take the

lowest values for high-variance data. However, the larger the refined parameters, the lower the R-value. Note that the experimental measurement of F_o^2 can also take negative values. This is because the background takes values higher than the peak as a result of statistical fluctuations. Additionally, if the data is slightly distorted, the shift can be greater than zero.

Both SHELXL and SHELXT were found to be easy to use inside the GUI of Olex2. Olex2 has an extensive GUI for 3D view of molecular structures, which was highly valuable when refining the structure and inspecting the electrostatic potential map.

5.3.7 Fileformat changes

An issue related to the ease of use of this data processing procedure, is the repeated switching between different file formats. Each program uses different input files, whereas some require both an instructional input file and data input in a different format, and produces output files of even more new formats. A full overview of the different file formats used throughout the procedure is shown in Table 5.2. A reduction of the number of file format changes and

Table 5.2: Input and output file formats corresponding to the programs used.

Program	Instructional input	Data input	Output
REDp	.3ded	.mrc	.img
XDS	.INP	.img	.hkl, .lp, .cbf
XPREP		.hkl	.ins, .hkl, .prp
SHELXT	.ins	.hkl	.res, .fcf, .ins, .lxt
SHELXL	.ins	.hkl	.ins, .lst, .cif

number of switches between programs may enhance the interoperability and simplify the use for new users. By switching to a broader use of versatile Python libraries such as EDtools and `HyperSpy`, one could possibly be going directly from Instamatic-output files `.img`, `.tiff` or `.mrc` to `.hkl` and `.ins`. As mentioned previously in Section 5.3.4, using EDtools instead of XPREP is possible. In Section 5.3.3, the use of REDp is also debated. This would reduce the amount of file changes.

5.3.8 Merging data

Using XSCALE is a rather simple method for merging data. The reflections (`.hkl`) are merged, by adding all reflections to the same `.hkl` file. Evidently, for the merged dataset from Section 4.2.2 of MOR1 and MOR2, 1521 of the reflections are common for both datasets. The resulting `.hkl` file will then contain twice the amount of these *hkl* reflections, with corresponding scaled intensities from the two constituent datasets. Table 4.6 shows the well-known trade-off between completeness and R-values. A completeness of 100% is usually not needed, but a higher completeness is desired primarily to solve the structure correctly. The completeness of the merged dataset was indeed higher than for the separate input datasets. However, more reflections will give higher R-values as there statistically will be more outliers deviating from the calculated structure factors. The merged R-values in Table 4.6 are in between the R-values of the separate input datasets, although for both R_1 and wR_2 the merged value is closer to the lowest separate value.

From these observations, it can be assumed that merging of datasets in general results in higher completeness. However, the R-values are averaged between the datasets and can be worse than some of the better input datasets.

5.3.9   Using XRD tools for ED

Several XRD-specific tools are used in the current procedure; XPREP, XDS, SHELXT, SHELXL, and Olex2. Section 4.1.2 describes how XPREP by default assumes it is dealing with XRD data and how the user manually has to change the wavelength and scattering factors to apply to electrons instead. XDS and SHELXT pose no specific issues when handling ED, except worsened statistical characteristics, such as the R-values mentioned in Section 5.3.6. However, SHELXL and the default refinement program in Olex2, olex2.refine, both assume kinematic diffraction [110]. Additionally, SHELXL will automatically reject systematically absent reflections [27]. Hence, it can be favorable to use unmerged data, i.e. where Friedel-opposites* are not merged, for enhanced data completeness statistics in the `.cif`-file. Dynamical approaches to ED-data is further discussed in Section 5.3.9.

To clarify, for electrons we often refer to the electrostatic potential map, rather than the electron density map which is used for X-rays. This is due to the electron being a charged particle, whereas the entire charge distribution including nuclear and electron charges contribute to the energy landscape along with the electron density distribution. Hence, electrostatic potential maps will have to be interpreted differently than pure electron density maps. Electron density is the probability distribution of electron presence in a volume element, a map over this density will hence never be negative. The electrostatic potential, however, may have minima with negative values if the region is attractive for positive charged atoms [112]. In refinement, an electrostatic potential map can at first glance seem more noisy and thus be more difficult to interpret. However, possibilities for seeing weakly scattering elements, like hydrogen, open up [42].

Dynamical analysis

The kinematical approximation, as described in Section 2.2.4, assumes electrons are scattered only once and intensities are proportional to the structure factor amplitude squared. Electron diffraction causes dynamical effects that, in addition to inelastic scattering and defects, result in larger deviations from the kinematical assumptions [113]. The R-factors become worse and the quality of the structural model decreases. Dynamical theory of electron diffraction non-linear deviations from kinematical theory.

Dynamical refinement routines which can be applied to already established experimental 3D ED routines have been developed [42]. Results from these studies show improved fits between experimental data and calculated models, improved accuracy of atomic position refinement, and considerable reduction of noise in electrostatic potential maps. This can be of great importance for discovering new structures and revealing absolute configurations of complex pharmaceutical compounds.

For further analysis of zeolites, with guest molecules in pores, an analysis accounting for dynamical effects is preferred. As dynamical refinement reduces the noise level in potential difference maps, weak structural features will be easier to observe, making determination of guest molecule location in MOFs and zeolites more precise. Furthermore, it will open the possibility for hydrogen atom positions to be revealed [42]. A dynamical analysis can however, be a time consuming task and is outside the scope of this work.

---

*Friedel's law applies only to centro-symmetric unit cells. This law states that intensities of "opposite" indices, like (h, k ,l) and (-h, -k, -l) often referred to as Friedel pair, are about equal. This can cause the real crystal to have lower symmetry than the result of a diffraction experiment would imply [111].

Alternative workflows

Other workflows for 3D ED data processing have been developed in different research groups around the world. As alternatives to REDp, other dedicated electron diffraction software that can be used are eADT or ADT3D developed by Kolb et al. [12], or PETS2 developed by Palatinus et al. [14]. The differences between PETS2 and the procedure from this work, include that PETS2 is only available for Windows operational system and uses .tiff files as input. Additionally, PETS2 has an automatic rotation axis finder and ED specific distortion correction [14]. The largest procedural difference however, is the dynamical refinement. Indeed, PETS2 has embedded the program JANA2020 that does dynamical refinement.

However, most 3D ED processing workflows still rely on XRD specific software. XDS is used in the workflow studied here, although many other programs with similar functionality are available as well, like DIALS [24], CrysAlis^{Pro} (by Rigaku)[18], Apex (by Bruker), and MOSFLM [25]. The reason why there are so many options for processing XRD data, is that XRD is a well established and widely used technique. The developers of DIALS are currently optimizing their program for electron diffraction. Note that the currently used version of Instamatic actually produces input files for DIALS and PETS2 in addition to the XDS and REDp-files used in the present work [22].

EDtools has in this project mostly been used for automated and batch data processing. However, this processing is still based on XDS, EDtools just runs XDS internally. Another alternative for automated 3D ED data processing is the Python-based library Scipion-ED [72]. Scipion-ED is based on the already existing open source, Python-based library Scipion, and DIALS. As of 2022, Scipion-ED does not have DIALS embedded, but rather calls DIALS through a shell similarly to what EDtools does with XDS. At the time of writing, there are no existing dedicated Python library for full crystal structure solution of electron diffraction data. However, both EDtools and Scipion-ED cover parts of the procedure.

PLATON is an additional software that is partly embedded in Olex2. This is a tool for analysis, structure validation and visualization of SCXRD data [114, 115]. A big difference with PLATON compared to e.g. the SHELX packages is that it analyzes the structure (`.cif` file), not the intensities (`.hkl` file). PLATON can be used for example if one suspects the refined structure is solved in a space group of too low symmetry. The program will then analyze the structure and find higher symmetries. Furthermore, PLATON will be able to transform the unit cell to the new symmetry and generate a new `.hkl`-file within the new space group.

The data collection method can also be changed if favorable. Other protocols such as serial electron diffraction (serialED) and stepwise rotation electron diffraction (RED) could be used. Especially serialED have been predicted to have a lot of potential and is under constant development. This method tracks multiple particles, but only small anular ranges from each particle, and later combines these patterns to form a stack of frames. This method is useful for beam sensitive materials as it reduces the dose on each particle and thus avoid amorphisation. In other experiments [65], serialED is used only for finding suited particles on the grid and then cRED is applied for the rest of the data collection procedure.

To summarize, this section has discussed the data processing procedure, as described in Chapter 4. For the dataset requirements, it is found that 3D ED data stacks should have minimum 21 frames and every other frame can maximum be excluded for dataset characteristics to be sufficient. The data obtained at NTNU suffer from slight elliptical distortions that should be corrected for in preprocessing. Potential short-cuts for the data processing workflow are suggested, favoring an increasing use of Python based software. Refinement is found to be the most rigorous and important part of the procedure. The tools

for structure solution and refinement, SHELXT and SHELXL, were easy to use, especially when run in Olex2 GUI, and effective, despite kinematical assumptions. Furthermore, the differences between XRD and ED have been addressed, whereas a short review of how dynamical analysis can be a solution to the problems arising from these kinematical (XRD) assumptions was assessed. Additional tools were introduced as alternatives to the current workflow.

Chapter 6

Future work

As become clear from the discussion in Chapter 5, work remains in order to fully establish 3D ED at NTNU. As of now, a manual for the full data processing procedure exists, which documents the established routines and should give more efficient on-boarding of new users. This work has illustrated the importance of correct data collection in order for the data processing method to function properly. Preferably, a manual for data collection should also be developed. The new users would then avoid the issues that have been discussed in this work, such as saturation of pixels, distortion, and timing issues. If further enhancements are made to the data processing method, the data processing manual would have to be updated. Changes in data processing development of this field are expected in the future, for example more open-source python packages as has happened for other TEM techniques (HyperSpy, Pyxem [107]). Thus, the manual should be constantly updated on the trends in the field of 3D ED data processing.

When this complete data collection and processing procedure is established at NTNU, a sea of opportunities opens up. An example of usage is to further investigate the zeolites used in this experiment and look for differences in lattice parameters and guest molecules in these samples. The guest molecule analysis is done in refinement. First, the framework is solved, and then the pores of the framework can be inspected along with the electron potential map. Positive ions, that is, K^+, Na^+, Cu^{+2}, and even H^+, should be possible to detect the position of from the electrostatic potential map. Inspecting H+ may be the most challenging, as the signal is weak and require dynamical refinements. In the current routine, based on X-rays, dynamical refinement play a minor role and alternative routes as PETS2 [14] might need to be considered.

In this work, diffraction spots are found through peak finding in both REDp and XDS. However, both of these peak finding methods rely on input parameters such as background pixels, peak to background threshold, and (for XDS) strong pixels and minimum number of pixels in a spot. If these are not correct, a varying number of diffraction spots can be detected. This input parameter requirement can be unfavorable for automation of the method. As an alternative to these peak finding algorithms, a neural network (NN) can be trained and applied. However, this will require knowing the peak positions in patterns accurately, which previously has been demonstrated for 4DSTEM strain analysis [116]. With machine learning (ML) and artificial intelligence being hot topics, it is remarkable that it is not yet utilized for 3D ED*. ML have been utilized for other structure analysis techniques for analyzing diffraction patterns [117, 118]. Using more ML approaches could have potential for decreasing the number of programs in the workflow of the method.

One of the main advantages of 3D ED is that it is a low dose technique [32]. Further

---

*In Instamatic [22], neural networks are in fact embedded in the program.

studies of beam sensitive materials, e.g. zeolites or MOFs, may require a more quantitative analysis of the dose applied onto the material. MOFs are materials based on zeolittic structural principles, where metal ions are inserted to alter properties. These are generally more unstable than zeolites. Like zeolites, the MOF properties are correlated with the structural composition. Plant-based MOFs have been used for the removal of pharmaceutical compounds in wastewater [119]. In recent years, an exponentially growing research interest in MOFs for use in waste water treatment has been observed [120]. Physical quantities like dose (C/nm^2) or dose rate (C/nm^2s) can be introduced, alternatively dose per frame to predict how many frames can be obtained before the material is damaged. Dose differences between different variations of the same materials could be interesting to look at. An initial practical experiment can involve determining the beam current in the ARM, using a Faraday cup [40], or to do absolute counting using the DED so that the dose or dose rate in a tilt series can be determined and balanced against other parameters, as critical dose for a material and sufficient signal-to-noise-ratio in the frames.

Changes in the software used could be favorable for future data processing. EDtools and Instamatic v2.0.0, for example, both being developed by Stockholm University [22], always assume 200 kV electrons. In this experiment, 80 kV is used for most datasets, which require manual change of wavelength parameters during `.ins`-file preparation and in XDS-input file preparation. If all 3D ED experiments at NTNU in the future will be obtained at 80 kV or other wavelengths that deviate from the Stockholm default of 200 kV, it is suggested that the Instamatic code is changed to read out the voltage/wavelength of electrons directly from the experiment. This suggestion would simply save the data analyzer some time. Note that there is a newer version of Instamatic available, i.e. v2.0.1, that might have solved this issue. The original developer of Instamatic, Stef Smeets, helped in 2023 with the integration of the Merlin DED into Instamatic that made 3D ED more practical for the NTNU setup.

For the experimental setup, several features need further investigation as became clear from Chapter 5. Specific experiments involving comparison and mapping out the tilt step homogeneity of holders are suggested. Differences in practical rotation range, stability, and ease of use could be parameters to track. Another possible comparison study is to observe the effect of different substrates on the background. In the current study standard C-supports were used, but for low-dose experiments with poor scatters, thinner C-support might make a difference. In practice, datasets with different substrates have already been used in this work, although too many other parameters have been varied to study the effect of different substrates specifically. Holey carbon films were observed to enhance uniformity in background for this experiment, which could be further investigated by comparing it to other films or varying the holes/mesh size and form of the carbon film. The stability of the goniometer should be tested by plotting the gonio response as a function of the input tilt. Within the TEM group at NTNU, well established template based orientation analysis might be useful for this purpose [121]. A well-known material can be used and the indexing can be done for each tilt. Thereby the trace in orientation space can be made and the misorientation between individual frames can be analyzed. An assumption in the current processing is that the steps are equal, but according to the issues seen in Chapter 5, this has to be experimentally verified. In general, it could be favorable for Instamatic to read out the actual tilt values for each tilt instead of the primitive oscillation range calculation that only tracks start tilt and stop tilt. Stepwise tilt is a common parameter to track for other TEM experiments, e.g. when 3D ED was performed initially without Instamatic, tilt was tracked. However, it is assumed that the primitive oscillation range calculation is done in order to fit the XRD-specific programs XDS (and DIALS) that do not work for irregular tilt steps. Fewer programs in the pipeline and enhanced interoperability is desired for the

data processing workflow. The additional timing problem, related to the acquisition time being much higher ($\sim$300 ms) than the exposure time, should be investigated further. This problem can potentially be connected to the large number of defocused frames.

Going further, the effect of elliptical distortion correction should be tested by running a distortion corrected dataset through the entire data processing procedure. It would be favorable to see whether or not it has an effect on any of the pipelined programs, for example enhanced R-values or completeness. Another type of distortion to consider is the potential distortion caused by the image corrector.

SerialED was described in Section 5.3.9 as a useful data collection tool for finding suitable particles. In previous work [1], the tedious procedure of finding suited particles was addressed. Using serialED to find suitable particles for cRED might solve this issue. This would require switching mode in Instamatic (to serialED) and familiarizing oneself with the required knowledge for serialED data collection. Stockholm University knows this technique well and could be expected to provide help for setting this up. serialED is under constant development, and the interpretation of serialED data is expected to improve significantly in the coming years.

In general, take serialED as an example, a trend towards more automation is observed in the field of 3D ED and in general for TEM, which would favor fewer programs (and changes in fileformats as mentioned in Section 5.2). For commercial dedicated systems as Eldico [17], Tescan Tensor, and Rigaku [18], this streamlining is achieved. Less black-box software and more open source code would be preferable for a broader understanding of the whole process. Developing a Python-based library for full structure analysis could be useful in the future. As TEM oriented packages as HyperSpy and pyXem demonstrate [107], contributions can be made from the community, largely by junior academics, PhDs and master students. In this way, many people can contribute to a larger common project.

Chapter 7

Conclusions

The objectives of this project were to establish a data processing procedure for 3D ED crystal structure analysis and optimizing the data acquisition process. During the project period, the workflow was successfully tested by applying it to zeolite reference datasets from Stockholm University, and a manual covering this data processing procedure was then developed. The 3D ED data acquisition procedure with associated parameters was optimized by applying the successful data processing procedure to 3D ED data obtained with the experimental setup at NTNU.

All programs in the data processing workflow, i.e. REDp, XDS, XPREP, EDtools, SHELXT, SHELXL and Olex2, have been thoroughly examined. Bottlenecks in the workflow have been investigated, and potential solutions for an improved data processing routine have been suggested. Considerable elliptical distortion was observed and recommended corrections for this with e.g. Python preprocessing is recommended prior to further data processing. The fact that XPREP is licensed was found to be impractical, hence EDtools was explored as an open-source alternative for SHELXT-input file preparation. For this purpose, and for additional features, such as merging of data, EDtools was found to perform well. The classic XRD tools for structure solution and refinement, i.e. SHELXT and SHELXL, were simple to use and effective. Kinematical assumptions were sufficient for this work, however, dynamical analyses might be necessary when inspecting structures requiring finer detail. Refinement of the crystal structure is found to be the most demanding part of the procedure, depending on the data quality and complexity of the structure in question.

The experimental cRED data obtained with the NTNU setup, i.e. a Jeol JEM ARM200F with a 4R Merlin EM DED detector and Instamatic software, were observed to be problematic. Hence, different data collection parameters were tested, e.g. acceleration voltage and choice of condenser aperture and spot size. Problems with saturated pixels were solved by changing to a higher detector bit depth. Permanent background noise from the untrusted rectangle was removed by applying flatfield correction. The lower acceleration voltage (80 kV) was deemed favorable due to sharper PSFs and less background than higher voltages (200 kV). A smaller radius spot size was observed to reduce the intensity in frames and thereby the dose, which could be favorable to avoid beam damage.

Beside optimizing the experimental data collection parameters, the written user manual will hopefully also proof valuable. By going through the whole solution and refinement procedure, it was demonstrated that implementing the technique on an existing TEM requires correlating the data acquisition and processing procedure. Learning from these findings, concrete issues that had to be tackled were compromised intensities (saturation) and intensity reduction for minimizing dose. The problems related to timing and goniometer stability remain unsolved and need to be overcome for 3D ED structure analysis to be fully functional

at NTNU. Future experiments can perform data analysis that includes dynamical diffraction refinement on more advanced structures, e.g. MOF and COF, to investigate guest molecules in the porous structures. A more quantitative analysis for dose determination should be conducted for further work with beam sensitive materials.

To conclude, 3D ED is a technique that is expected to have considerable potential for further use at NTNU, especially now that a well-functioning data processing procedure is established, although work still remains to solve the issues with the data acquisition procedure identified in this work.

Bibliography

[1] A. Teien, "3D electron diffraction tomography: From electron diffraction tomo datasets to detailed structural information." TFY4520 Nanotechnology Specialization Project, Department of Physics, NTNU, 2023.

[2] S. Smeets and M. O. Cichocka, "Continuous rotation electron diffraction data for Zeolite Mordenite," 2018.

[3] C. Kittel and P. McEuen, *Introduction to Solid State Physics*. Wiley, 2018.

[4] C. Giacovazzo, C. Giacovazzo, H. L. Monaco, G. Artioli, D. Viterbo, M. Milanesio, G. Gilli, P. Gilli, G. Zanotti, G. Ferraris, and M. Catti, *Fundamentals of Crystallography*. Oxford University Press, 2011.

[5] A. West, *Solid State Chemistry and its Applications*. Wiley, 2014.

[6] M. De Graef, *Introduction to Conventional Transmission Electron Microscopy*. Cambridge University Press, 2003.

[7] P. Misra, *Physics of Condensed Matter*. Elsevier Science, 2011.

[8] A. J. C. Wilson and V. Geist, "International tables for crystallography. volume C: Mathematical, physical and chemical tables.," *Dordrecht ; Boston ; London : Published for the International Union of Crystallography by Kluwer Academic Publishers*, vol. 28, no. 1, pp. 110–110, 1993.

[9] Z. Huang, T. Willhammar, and X. Zou, "Three-dimensional electron diffraction for porous crystalline materials: structural determination and beyond," *Chemical Science*, vol. 12, 2020.

[10] M. Gemmi, E. Mugnaioli, T. E. Gorelik, U. Kolb, L. Palatinus, P. Boullay, S. Hovmöller, and J. P. Abrahams, "3D electron diffraction: The nanocrystallography revolution," *ACS Central Science*, vol. 5, pp. 1315–1329, 2019. PMID: 31482114.

[11] B. L. Nannenga and T. Gonen, "The cryo-em method microcrystal electron diffraction (microed)," *Nature Methods*, vol. 16, pp. 369–379, 2019.

[12] U. Kolb, Y. Krysiak, and S. Plana-Ruiz, "Automated electron diffraction tomography – development and applications," *Acta Crystallographica Section B*, vol. 75, no. 4, pp. 463–474, 2019.

[13] E. Mugnaioli, T. Gorelik, and U. Kolb, "Ab initio structure solution from electron diffraction data obtained by a combination of automated diffraction tomography and precession technique," *Ultramicroscopy*, vol. 109, no. 6, pp. 758–765, 2009.

[14] L. Palatinus, P. Brázda, M. Jelínek, J. Hrdá, G. Steciuk, and M. Klementová, "Specifics of the data processing of precession electron diffraction tomography data and their implementation in the program *PETS2.0*," *Acta Crystallographica Section B*, vol. 75, no. 4, pp. 512–522, 2019.

[15] L. Palatinus, P. Brázda, P. Boullay, O. Perez, M. Klementová, S. Petit, V. Eigner, M. Zaarour, and S. Mintova, "Hydrogen positions in single nanocrystals revealed by electron diffraction," *Science*, vol. 355, no. 6321, pp. 166–169, 2017.

[16] R. F. Service, "Molecular ct scan could speed drug discovery," *Science*, vol. 362, no. 6413, pp. 389–389, 2018.

[17] P. Simoncic, E. Romeijn, E. Hovestreydt, G. Steinfeld, G. Santiso-Quiñones, and J. Merkelbach, "Electron crystallography and dedicated electron-diffraction instrumentation," *Acta Crystallographica Section E*, vol. 79, no. 5, pp. 410–422, 2023.

[18] S. Ito, F. J. White, E. Okunishi, Y. Aoyama, A. Yamano, H. Sato, J. D. Ferrara, M. Jasnowski, and M. Meyer, "Structure determination of small molecule compounds by an electron diffractometer for 3D ed/microed," *CrystEngComm*, vol. 23, pp. 8622–8630, 2021.

[19] P. S. Shen, "The 2017 Nobel Prize in chemistry: cryo-EM comes of age," *Anal Bioanal Chem*, vol. 410, no. 8, pp. 2053–2057, 2018.

[20] X. Zou and S. Hovmöller, "Electron crystallography: imaging and single-crystal diffraction from powders," *Acta Crystallographica Section A*, vol. 64, no. 1, pp. 149–160, 2008.

[21] D. Zhang, P. Oleynikov, S. Hovmöller, and X. Zou, "Collecting 3D electron diffraction data by the rotation method," *Zeitschrift für Kristallographie*, vol. 225, no. 2-3, pp. 94–102, 2010.

[22] S. Smeets, B. Wang, and E. Hogenbirk, "instamatic-dev/instamatic: 1.7.0," 2021.

[23] T. Yang, T. Willhammar, H. Xu, X. Zou, and Z. Huang, "Single-crystal structure determination of nanosized metal-organic frameworks by three-dimensional electron diffraction," *Nature protocols*, vol. 17, no. 10, p. 2389—2413, 2022.

[24] G. Winter, J. Beilsten-Edmands, N. Devenish, M. Gerstel, R. J. Gildea, D. McDonagh, E. Pascal, D. G. Waterman, B. H. Williams, and G. Evans, "Dials as a toolkit," *Protein Science*, vol. 31, no. 1, pp. 232–250, 2022.

[25] A. G. W. Leslie, "The integration of macromolecular diffraction data," *Acta Crystallographica Section D*, vol. 62, no. 1, pp. 48–57, 2006.

[26] A. Saha, S. S. Nia, and J. A. Rodríguez, "Electron diffraction of 3D molecular crystals," *Chemical Reviews*, vol. 122, pp. 13883–13914, 2022.

[27] G. M. Sheldrick, "Crystal structure refinement with *SHELXL*," *Acta Crystallographica Section C*, vol. 71, no. 1, pp. 3–8, 2015.

[28] G. T. Kadja, N. T. Culsum, and R. M. Putri, "Recent advances in the utilization of zeolite-based materials for controlled drug delivery," *Results in Chemistry*, vol. 5, p. 100910, 2023.

[29] J. Garcia-Martinez, C. Xiao, K. A. Cychosz, K. Li, W. Wan, X. Zou, and M. Thommes, "Evidence of intracrystalline mesostructured porosity in zeolites by advanced gas sorption, electron tomography and rotation electron diffraction," *ChemCatChem*, vol. 6, no. 11, pp. 3110–3115, 2014.

[30] J. Zečević, A. M. J. van der Eerden, H. Friedrich, P. E. de Jongh, and K. P. de Jong, "Heterogeneities of the nanostructure of platinum/zeolite y catalysts revealed by electron tomography," *ACS Nano*, vol. 7, pp. 3698–3705, 2013.

[31] V. Girelli Consolaro, V. Rouchon, and O. Ersen, "Electron beam damages in zeolites: A review," *Microporous and Mesoporous Materials*, vol. 364, p. 112835, 2024.

[32] Y. Li, Y. Ling, and Y. Ma, "Application of three-dimensional electron diffraction in structure determination of zeolites," *Chinese Journal of Structural Chemistry*, vol. 43, no. 4, p. 100237, 2024.

[33] Y. Seto and M. Ohtsuka, "ReciPro: free and open-source multipurpose crystallographic software integrating a crystal model database and viewer, diffraction and microscopy simulators, and diffraction data analysis tools," *Journal of Applied Crystallography*, vol. 55, no. 2, pp. 397–410, 2022.

[34] X. Li, "Computer programs for unit-cell determination in electron diffraction experiments," *Ultramicroscopy*, vol. 102, no. 4, pp. 269–277, 2005.

[35] D. Schwarzenbach, *Crystallography*. Wiley, 1996.

[36] Z. Dauter and M. Jaskolski, "How to read (and understand) Volume A of *International Tables for Crystallography*: an introduction for nonspecialists," *Journal of Applied Crystallography*, vol. 43, no. 5 Part 2, pp. 1150–1171, 2010.

[37] IUCr, *International Tables for Crystallography, Volume A: Space Group Symmetry*. International Tables for Crystallography, Dordrecht, Boston, London: Kluwer Academic Publishers, 5. revised ed., 2002.

[38] J. C. Cockcroft, "High-resolution space group diagrams and tables," 1999. [Accessed 2024-03-10].

[39] J. S. Nilsen, *Advanced Transmission Electron Microscopy of III-V Nanowires*. PhD thesis, NTNU, 2020.

[40] D. B. Williams and C. B. Carter, *Scattering and Diffraction*, pp. 23–38. Boston, MA: Springer US, 2009.

[41] W. H. S. Bragg and W. L. S. Bragg, "The reflection of x-rays by crystals," *Nature*, vol. 91, pp. 477–477, 1913.

[42] P. B. Klar, Y. Krysiak, H. Xu, G. Steciuk, J. Cho, X. Zou, and L. Palatinus, "Accurate structure models and absolute configuration determination using dynamical effects in continuous-rotation 3D electron diffraction data," *Nat. Chem.*, vol. 15, no. 6, pp. 848–855, 2023.

[43] G. L. Taylor, "Introduction to phasing," *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 4, pp. 325–338, 2010.

[44] P. A. Karplus and K. Diederichs, "Linking crystallographic model and data quality," *Science*, vol. 336, no. 6084, pp. 1030–1033, 2012.

[45] L. Palatinus, "The charge-flipping algorithm in crystallography," *Acta Crystallographica Section B*, vol. 69, no. 1, pp. 1–16, 2013.

[46] G. M. Sheldrick, "*SHELXT* – Integrated space-group and crystal-structure determination," *Acta Crystallographica Section A*, vol. 71, no. 1, pp. 3–8, 2015.

[47] P. Evans and A. McCoy, "An introduction to molecular replacement," *Acta Crystallogr D Biol Crystallogr*, vol. 64, no. Pt 1, pp. 1–10, 2007.

[48] G. Oszlányi and A. Süto, "*Ab initio* structure solution by charge flipping," *Acta Crystallographica Section A*, vol. 60, no. 2, pp. 134–141, 2004.

[49] E. Kreyszig, *Advanced engineering mathematics 9th edition*. Wiley India Pvt. Limited, 2011.

[50] G. H. Michler and W. Lebek, *Electron Microscopy of Polymers*, ch. 3, pp. 37–53. John Wiley & Sons, Ltd, 2016.

[51] J. Zuo and J. Spence, *Advanced Transmission Electron Microscopy: Imaging and Diffraction in Nanoscience*. Springer New York, 2016.

[52] L. Reimer and H. Kohl, *Transmission Electron Microscopy: Physics of Image Formation*. Springer Series in Optical Sciences, Springer New York, 2008.

[53] B. D. A. Levin, "Direct detectors and their applications in electron microscopy for materials science," *Journal of Physics: Materials*, vol. 4, no. 4, p. 042005, 2021.

[54] R. Egerton and M. Watanabe, "Spatial resolution in transmission electron microscopy," *Micron*, vol. 160, p. 103304, 2022.

[55] R. Egerton, "Dose measurement in the TEM and STEM," *Ultramicroscopy*, vol. 229, p. 113363, 2021.

[56] R. Egerton, "Radiation damage to organic and inorganic specimens in the tem," *Micron*, vol. 119, pp. 72–87, 2019.

[57] R. Henderson, "From electron crystallography to single particle cryoem (nobel lecture)," *Angewandte Chemie International Edition*, vol. 57, no. 34, pp. 10804–10825, 2018.

[58] R. Vincent and P. Midgley, "Double conical beam-rocking system for measurement of integrated electron diffraction intensities," *Ultramicroscopy*, vol. 53, no. 3, pp. 271–282, 1994.

[59] A. S. Eggeman and P. A. Midgley, "Chapter 1 - precession electron diffraction," in *Advances in Imaging and Electron Physics* (P. W. Hawkes, ed.), vol. 170 of *Advances in Imaging and Electron Physics*, pp. 1–63, Elsevier, 2012.

[60] T. E. Weirich, R. Ramlau, A. Simon, S. Hovmöller, and X. Zou, "A crystal structure determined with 0.02 å accuracy by electron microscopy," *Nature*, vol. 382, pp. 144–146, 1996.

[61] L. J. Clark, G. Bu, B. L. Nannenga, and T. Gonen, "Microed for the study of protein–ligand interactions and the potential for drug discovery," *Nature Reviews Chemistry*, vol. 5, pp. 853–858, 2021.

[62] A. A. Bardin, A. Haymaker, F. Banihashemi, J. Y. Lin, M. W. Martynowycz, and B. L. Nannenga, "Focused ion beam milling and microed structure determination of metal-organic framework crystals," *Ultramicroscopy*, vol. 257, p. 113905, 2024.

[63] U. Kolb, T. Gorelik, C. Kübel, M. Otten, and D. Hubert, "Towards automated diffraction tomography: Part i - data acquisition," *Ultramicroscopy*, vol. 107, pp. 507–13, 2007.

[64] M. Gemmi, M. G. I. La Placa, A. S. Galanis, E. F. Rauch, and S. Nicolopoulos, "Fast electron diffraction tomography," *Journal of Applied Crystallography*, vol. 48, no. 3, pp. 718–727, 2015.

[65] M. O. Cichocka, J. Ångström, B. Wang, X. Zou, and S. Smeets, "High-throughput continuous rotation electron diffraction data acquisition *via* software automation," *Journal of Applied Crystallography*, vol. 51, no. 6, pp. 1652–1661, 2018.

[66] W. Kabsch, "Automatic indexing of rotation diffraction patterns," *Journal of Applied Crystallography*, vol. 21, no. 1, pp. 67–72, 1988.

[67] W. Wan, J. Sun, J. Su, S. Hovmöller, and X. Zou, "Rotation electron diffraction (red) software," 2013. https://doi.org/10.5281/zenodo.2545322 [Accessed 2024-02-20].

[68] E. van Genderen, M. T. B. Clabbers, P. P. Das, A. Stewart, I. Nederlof, K. C. Barentsen, Q. Portillo, N. S. Pannu, S. Nicolopoulos, T. Gruene, and J. P. Abrahams, "*Ab initio* structure determination of nanocrystals of organic pharmaceutical compounds by electron diffraction at room temperature using a Timepix quantum area direct electron detector," *Acta Crystallographica Section A*, vol. 72, no. 2, pp. 236–242, 2016.

[69] T. Matsumoto, A. Yamano, T. Sato, J. Ferrara, F. White, and M. Meyer, ""What is This?" A structure analysis tool for rapid and automated solution of small molecule structures," *Journal of Chemical Crystallography*, vol. 51, pp. 1–13, 2021.

[70] "Tescan tensor," 2024. https://info.tescan.com/stem [Accessed 2024-05-20].

[71] "Digistar-3," 2022. https://nanomegas.com/digistar-3/ [Accessed 2024-05-23].

[72] V. E. G. Bengtsson, L. Pacoste, J. M. de la Rosa-Trevin, G. Hofer, X. Zou, and H. Xu, "*Scipion-ED*: a graphical user interface for batch processing and analysis of 3D ED/MicroED data," *Journal of Applied Crystallography*, vol. 55, no. 3, pp. 638–646, 2022.

[73] A. Mastinu, A. Kumar, G. Maccarinelli, S. A. Bonini, M. Premoli, F. Aria, A. Gianoncelli, and M. Memo, "Zeolite clinoptilolite: Therapeutic virtues of an ancient mineral," *Molecules*, vol. 24, no. 8, p. 1517, 2019.

[74] I. V. Pekov, A. A. Grigorieva, A. G. Turchkova, and E. V. Lovskaya, "Natural ion exchange in microporous minerals: Different aspects and implications," in *Minerals as Advanced Materials I*, pp. 7–15, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[75] C. Algieri and E. Drioli, "Zeolite membranes: Synthesis and applications," *Separation and Purification Technology*, vol. 278, p. 119295, 2021.

[76] S. Wang and Y. Peng, "Natural zeolites as effective adsorbents in water and wastewater treatment," *Chemical Engineering Journal*, vol. 156, no. 1, pp. 11–24, 2010.

[77] M. Jiménez-Reyes, P. T. Almazán-Sánchez, and M. Solache-Ríos, "Radioactive waste treatments by using zeolites. a short review," *J. Environ. Radioact.*, vol. 233, no. 106610, p. 106610, 2021.

[78] D. G. Boer, J. Langerak, and P. P. Pescarmona, "Zeolites as selective adsorbents for co2 separation," *ACS Applied Energy Materials*, vol. 6, pp. 2634–2656, 2023.

[79] H. M. Aly, M. E. Moustafa, and E. A. Abdelrahman, "Synthesis of mordenite zeolite in absence of organic template," *Advanced Powder Technology*, vol. 23, no. 6, pp. 757–760, 2012.

[80] S. Prodinger, I. C. Berdiell, T. Cordero-Lanzac, O. R. Bygdnes, B. G. Solemsli, K. Kvande, B. Arstad, P. Beato, U. Olsbye, and S. Svelle, "Cation-induced speciation of port-size during mordenite zeolite synthesis," *J. Mater. Chem. A*, vol. 11, pp. 21884–21894, 2023.

[81] "Si crystal structure: Datasheet from 'Pauling file multinaries edition – 2022 in SpringerMaterials (`https://materials.springer.com/isp/crystallographic/docs/sd{_}0304951`)." [Accessed 2023-12-16].

[82] "GaSb crystal structure: Datasheet from 'Pauling file multinaries edition – 2022' in SpringerMaterials (`https://materials.springer.com/isp/crystallographic/docs/sd{_}0313947`)." [Accessed 2023-12-16].

[83] M. Cooper and K. Robinson, "The crystal structure of the ternary alloy α(AlMnSi)," *Acta Crystallographica*, vol. 20, pp. 614–617, 1966.

[84] W. M. Meier, "The crystal structure of mordenite (ptilolite)*," *Zeitschrift für Kristallographie*, vol. 115, no. 5-6, pp. 439–450, 1961.

[85] W. Brehm, J. Triviño, J. M. Krahn, I. Usón, and K. Diederichs, "*XDSGUI*: a graphical user interface for *XDS*, *SHELX* and *ARCIMBOLDO*," *Journal of Applied Crystallography*, vol. 56, no. 5, pp. 1585–1594, 2023.

[86] "Xprep," 2011. Bruker AXS Inc., Madison, Wisconsin, USA.

[87] "edtools — edtools 1.0.4 documentation," 2021. `https://edtools.readthedocs.io/en/latest/` [Accessed 2024-02-12].

[88] O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, and H. Puschmann, "*OLEX2*: a complete structure solution, refinement and analysis program," *Journal of Applied Crystallography*, vol. 42, no. 2, pp. 339–341, 2009.

[89] A. Thorn, B. Dittrich, and G. M. Sheldrick, "Enhanced rigid-bond restraints," *Acta Crystallographica Section A*, vol. 68, no. 4, pp. 448–451, 2012.

[90] A. L. Lamb, T. J. Kappock, and N. R. Silvaggi, "You are lost without a map: Navigating the sea of protein structures," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1854, no. 4, pp. 258–268, 2015.

[91] T. J. Slater, A. Janssen, P. H. Camargo, M. G. Burke, N. J. Zaluzec, and S. J. Haigh, "STEM-EDX tomography of bimetallic nanoparticles: A methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[92] L. Jones, "Quantitative adf stem: acquisition, analysis and interpretation," *IOP Conference Series: Materials Science and Engineering*, vol. 109, p. 012008, jan 2016.

[93] P. Brázda, M. Klementová, Y. Krysiak, and L. Palatinus, "Accurate lattice parameters from 3D electron diffraction data. I. Optical distortions," *IUCrJ*, vol. 9, no. 6, pp. 735–755, 2022.

[94] A. Basha, G. Levi, L. Houben, T. Amrani, I. Goldfarb, and A. Kohn, "Evaluating direct detection detectors for short-range order characterization of amorphous materials by electron scattering," *Ultramicroscopy*, vol. 249, p. 113737, 2023.

[95] I. MacLaren, T. A. Macgregor, C. S. Allen, and A. I. Kirkland, "Detectors—The ongoing revolution in scanning transmission electron microscopy and why this important to material characterization," *APL Materials*, vol. 8, no. 11, p. 110901, 2020.

[96] B. Wang, X. Zou, and S. Smeets, "Automated serial rotation electron diffraction combined with cluster analysis: an efficient multi-crystal workflow for structure determination," *IUCrJ*, vol. 6, no. 5, pp. 854–867, 2019.

[97] M. Buxhuku, V. Hansen, P. Oleynikov, and J. Gjønnes, "The determination of rotation axis in the rotation electron diffraction technique," *Microsc. Microanal.*, vol. 19, no. 5, pp. 1276–1280, 2013.

[98] E. H. HIRSCH, "Stability of zeolites under electron irradiation," *Nature*, vol. 293, pp. 759–759, 1981.

[99] S. Wang, L. Wang, and R. Ewing, "Electron and ion irradiation of zeolites," *Journal of Nuclear Materials*, vol. 278, no. 2, pp. 233–241, 2000.

[100] R. Csencsits and R. Gronsky, "Damage of zeolite y in the tem and its effects on tem images," *Ultramicroscopy*, vol. 23, no. 3, pp. 421–431, 1987.

[101] M. Treacy and J. Newsam, "Electron beam sensitivity of zeolite l," *Ultramicroscopy*, vol. 23, no. 3, pp. 411–419, 1987.

[102] M. J. Dobro, L. A. Melanson, G. J. Jensen, and A. W. McDowall, "Plunge freezing for electron cryomicroscopy," in *Methods in Enzymology*, Methods in enzymology, pp. 63–82, Elsevier, 2010.

[103] J. M. Rodenburg, "Understanding transmission electron microscope alignment: A tutorial," *Microscopy and analysis*, vol. 18, 2004.

[104] R. Egerton, "Choice of operating voltage for a transmission electron microscope," *Ultramicroscopy*, vol. 145, pp. 85–93, 2014. Low-Voltage Electron Microscopy.

[105] T. Muller-Reichert and G. Pigino, *Three-Dimensional Electron Microscopy*. Elsevier Science, 2019.

[106] J. A. Seibert, J. M. Boone, and K. K. Lindfors, "Flat-field correction technique for digital detectors," in *Medical Imaging 1998: Physics of Medical Imaging* (J. T. Dobbins and J. M. Boone, eds.), vol. 3336 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 348–354, 1998.

[107] "pyxem/pyxem: v0.16.0," 2023. https://doi.org/10.5281/zenodo.10125648 [Accessed 24-05-10].

[108] G. Capitani, P. Oleynikov, S. Hovmöller, and M. Mellini, "A practical method to detect and correct for lens distortion in the TEM," *Ultramicroscopy*, vol. 106, no. 2, pp. 66–74, 2006.

[109] I. Nederlof, E. van Genderen, Y.-W. Li, and J. P. Abrahams, "A Medipix quantum area detector allows rotation electron diffraction data collection from submicrometre three-dimensional protein crystals," *Acta Crystallographica Section D*, vol. 69, no. 7, pp. 1223–1230, 2013.

[110] L. J. Bourhis, O. V. Dolomanov, R. J. Gildea, J. A. K. Howard, and H. Puschmann, "The anatomy of a comprehensive constrained, restrained refinement program for the modern computing environment - olex2 dissected," *Acta Crystallogr. A Found. Adv.*, vol. 71, no. Pt 1, pp. 59–75, 2015.

[111] "Iucr dictionary: Friedel's law.," 2017. https://dictionary.iucr.org/Friedel%27s_law [Accessed 2024-03-10].

[112] J. Wang and P. B. Moore, "On the interpretation of electron microscopic maps of biological macromolecules," *Protein Sci.*, vol. 26, no. 1, pp. 122–129, 2017.

[113] T. Gruene, J. J. Holstein, G. H. Clever, and B. Keppler, "Establishing electron diffraction in chemical crystallography," *Nat. Rev. Chem.*, vol. 5, no. 9, pp. 660–668, 2021.

[114] S. Brennan and P. L. Cowan, "A suite of programs for calculating x-ray absorption, reflection, and diffraction performance for a variety of materials at arbitrary wavelengths," *Review of Scientific Instruments*, vol. 63, no. 1, pp. 850–853, 1992.

[115] A. L. Spek, "Structure validation in chemical crystallography," *Acta Crystallographica Section D*, vol. 65, no. 2, pp. 148–155, 2009.

[116] J. Munshi, A. Rakowski, B. H. Savitzky, S. E. Zeltmann, J. Ciston, M. Henderson, S. Cholia, A. M. Minor, M. K. Y. Chan, and C. Ophus, "Disentangling multiple scattering with deep learning: application to strain mapping from electron diffraction patterns," *npj Computational Materials*, vol. 8, p. 254, 2022.

[117] J. Chen, H. Zhang, C. B. Wahl, W. Liu, C. A. Mirkin, V. P. Dravid, D. W. Apley, and W. Chen, "Automated crystal system identification from electron diffraction patterns using multiview opinion fusion machine learning," *Proceedings of the National Academy of Sciences*, vol. 120, no. 46, p. e2309240120, 2023.

[118] X. Liu, K. Amini, A. Sanchez, B. Belsa, T. Steinle, and J. Biegert, "Machine learning for laser-induced electron diffraction imaging of molecular structures," *Communications Chemistry*, vol. 4, p. 154, 2021.

[119] E. Svensson Grape, A. J. Chacón-García, S. Rojas, Y. Pérez, A. Jaworski, M. Nero, M. Åhlén, E. Martínez-Ahumada, A. E. Galetsa Feindt, M. Pepillo, M. Narongin-Fujikawa, I. A. Ibarra, O. Cheung, C. Baresel, T. Willhammar, P. Horcajada, and A. K. Inge, "Removal of pharmaceutical pollutants from effluent by a plant-based metal–organic framework," *Nature Water*, vol. 1, pp. 433–442, 2023.

[120] M. N. Naseer, J. Jaafar, H. Junoh, A. A. Zaidi, M. Kumar, A. Alqahtany, R. Jamil, S. H. Alyami, and N. A. Aldossary, "Metal-Organic frameworks for wastewater decontamination: Discovering intellectual structure and research trends," *Materials (Basel)*, vol. 15, no. 14, 2022.

[121] N. Cautaerts, P. Crout, H. W. Ånes, E. Prestat, J. Jeong, G. Dehm, and C. H. Liebscher, "Free, flexible and fast: Orientation mapping using the multi-core and gpu-accelerated template matching capabilities in the python-based open source 4d-stem analysis toolbox pyxem," *Ultramicroscopy*, vol. 237, p. 113517, 2022.

BIBLIOGRAPHY

Appendix A

Supplementary material

A.1   Experimental data

Table A.1: Overview of Si datasets with experimental parameters used.

	Si_old	Si2	Si4_8	Si4_9	Si5_10
Voltage [kV]	80	80	80	80	80
Camera length* [cm]	100	150	150	120	120
Bitdepth [bit]	12	12	12	12	12
Holder	-	G	G	G	G
Number of frames	168	170	384	484	424
Defocused images	-	-	127	160	140
Exposure time[†] [s]	0.5	0.5	0.5	0.5	0.5
Acquisition time[†] [s]	-	-	0.8	0.799	0.799
Total acquisition time [s]	106.45	140.64	510.12	642.096	562.479
Image interval [s]	5	5	5	5	5
Tilt speed	-	-	-	-	-
Spot size	1	1	1	1	1
Condenser aperture	-	-	-	-	-
Oscillation range [°]	0.3047	0.4001	0.165	0.1372	0.1593
Flatfield applied [y/n]	n	n	n	n	n
Rotation direction [+/-]	-	+	-	+	-
Rotation range[‡] [°]	106.45	84.82	105.28	110.34	112.15

---

*nominal camera length

[†]per frame

[‡]+ rotation from positive to negative angles, - from negative to positive

Table A.2: Overview of GaSb datasets with experimental parameters used.

	GaSb1_1	GaSb1_5
Voltage [kV]	80	80
Camera length* [cm]	120	120
Bitdepth [bit]	12	12
Holder	G	G
Number of frames	304	463
Defocused images	100	154
Exposure time† [s]	0.5	0.5
Acquisition time† [s]	0.8	0.799
Total acquisition time [s]	403.047	615.795
Image interval [s]	5	5
Tilt speed	-	-
Spot size	1	1
Condenser aperture	-	-
Oscillation range [°]	0.1631	0.1405
Flatfield applied [y/n]	n	n
Rotation direction [+/-]	-	+
Rotation range‡ [°]	82.19	108.36

Table A.3: Overview of NaMOR datasets with experimental parameters used.

	Na-MOR-9	Na-MOR-10	Na-MOR-11	Na-MOR-13
Voltage [kV]	80	80	80	80
Camera length* [cm]	100	80	80	60
Bitdepth [bit]	12	12	12	12
Holder	-	G	G	G
Number of frames	-	113	100	-
Defocused images	-	37	33	-
Exposure time† [s]	0.5	0.5	0.5	0.5
Acquisition time† [s]	-	0.803	0.801	-
Total acquisition time [s]	-	150.249	132.958	-
Image interval [s]	5	5	5	5
Tilt speed	-	-	-	-
Spot size	2	2	2	2
Condenser aperture	1	1	2	-
Oscillation range [°]	-	0.4876	0.5264	-
Flatfield applied [y/n]	n	n	n	n
Rotation direction [+/-]	N/A	+	-	N/A
Rotation range‡ [°]	-	91.19	87.39	-

Table A.4: Overview of NaMOR datasets with experimental parameters used.

	Na-MOR-15	Na-MOR-16	Na-MOR-17	Na-MOR-18
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	100	100	100
Bitdepth [bit]	12	12	12	12
Holder	G	G	-	G
Number of frames	118	133	133	137
Defocused images	39	44	43	45
Exposure time† [s]	0.5	0.5	0.5	0.5
Acquisition time† [s]	0.801	0.8	0.8	0.803
Total acquisition time [s]	156.945	176.734	175.198	182.176
Image interval [s]	5	5	5	5
Tilt speed	-	-	-	-
Spot size	2	2	2	2
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range [°]	0.5003	0.5187	0.5213	0.4972
Flatfield applied [y/n]	n	n	n	n
Rotation direction [+/-]	+	-	-	+
Rotation range‡ [°]	98.06	114.63	114.17	112.86

Table A.5: Overview of NaMOR datasets with experimental parameters used.

	Na-MOR-19	Na-MOR-21	Na-MOR-22	Na-MOR-24
Voltage [kV]	200	200	200	200
Camera length* [cm]	100	200	150	120
Bitdepth [bit]	12	12	12	12
Holder	G	G	G	G
Number of frames	-	112	121	-
Defocused images	-	37	40	-
Exposure time† [s]	0.5	0.5	0.5	0.5
Acquisition time† [s]	-	0.801	0.801	-
Total acquisition time [s]	-	148.946	160.96	-
Image interval [s]	5	5	5	5
Tilt speed	-	-	-	-
Spot size	2	2	2	2
Condenser aperture	CA1	CA1	CA1	CA1
Oscillation range [°]	-	0.5136	0.482	-
Flatfield applied [y/n]	n	n	n	n
Rotation direction [+/-]	+	-	+	-
Rotation range‡ [°]	-	95.53	96.89	-

Table A.6: Overview of NaMOR datasets with experimental parameters used.

Dataset name	Na-MOR-25	Na-MOR-27	Na-MOR-28	Na-MOR-31
Voltage [kV]	200	200	200	200
Camera length [cm]	120	150	150	150
Bitdepth [bit]	12	12	12	12
Holder	G	G	G	G
Number of frames	-	-	101	89
Defocused images	-	-	33	29
Exposure time† [s]	0.5	0.5	0.5	0.5
Acquisition time† [s]	-	-	0.805	0.806
Total acquisition time [s]	-	-	134.471	118.514
Image interval	5	5	5	5
Tilt speed	-	-	-	-
Spot size	2	2	2	3
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range	-	-	0.4882	0.5276
Flatfield applied [y/n]	n	n	n	n
Rotation direction [+/-]	N/A	N/A	+	-
Rotation range [°]	-	-	81.53	77.55

Table A.7: Overview of the studied α-Al(Mn, Fe)Si datasets with experimental parameters used.

	α-Al(Mn,Fe)Si$_1$	α-Al(Mn,Fe)Si$_3$	α-Al(Mn,Fe)Si$_4$	α-Al(Mn,Fe)Si$_5$
Voltage [kV]	200	200	200	200
Camera length* [cm]	150	120	100	80
Bitdepth [bit]	12	12	12	12
Holder	G	G	G	G
Number of frames	-	-	-	-
Defocused images	-	-	-	-
Exposure time† [s]	0.5	0.5	0.5	0.5
Acquisition time† [s]	-	-	-	-
Total acquisition time [s]	-	-	-	-
Image interval [s]	5	5	5	5
Tilt speed	-	-	-	-
Spot size	3	3	3	3
Condenser aperture	CA1	CA1	CA2	CA2
Oscillation range [°]	-	-	-	-
Flatfield applied [y/n]	n	n	n	n
Rotation direction‡ [+/-]	N/A	N/A	N/A	N/A
Rotation range [°]	-	-	-	-

Table A.8: Overview of the studied α-Al(Mn, Fe)Si datasets with experimental parameters used.

	α-Al(Mn,Fe)Si$_7$	α-Al(Mn,Fe)Si$_8$	α-Al(Mn,Fe)Si$_{10}$
Voltage [kV]	200	200	200
Camera length* [cm]	60	80	100
Bitdepth [bit]	12	12	12
Holder	G	G	G
Number of frames	-	-	-
Defocused images	-	-	-
Exposure time† [s]	0.5	0.5	0.5
Acquisition time† [s]	-	-	-
Total acquisition time [s]	-	-	-
Image interval [s]	5	5	5
Tilt speed	-	-	-
Spot size	3	3	3
Condenser aperture	CA2	CA2	CA2
Oscillation range [°]	-	-	-
Flatfield applied [y/n]	n	n	n
Rotation direction‡ [+/-]	N/A	N/A	N/A
Rotation range [°]	-	-	-

Table A.9: Overview of the studied K-MOR datasets with experimental parameters used.

	K-MOR-5	K-MOR-14	K-MOR-17	K-MOR-18
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	-	350	244	358
Defocused images	-	57	40	59
Exposure time[†] [s]	0.5	0.1	0.1	0.1
Acquisition time[†] [s]	-	0.307	0.309	0.308
Total acquisition time [s]	-	177.25	124.91	183.16
Image interval [s]	5	10	10	10
Tilt speed	2x	2x	2x	2x
Spot size	2	2	2	2
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range [°]	-	0.1874	0.1911	0.2008
Flatfield applied [y/n]	y	y	y	y
Rotation direction[‡] [+/-]	N/A	+	+	-
Rotation range [°]	98.26	108.32	77.19	119.28

Table A.10: Overview of the studied K-MOR datasets with experimental parameters used.

	K-MOR-20	K-MOR-24	K-MOR-27	K-MOR-31
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	-	143	147	123
Defocused images	-	70	72	60
Exposure time† [s]	0.1	0.1	0.1	0.1
Acquisition time† [s]	-	0.517	0.517	0.518
Total acquisition time [s]	-	182.106	187.206	156.366
Image interval [s]	10	10	10	10
Tilt speed	2x	2x	2x	2x
Spot size	2	2	2	2
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range [°]	-	0.3353	0.3354	0.3242
Flatfield applied [y/n]	y	y	y	y
Rotation direction‡ [+/-]	N/A	-	-	+
Rotation range [°]	123.72	118.01	121.4	97.91

Table A.11: Overview of the studied Cu-MOR-K datasets with experimental parameters used.

	Cu-MOR-K-1	Cu-MOR-K-2	Cu-MOR-K-3	Cu-MOR-K-4
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	141	118	136	110
Defocused images	23	19	22	17
Exposure time† [s]	0.1	0.1	0.1	0.1
Acquisition time† [s]	0.308	0.313	0.312	0.31
Total acquisition time [s]	71.766	60.713	69.838	55.095
Image interval [s]	10	10	10	10
Tilt speed	2x	2x	2x	2x
Spot size	2	2	2	2
Condenser aperture	CA2	CA2	CA2	CA2
Oscillation range [°]	0.5583	0.5695	0.5692	0.5653
Flatfield applied [y/n]	y	y	y	y
Rotation direction‡ [+/-]	-	+	-	+
Rotation range [°]	130.09	110.49	127.51	100.63

Table A.12: Overview of the studied Cu-MOR-K datasets with experimental parameters used.

	Cu-MOR-K-7	Cu-MOR-K-8	Cu-MOR-K-15	Cu-MOR-K-22
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	135	130	340	54
Defocused images	22	21	56	25
Exposure time† [s]	0.1	0.1	0.1	0.1
Acquisition time† [s]	0.308	0.312	0.308	0.524
Total acquisition time [s]	68.738	66.738	173.87	67.015
Image interval [s]	10	10	10	5
Tilt speed	2x	2x	1x	2x
Spot size	2	2	2	2
Condenser aperture	CA3	CA3	CA3	CA3
Oscillation range [°]	0.5605	0.5711	0.2042	0.9532
Flatfield applied [y/n]	y	y	y	y
Rotation direction‡ [+/-]	-	+	-	+
Rotation range [°]	124.995	122.21	115.18	122.01

Table A.13: Overview of the studied Cu-MOR-K datasets with experimental parameters used.

	Cu-MOR-K-28	Cu-MOR-K-30	Cu-MOR-K-31	Cu-MOR-Na-4
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	45	54	53	150
Defocused images	21	25	25	73
Exposure time† [s]	0.1	0.1	0.1	0.1
Acquisition time† [s]	0.522	0.524	0.521	0.518
Total acquisition time [s]	55.869	67.008	66.225	190.729
Image interval [s]	5	5	5	5
Tilt speed	3x	3x	3x	2x
Spot size	2	2	2	2
Condenser aperture	CA3	CA3	CA3	CA3
Oscillation range [°]	0.9494	0.9496	0.9535	0.3214
Flatfield applied [y/n]	y	y	y	y
Rotation direction‡ [+/-]	+	-	+	+
Rotation range [°]	101.59	121.55	121.1	118.27

Table A.14: Overview of Cu-MOR-Na datasets with experimental parameters used.

	Cu-MOR-Na-5	Cu-MOR-Na-8	Cu-MOR-Na-9	Cu-MOR-Na-11
Voltage [kV]	80	80	80	80
Camera length* [cm]	120	120	120	120
Bitdepth [bit]	16	16	16	16
Holder	J	J	J	J
Number of frames	129	47	48	54
Defocused images	63	22	22	25
Exposure time† [s]	0.1	0.1	0.1	0.1
Acquisition time† [s]	0.518	0.521	0.525	0.524
Total acquisition time [s]	164.134	58.357	59.297	67.122
Image interval [s]	5	5	5	5
Tilt speed	3x	3x	3x	3x
Spot size	2	2	2	2
Condenser aperture	CA3	CA3	CA3	CA3
Oscillation range [°]	0.338	0.9508	0.9626	0.9563
Flatfield applied [y/n]	y	y	y	y
Rotation direction‡ [+/-]	-	+	-	-
Rotation range [°]	107.15	106.49	108.77	122.41

Table A.15: Overview of Cu-MOR-Na datasets with experimental parameters used.

	Cu-MOR-Na-14	Cu-MOR-Na-15	Cu-MOR-Na-16
Voltage [kV]	80	80	80
Camera length* [cm]	120	120	120
Bitdepth [bit]	16	16	16
Holder	J	J	J
Number of frames	53	49	45
Defocused images	25	23	21
Exposure time[†] [s]	0.1	0.1	0.1
Acquisition time[†] [s]	0.521	0.521	0.522
Total acquisition time [s]	66.152	60.954	55.865
Image interval [s]	5	5	5
Tilt speed	3x	3x	3x
Spot size	2	2	2
Condenser aperture	CA3	CA3	CA3
Oscillation range [°]	0.9455	0.9378	0.9514
Flatfield applied [y/n]	y	y	y
Rotation direction[‡] [+/-]	+	-	+
Rotation range [°]	120.08	109.72	101.8

Table A.16: Overview of Cu-MOR-Na datasets with experimental parameters used.

	Cu-MOR-Na-18	Cu-MOR-Na-20	Cu-MOR-Na-22
Voltage [kV]	80	80	80
Camera length* [cm]	120	120	120
Bitdepth [bit]	16	16	16
Holder	J	J	J
Number of frames	50	51	52
Defocused images	23	24	24
Exposure time† [s]	0.1	0.1	0.1
Acquisition time† [s]	0.524	0.521	0.525
Total acquisition time [s]	61.85	63.546	64.602
Image interval [s]	5	5	5
Tilt speed	3x	3x	3x
Spot size	2	2	2
Condenser aperture	CA3	CA3	CA3
Oscillation range [°]	0.9552	0.9388	0.9607
Flatfield applied [y/n]	y	y	y
Rotation direction‡ [+/-]	-	+	-
Rotation range [°]	112.71	114.53	118.17

Table A.17: Overview of H-MOR-K datasets with experimental parameters used.

	H-MOR-K-1	H-MOR-K-4	H-MOR-K-6
Voltage [kV]	80	80	80
Camera length* [cm]	120	120	120
Bitdepth [bit]	16	16	16
Holder	J	J	J
Number of frames	47	52	131
Defocused images	22	24	64
Exposure time† [s]	0.1	0.1	0.1
Acquisition time† [s]	0.522	0.525	0.518
Total acquisition time [s]	58.499	64.524	166.815
Image interval [s]	5	5	5
Tilt speed	3x	3x	1x
Spot size	2	2	2
Condenser aperture	CA3	CA3	CA3
Oscillation range [°]	0.9454	0.9566	0.3271
Flatfield applied [y/n]	y	y	y
Rotation direction‡ [+/-]	+	-	+
Rotation range [°]	105.89	117.66	105.34

A.2   List of Github-scripts

1. Elliptical distortion correction: `https://github.com/aurorateien/master-thesis/blob/main/AuCalibration.ipynb`

2. Removal of untrusted rectangle: `https://github.com/aurorateien/master-thesis/blob/main/removeRectangle.ipynb`

3. cRED conversion without Instamatic: `https://github.com/TEM-Gemini-Centre/cRED/blob/main/cRED_conversion.ipynb`

A.3   Manual

The following pages contain the user manual developed in this work. Note that the page numbering is from now on nulled.

NTNU
Kunnskap for en bedre verden

Manual for 3D electron diffraction data processing

Author:

Aurora Teien

Last updated June, 2024
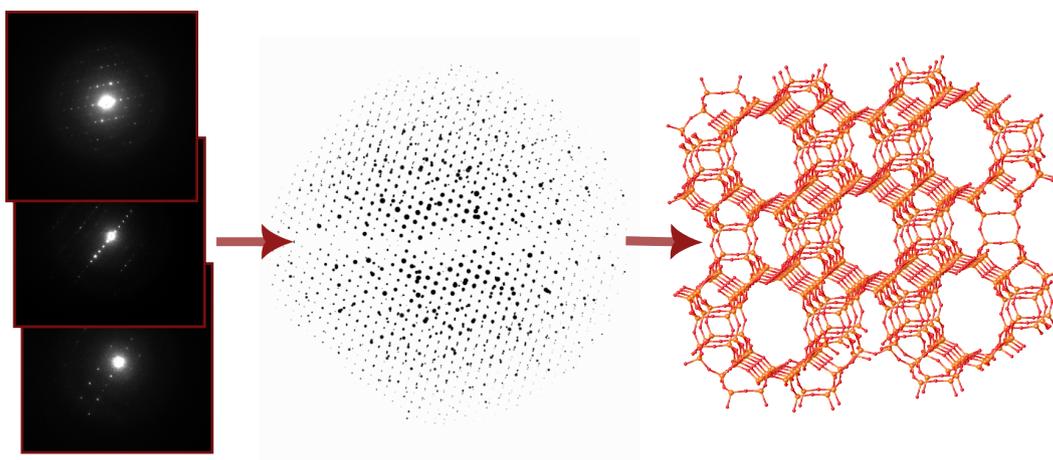
Contents

1 Introduction

3D electron diffraction (3D ED) is an up- and coming technique for crystal structure analysis of sub-micron crystals [1]. The technique can be performed on dedicated microscopes like Eldico etc. or in a regular transmission electron microscope (TEM) [2]. The data collection process is fast (seconds to minutes) and low dose compared to (S)TEM imaging techniques, although post-processing of 3D ED data might follow rather tedious and time-consuming routines that might appear like black box magic [3]. A possible 3D ED data processing workflow is described in this manual. Five softwares with different functions and purposes are used successively: REDp, XDS, XPREP, OLEX2 with SHELXT and SHELXL. These are described in detail in this manual. This manual first goes through a short description of all the programs, including how to install them. Next, a short checklist of the full procedure is described. Further, guides describing how to use each individual program are provided. The guides are divided into categories that describe their purpose. Problems that might arise during data processing are marked with warning signs ($\triangle$) with potential solutions to the problem.

This manual is based on the procedure described in the protocol developed by Yang et al. [4]. The purpose of making such a manual, when the above protocol already exists, is to adapt this method to NTNU-specific equipment and furthermore make the process easier for new users.

This manual was developed by Aurora Teien as part of her Master thesis in the spring of 2024. The document will be available from NTNU-open October 2024. The information in this manual is described more extensively in her specialization project and Master's thesis.

2   Tools

The following section describes the tools used in a typical data processing routine for 3D ED data. A visual representation of the full workflow covered in this manual is shown in Figure 1. The raw data can first be preprocessed in Python, before REDp and/or XDS are used for structure determination. XPREP can also be used for the preparation of SHELXT files, which is used for structure solution, and finally SHELXL and OLEX2 are used for refinement and visualization of the crystal structure.
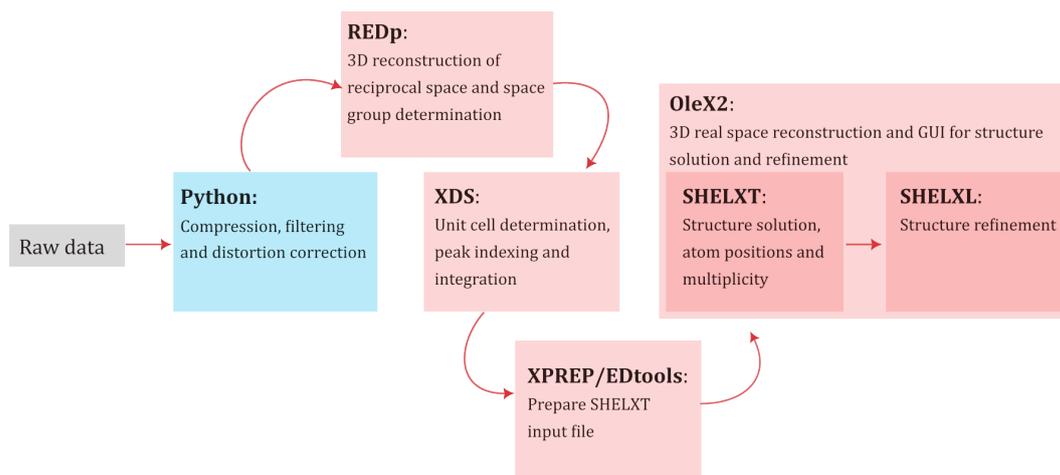
Figure 1: Overview of the workflow and different softwares used for data processing of 3D ED.

The programs might have different input and output file formats. An overview of the required input file formats for each program is provided in Table 1. Note that most of the programs take both an instructional input and a data input, where the instructional input tells the program what to do and contains metadata, while the data input usually contain the individual frames or intensities.

Table 1: Programs with corresponding input and output file formats.

Program	Instructional input	Data input	Output
REDp	.3ded	.mrc	.img
XDS	.INP	.img	.hkl, .lp, .cbf
XPREP		.hkl	.ins, .hkl, .prp
SHELXT	.ins	.hkl	.res, .fcf, .ins, .lxt
SHELXL	.ins	.hkl	.ins, .lst, .cif

2.1 Instamatic

Data collection can be done manually or with the use of Instamatic. Instamatic is a Python-based program for automated collection of 3D ED data [5]. The latest version is available from Zenodo, https://zenodo.org/records/5175957.

Alternatively, it can be directly imported as a Python library through the command `pip install instamatic` in a desired `conda` environment.

2.2 Python preprocessing

Manual data collection requires pre-processing in Python before using the subsequent data processing tools. The raw data from a manually obtained 3D ED tilt series have the generic file format `.mib`. Such data sets are often big (>1 GB) and need to be compressed, rebinned, and converted to other file formats to be used in further processing. As can be seen in Table 1, a series of different file formats are being used for the different softwares. In addition the defocused frames should be removed from the particle tracking during data acquisition. This is all done in the `cRED_conversion.ipynb` script available from TEM Gemini Centre GitHub, see
https://github.com/TEM-Gemini-Centre/developments/blob/cRED-dev/cRED/cRED_conversion.ipynb

However, if the data are obtained using Instamatic, then the data should be fine without further preprocessing. The data stacks are automatically filtered for defocused frames, and converted into useful formats (both for REDp and XDS as well as the frames being saved in `.tiff` format). Flatfield correction in Instamatic can correct for permanent noise or imaging defects such as dead pixels and bright untrusted rectangles between the detector quadrants.

2.2.1 EDtools

EDtools is a Python-based library. It is used for two main purposes in this manual. First of all, it can be used for batch data processing, where multiple datasets can run simultaneously. Secondly, it can be used as an alternative to XPREP (which is not open-source) [4]. EDtools can prepare input files for SHELXT. Download is available by using `pip install edtools` or download from GitHub, https://github.com/instamatic-dev/edtools

2.3 REDp

Rotation Electron Diffraction processing (REDp) is a software developed to visualize and process 3DED data obtained by continuous rotation electron diffraction (cRED) [6]. For each ED frame, the program is able to do shift correction, as well as peak finding and identifying diffraction spots in 3D reciprocal space. Further, the program can be used to

determine lattice parameters and hkl-indexing as part of the full 3D mapping of points in reciprocal lattice. REDp includes a 3D viewer and gives intensities and hkl-indices as output. However, the hkl-files produced from REDp are usually not used, as XDS rather is used for this purpose.

REDp can be downloaded from

https://zenodo.org/records/2545322#.XEW5sC0o8uU

Instructions for usage is given in a pdf included in the downloaded REDp package. If you experience problems with missing .dll-files after download, make sure your computer has the right version of Microsoft Visual C++ (I needed the 2010, 2012 and 2013 version). Try other versions if these do not work. Download is free, from

https://www.microsoft.com/en-us/download/details.aspx?id=40784.

2.4 XDS

X-ray Diffraction Software (XDS) can be downloaded from [7] and runs only on the Linux operating system. For Windows, a Linux subsystem is available from:

https://learn.microsoft.com/en-us/windows/wsl/install

The software requires an input file of format .INP that gives instructions and information to the program. XDS can also generate a default version of this type of file based on a single frame of the dataset. XDS itself is only terminal-based and provides many resulting output files when running properly. A more user-friendly user interface called 'XDSGUI' can be downloaded for better visualization of the process [8].

A complete list of all input parameters and their functionality is available on the official website,

https://xds.mr.mpg.de/html_doc/xds_parameters.html

A more detailed tutorial for XDS is available through

https://wiki.uni-konstanz.de/pub/MWhitley_CSHL-2018_XDS-Tutorial.pdf.

2.5 XPREP

XPREP was developed by George Sheldrick and prepares the input files for SHELXT. Previously, this was part of the SHELX package, but it is now licensed and sold by Bruker. The program is terminal-based. A detailed description of how it works is available in this tutorial:

https://xray.uky.edu/Tutorials/sucrose-routine/shelxtl_routine_xprep.html

The program should run in the same folder as the target file. To run the program from any folder or path, add the parent folder for XPREP to your system PATH by following the steps in this guide:

https://www.architectryan.com/2018/03/17/add-to-the-path-on-windows-10/.

2.6 SHELXT and SHELXL

SHELX is a series of packages that uses direct or dual-space methods for crystal structure solution and refinement [9]. SHELXT specifically is one of the packages that specializes in solving small molecular structures. SHELXL is a structure refinement program for diffraction data [10]. Both SHELXT and SHELXL run in a command terminal. They can also be run inside the Olex2 GUI. Before downloading one has to register as a user and answer a very doable control question about crystallography. Download is available through: https://shelx.uni-goettingen.de/download.php

2.7 OLEX2

OLEX2 is an open-source software developed for visualization, determination, and analysis of molecular crystal structures [11]. This software has a much more extensive user interface than the previously described programs, including a 3D view of the molecular structure in the unit cell.
Download is available from
https://www.olexsys.org/olex2/docs/getting-started/installing-olex2/

2.8 Other useful tools

ReciPro is an open-source software that combines crystallographic databases with diffraction pattern simulation. This can be useful, for example, when indexing diffraction patterns from slices in REDp [12]. It can be installed from,
https://github.com/seto77/ReciPro/releases/tag/v.4.878
ImageJ is a useful tool for image processing. It is widely used for a broad range of purposes, but in this manual it is used only to count the pixels between certain features in an image. This can also be done by other image processing tools. ImageJ can be downloaded from,
https://imagej.net/ij/download.html
VESTA is a 3D visualization tool for crystallographic structures [13]. It might also be used for electronic state calculations. Download is available through,
https://jp-minerals.org/vesta/en/download.html

2.9 Datasets used

1. Silicon dataset 5_10 were used for demonstrating REDp. Available for download from https://github.com/aurorateien/master-thesis.

2. Dataset Cu-MOR-K-31 were used for demonstrating how to find reflection conditions in REDp. Available for download from https://github.com/aurorateien/master-thesis.

3. The MOR2 mordenite dataset was used for XDS, structure solution and refinement. Files available for download from https://zenodo.org/records/1321880.

3 Checklist for solving a structure

Python preprocessing:

- Calibrate values such as pixel size and camera length

- (if necessary) elliptical distortion correction

REDp:

- Browse frames in datasets to evaluate the quality of the full dataset

- Determine axis of rotation (if you do not already know this)

- Check for overexposure, by inspecting the intensities in frames

- Find direct beam centre positions (x, y) to input in XDS

- Do peak search, check that the resolution (should be > 1 Å) stays acceptable throughout the dataset and find unit cell parameters

- Click on the `cctbx`-link, find the most symmetric space group and do transformation of the unit cell

- Search for extinction rules and compare to IUCr table to find space group

XDS:

- Edit input file XDS.INP: update incorrect info such as WAVELENGTH, ROTATION AXIS, DETECTOR_DISTANCE, UNTRUSTED_RECTANGLE, path to img-files

- Input info from REDp into XDS.INP: ORGX, ORGY and OVERLOAD

- Check that the program is running correctly (not crashing)

- Check SCALE under the INTEGRATE-tab and remove frames causing abrupt spikes in the plot with EXCLUDE_FRAMES

- Make sure that enough reflections are indexed. If necessary, adjust TRUSTED_REGION, EXCLUDE_ICE_RING, INCLUDE_RESOLUTION_RANGE, MAXIMUM_ERROR_OF _SPOT_POSITION and/or MAXIMUM_ERROR_OF _SPINDLE_POSITION

- Make sure that enough reflections are saved to hkl file. If necessary, adjust DELPHI

- After one successful run, update the parameters SPACE_GROUP and LATTICE_PARAMETERS according to the space group found in REDp and the lattice parameters suggested by the initial XDS run

- Convert XDS_ASCII.HKL to SHELX format using XDSCONV

- Merge multiple datasets with XSCALE if necessary (in the case of low completeness)

XPREP: (if you already have an `.ins` file for the correct space group/element, this step can be ignored)

- Find space group (with best combined figure of merit (CFOM))

- Merge identical reflections in .hkl file

- Input chemical composition

- Create `.ins` file for use in SHELX [1]

- Correct the wavelength and insert SFAC values for electrons in `.ins` file

Olex2:

- Run SHELXT to solve the structure

- Evaluate the solution and select alternative solutions (with higher R-value) if you think they are more correct. Alternatively, run a new solution with forced space group.

- Check that the chemical composition is correct and that the atoms are placed correctly according to the electron density map

- Run initial refinement using SHELXL

- Remove bonds and atoms that appear to be incorrectly placed in the GUI (e.g. chemically forbidden phenomena)

- Introduce restraints or constraints to limit bond lengths or angles in the model

- Run refinement until the statistical value $Shift$ converge towards a fixed minima, R-values are within acceptable range and electron density map fits with the displayed molecule.

---

[1] Can also be done with EDtools. See Section 4.4.1.

4 Detailed guides

4.1 Python preprocessing

4.1.1 Calibration

Both not and XDS require some calibrated values before usage. This can be done in Python preprocessing or manually by inspecting individual frames. The distance between two equivalent planes that correspond to single spots in the diffraction pattern is measured in terms of the number of pixels n_p with imaging tools like ImageJ. This quantity is proportional to the inverse spacing between the planes in real space d_{hkl}, namely the reciprocal distance between diffraction spots $1/d_{hkl}$. The calibrated camera length L, is related to the wavelength of the electron λ, the physical pixel size R and the scale $d = n_p * d_{hkl}$ through the following equation

$$Rd = \lambda L, \tag{1}$$

where the physical pixel size for Medipix direct electron detector (DED) R is 55 μm, the wavelength of the accelerated electrons used is 4.1757 Å for accelerating voltage of 80 kV (2.5079 Å for 200 kV). The scale value, $1/d$ in $\text{Å}^{-1}\text{px}^{-1}$, was used as pixel size in REDp. The calculated L (in mm) were used as DETECTOR_ DISTANCE in XDS. For example for a camera length 120 cm at 80 kV (and scale of $0.0056715\,\text{Å}^{-1}\text{px}^{-1}$) the detector distance is 232.24 mm.

For a well-known setup and acceleration voltage, the calibrated scales are tabulated from before, see https://github.com/TEM-Gemini-Centre/developments/blob/main/Calibration/calibrations.json.

4.1.2 Distortion correction

Optional preprocessing includes noise reduction and correction for elliptical distortions. The distortion in frames can be corrected for in different ways, either by Python scripts or directly in XDS. Using pyxem, follow the instructional steps of the demo from https://github.com/pyxem/pyxem-demos/blob/master/03%20Reference%20Standards%20-%20Dimension%20Calibrations%20-%20Rotation%20Calibrations.ipynb. Scripts for ellipticity checks and distortion correction are additionally available from https://github.com/aurorateien/master-thesis.

4.2 Determine lattice parameters and space group

To determine lattice parameters and space group of the crystal material, REDp is used first before XDS is applied. Dataset Si5_10 is used for the demonstration.

4.2.1 How to use REDp

REDp has a fairly logical GUI and in addition a thorough manual included in the package describing how to use it. Here is a brief step-by-step summary of REDp data processing for checking the data quality, determining the unit cell parameters, and space group of a sample material. Dataset Si5_10 is used for this purpose.

1. *File upload*: Find and open your `.ed3d` file using File → Open. The `.ed3d` file must be in the same folder as the corresponding `.mrc` data for each frame.

2. *Check metadata*: Important metadata is available under Control panel → Frames, see Figure 2. A table of all frames and their corresponding stage and beam tilt[2], as well as information about wavelength, pixel size, rotation axis, and intensity range are displayed here and should be correct before proceeding. The data can be altered directly in the GUI or by opening the `.3ded` input file in a text editor. See Figure 2. Only the latter of those options will cause permanent change, i.e. if one changes the metadata in the GUI it will not be saved for future use. The pixel size must be calibrated. The rotation axis has to be determined in the software if not known, and can be shown in the frame as a blue dashed line by clicking the "Show on EDP"-button; see Figure 2.
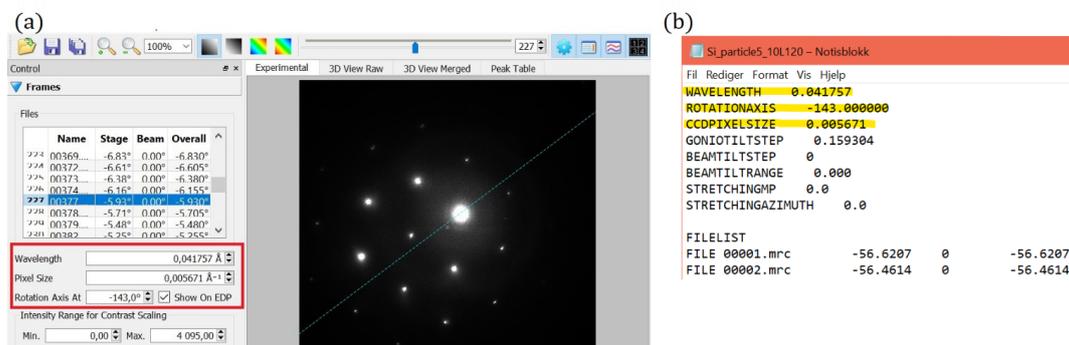
Figure 2: (a) Initial interface of REDp. The metadata in the red square needs to be checked and possibly updated and can be edited in the GUI directly or in the .3ded-file as shown in (b).

3. *Inspect intensity in frames:* Further the intensity in the frames can be inspected by holding and dragging the right mouse button, see Figure 3. The intensity in the specific pixel you point the mouse at will be visible. In addition, the intensities of the nearby pixels will be available in the "Text Data" tab to the lower right of the screen, moving simultaneously along with the mouse. The values are shown in text and coloured according to the amplitude of the intensity, red is the highest and blue

---

[2]Beam tilt is not used in cRED data acquisition, as the tilt is continuous. Previous methods (RED) used discrete tilts.

is the lowest. This can be used as a heat map revealing where the diffraction spots are located. Find the intensity of the direct beam; it will be used as the OVERLOAD-parameter in XDS.
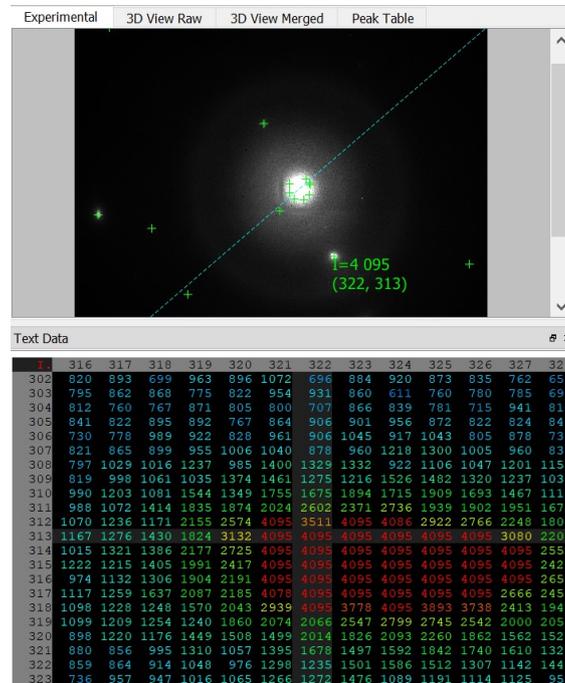

Figure 3: Intensity in frames shown as a heatmap in the text data around an overexposed diffraction spot. The text data can be found by default in the lower right corner.

4. *Shift evaluation*: Select the size of the region and move the center of the frame (marked with a red cross) to the main beam spot, as shown in Figure 4. This corrects deviations from the original position of the main beam and defines a common origin for all frames. The whole direct beam should be inside the yellow square for all frames throughout the whole tilt range. If the direct beam goes out of the chosen square, you should choose a larger one. The coordinates of the center of the beam can be used as input for XDS if used later.

5. *Peak search*: Set "Peak-B.G. Threshold", "Peak Radius", and "Background Radius". The intensity heatmap can be used to determine the "Peak Radius", i.e. how many pixels there are in a diffraction spot. The "Background Radius" defines the region over which the background intensity will be smoothed. The smoothed background is subtracted from the smoothed peaks, and intensity values above the "Peak-B.G. Threshold" are considered a peak. "Show found Peaks" highlights which spots are detected with a green cross, see Figure 5. Click run and a 3D view of the crystal should be visible in the "3D View Raw" tab. Adjust the parameters for the program to detect the desired amount of diffraction spots. The ideal parameters can vary with acceleration voltage used, camera length, and material studied.
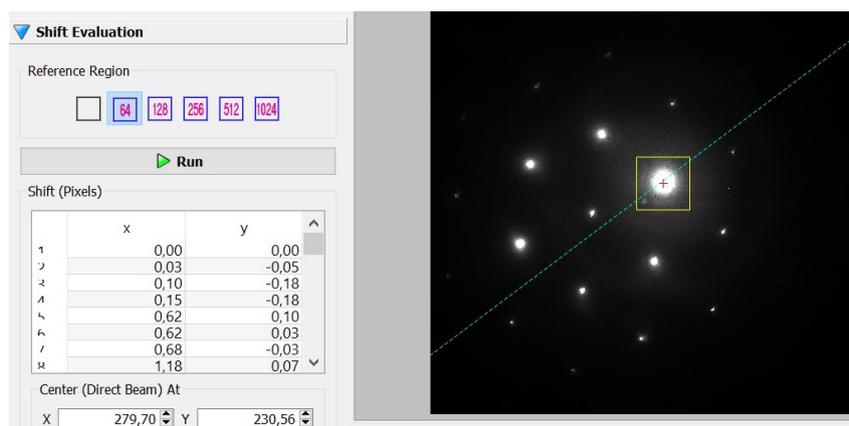
Figure 4: Beam shift correction with a reference region of 64x64 pixels shown as a yellow square in the diffraction pattern in Figure 4. The center of the direct beam is marked with a red cross.
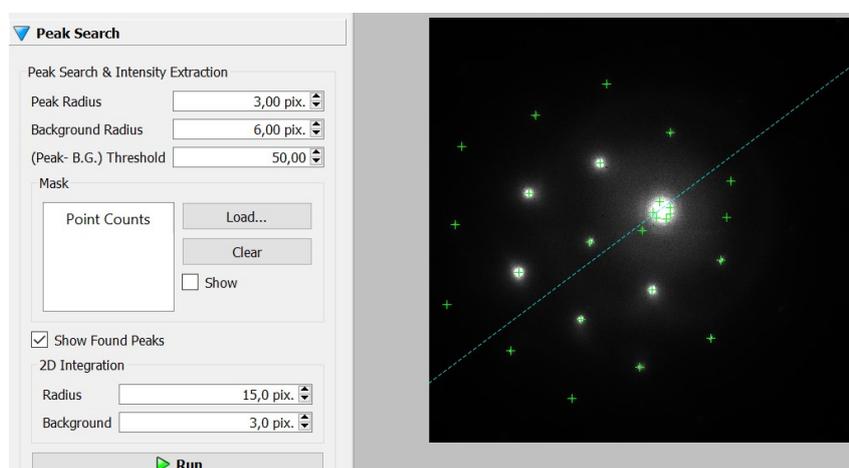

Figure 5: Peak search with found peaks visible as green crosses on the diffraction pattern in the frame.

6. *3D Peak Merging*: Peaks found in multiple consecutive frames are merged. Choose, for example, "Point" under "Intensity In Frames" and "Maximum" under "Intensity Through Frames". Adjust the values for max/min distances in 2D and 3D to merge found peaks that originate from the same diffraction spot. Change "Min. 3D Distance" from 10 pix to 5 pix if the unit cell is large. The new pattern should be visible in the tab "3D View Merged". If there are many detected spots originating from the direct beam, then the lower resolution range or intensity range may be adjusted to higher values. Too much merging can cause important reflections to disappear, although too little merging can cause a noisy 3D reconstruction, making it harder to extract important information. Turn up the "Max. Resolution Range" (from the default of 1 Å^{-1}) to see the full extent of your data resolution.
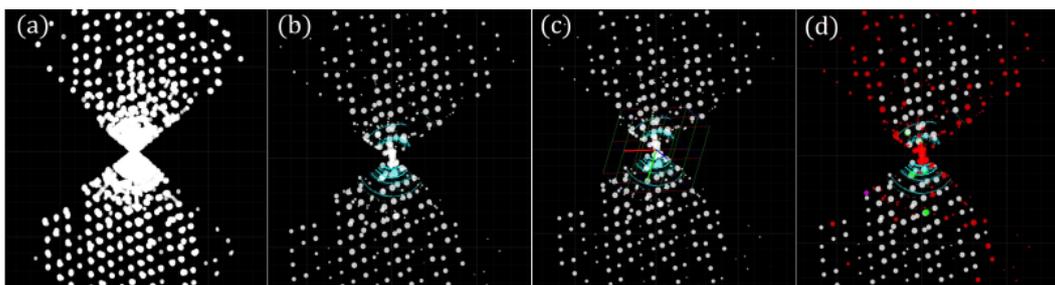
Figure 6: 3D visualizations of four stages of REDp processing: (a) raw after first peak search, (b) after merging, (c) with found primitive unit cell, (d) after indexing of correct space group. Blue spots are merged. Red spots are indexed.

7. *Unit Cell Determination and Indexing*: Click "Search for Unit Cell" to get the program to detect the positions of peaks to determine the unit cell. The resulting lattice parameters are shown in the "Messages" window, see Figure 7. Suggestions for alternative unit cells and space groups are available in a hyperlink to the CCTBX website in the output of the "Messages" window. See https://cci.lbl.gov/cctbx/lattice_symmetry.html. The space groups are listed from the highest to the lowest symmetry. Usually, one wants to choose the highest symmetry space group suggested. However, the suggested space group should not be taken for granted to be correct. Note also that the default angular tolerance of the website is 5 degrees. If the found primitive unit cell is more than 5 degrees off, potentially only lower symmetry suggestions are found (new round of merging can be necessary).
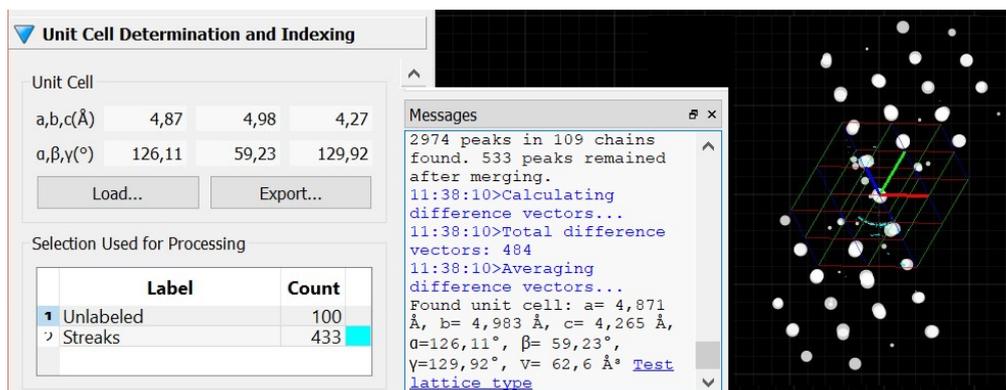
Figure 7: Detected unit cell visualized in the 3D View Merged tab and in table format in the control panel to the left.

8. *Transform unit cell:* Transform the unit cell according to the CCTBX results. In general, the highest symmetry suggested is used. Use the "Change of basis" information as input in the reciprocal "Unit Cell Transformation" in REDp as illustrated in Figure 8. The new unit cell will now appear. This transformed unit cell should be used to

find systematic absences and general reflection rules to determine the space group, which is described in section 4.2.2.


```
Input
=====

Unit cell: (4.16016, 4.10754, 4.13818, 120.852, 59.6593, 117.217)
Space group: P 1 (No. 1)

Angular tolerance: 5.000 degrees

Similar symmetries
==================

Symmetry in minimum-lengths cell: F m -3 m (2*z,x-y+z,x+y+z) (No. 225)
      Input minimum-lengths cell: (4.06974, 4.10754, 4.12783, 93.0475, 116.619, 119.199)
          Symmetry-adapted cell: (4.15237, 4.15237, 4.15237, 90, 120, 120)
             Conventional setting: F m -3 m (No. 225)
                       Unit cell: (5.87234, 5.87234, 5.87234, 90, 90, 90)
                  Change of basis: 1/2*x-1/2*y,1/2*y-1/2*z,1/2*x+1/2*z
                          Inverse: x+y+z,-x+y+z,-x-y+z
        Maximal angular difference: 3.930 degrees
```
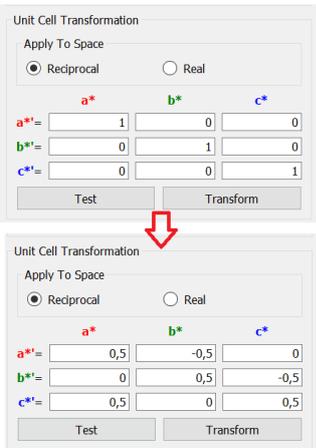
Figure 8: Transformation of unit cell from CCTBX result to REDp unit cell.

9. *(Optional) Index reflections:* If the unit cell is known, choose the right "Reflection Grouping" and click "Index Merged Peaks". "Output Indexing Results" colors the peaks that were indexed in the 3D viewer, see Figure 6(d), and presents the rest of the result in the "Peak Table" tab.

10. *(Optional) Further improvements*: The estimations of the lattice parameters can be further improved by using "Unit Cell Refinement". Other improvements related to the visualization of the three-dimensional data, such as rotation and filtering of spots, can be controlled in "3D View Control".

REDp is useful for reconstructing the reciprocal space and gives an overview of the stack of frames. Saturation of pixel intensity and other unwanted effects can be revealed, and extinction rules may hint as to which space group the material belongs to. After REDp, the structure solution tool XDS can be used for more accurate lattice parameter determination [7]. From REDp you should have found the following information to be used in XDS: average position of the direct beam (ORGX, ORGY), intensity of direct beam (OVERLOAD), rotation axis (ROTATION_ AXIS), and a tentative space group (INPUT_ SPACE_ GROUP). Input files for usage in XDS can be generated from REDp or directly during data collection via Instamatic.

Potential problems

⚠ *The green crosses do not correspond to the spots I see. How do I find the perfect parameters for peak search?*

How the reflections in the frames of a 3D ED stack are detected in REDp is determined by the "Peak Search" step. Use the parameters "Background radius", "Peak radius", and

"Peak-B.G. radius" to find the sweet spot for the detected number of reflections. The effect of the different parameters in "Peak Search" is visible in Figure 9. This figure shows the possible outcomes of to high or too low values of the three "Peak Search" parameters. A perfect number of detected spots is visible in Figure 5.
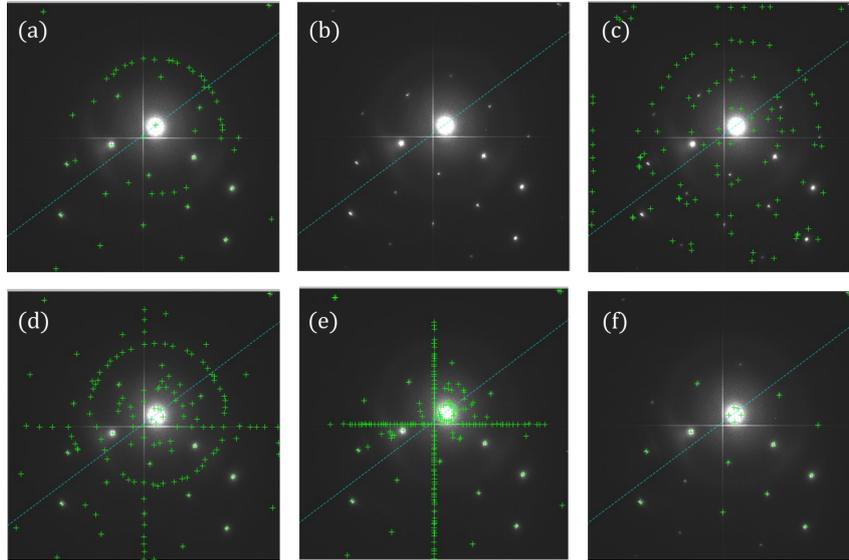

Figure 9: The effect of different parameters in "Peak Search". (a) Too high "Background Radius". (b) "Background Radius" equal to or too close to the "Peak Radius". (c) "Background Radius" too low, lower than the "Peak Radius". (d) Too low "Peak-B.G. Threshold". (e) Too low "Peak Radius". (f) Too high "Peak Radius" or threshold. The blue dashed line is the rotation axis, and the green crosses mark the detected diffraction peaks. The white cross marks the four detector segments.

⚠ *The determined lattice parameters are way larger than ideal.*
The determined lattice parameters are way larger than ideal (up to 10 times as big). This is often caused by too many fake reflections being detected close to the direct beam. Try increasing the value for the minimum resolution range while carefully watching the 3D viewer to see if the desired blobs of unwanted detected peaks disappear and not other important reflections. Heavier merging is also possible to do, increase maximum 2D distance for example. Removing initial noise with preprocessing in Python can also be an option.

⚠ *What to do when the diffraction spots near the direct beam is overexposed?*
Overexposing a material can be unwanted for many reasons. The material might be beam sensitive, and does not handle too heavy radiation for longer times. In addition, the programs used for data processing do not need too large intensity-differences in order to correctly determine where the diffraction peaks are located. Intensities around half of the direct beam (here direct beam is 4095, so half is around 2000) or less for the most

prominent diffraction spots are sufficient. Overexposure is most severe in the structure refinement step as the calculated and experimental structure factors will not match.

In order to achieve lower intensities, one has to change data collection parameters. Using a higher spot size, different condenser apertures, and changing the bit depth might correct for the issue.

⚠ *I get far less symmetric space group suggestions from the CCTBX results compared to the ground truth.*

If you get far less symmetric space group suggestions from the CCTBX results compared to the ground truth, there is reason to believe that you should change some parameters in REDp. The CCTBX-website is quite sensitive to decimals in the suggested unit cells and might suggest some different results for two quite similar unit cells.

⚠ *How can I check if I determined the right rotation axis?*

There are multiple ways to see if the rotation axis is correct or not. The first method is to run through the whole series of frames in REDp, with the rotation axis visualized, and see if the patterns behave like it moves "around" the currently set rotation axis. Secondly, when the "Peak search" is done in REDp, observe the initial raw 3D view, as this should be fairly symmetric if the rotation axis is correct. Go back to "Frames" and try changing the rotation axis to something different and observe how the order in the 3D view ceases to exist. The observed change is visible in Figure 10, where subfigure (a) has the correct rotation axis, and subfigure (b) has a wrong rotation axis.

Later, when using XDS, it is also possible to detect whether the rotation axis is correct or not, by inspecting FRAME.cbf. Under the tab 'tools' in XDSGUI choose the feature 'Show frame with predicted spots' and observe if the predicted spots align with the observed spots. Usually, a wrong rotation axis will additionally give fewer indexed peaks in IDXREF as well. The difference is visible in Figure 10, where subfigure (c) has a correct rotation axis and subfigure (d) has wrong rotation axis.

⚠ *The rotation axis is seemingly different for different datasets. Why?*

Usually, the rotation axis is constant for a given microscope manufacturer and holder. Hence, the rotation axis would not need to be determined every time a new data set is obtained, as long as the setup is the same. However, it has been observed that the rotation direction has an effect on the rotation axis. The default direction, as defined in Instamatic, is going from positive degrees to negative degrees, for example from +32 to -23 degrees. Going the opposite way, i.e. from negative angles to positive angles, can cause the rotation axis to be flipped. For the NTNU setup, this would imply that the rotation axis would go from -143° to -37°.
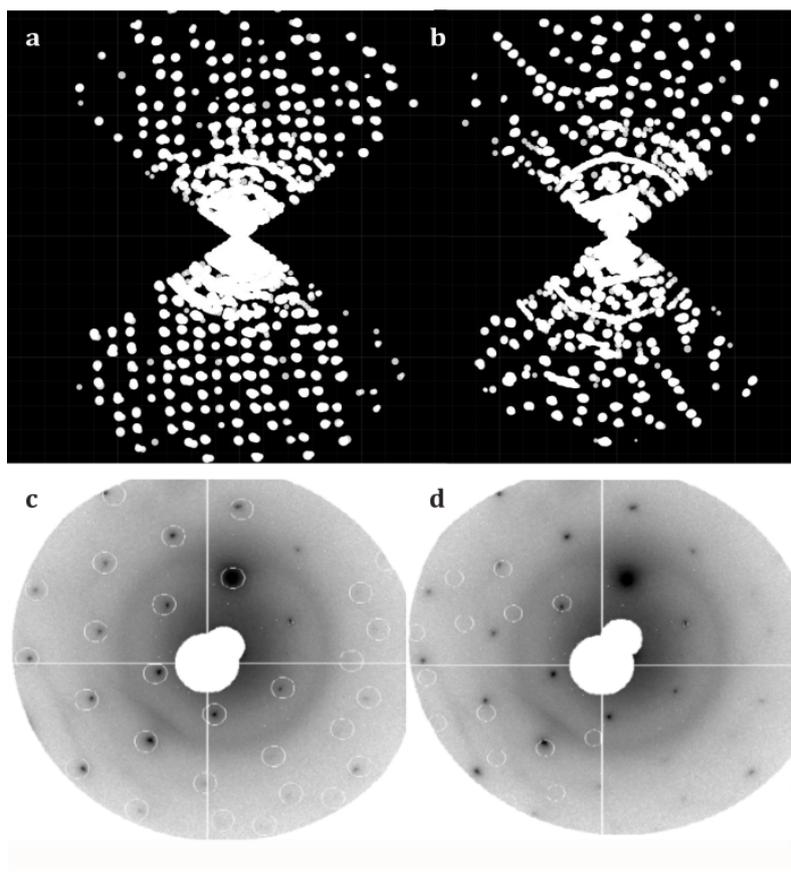
Figure 10: REDp 3D reconstruction with (a) correct rotation axis and (b) wrong rotation axis. XDS result FRAME.cbf visualization of predicted spots in a diffraction pattern with (c) correct rotation axis, and (d) wrong rotation axis.

4.2.2 Finding systematic absences

A possible way to determine the space group is by searching for systematic kinematic absences that are characteristic for certain space groups. To do this, we use REDp. More specifically we use the 3D raw reconstruction of reciprocal space, which can be viewed in the "3D View Raw" tab. In this section, dataset Cu-MOR-K-31 will be used, because the previously used dataset Si5_10 has some features that makes it unsuitable for this step (saturated intensities and untrusted rectangle). Perform the previously described steps by using the parameters in "InputREDp.txt" in the folder containing the dataset, and start here when transformation is done.

1. *Preprocess:* If there is a lot of noise or distortion in the dataset, it is preferential to preprocess the data before doing this step, because features like the untrusted rectangle can cause unwanted detected peaks. In REDp, the tab "3D View Control" will be used, to project the 3D reciprocal space in directions we desire and to make

it easier to see the extinct spots. A reasonable detected unit cell should have been determined in advance.

2. *Include all reflections:* It is important to include all reflections and therefore also the whole intensity range, therefore the "Intensity Range Min" should be set to 0 under the "3D Peak Merging" tab, and "Intensity Range Max" should be set high so that all reflections are visible, for instance by using the highest value allowed by the data bit depth (bit depth 12 allows intensities up to $2^{12} = 4096$).

3. *Tilt the view:* Click "View Along" and choose a reciprocal unit cell axis **a***, **b*** or **c***. In the example used in Figure 11, the reciprocal space is viewed along axis c*. Align the remaining unchosen directions with either the horizontal or vertical axis. In Figure 12, **b*** is aligned horizontally and **a*** is aligned vertically.
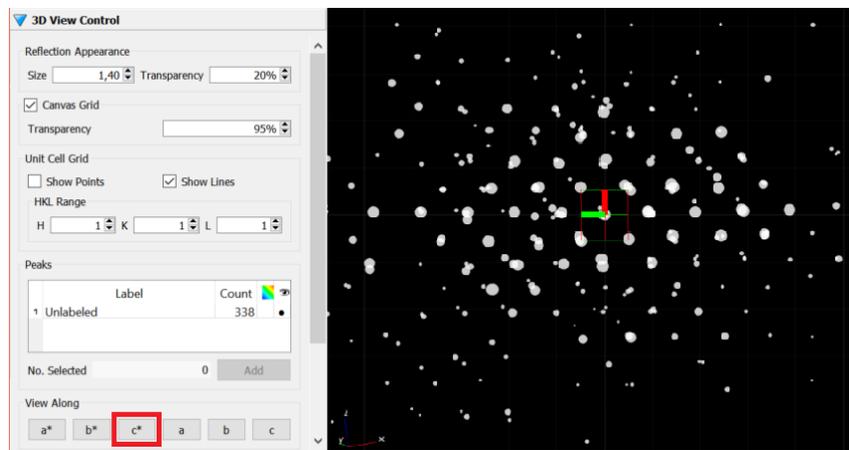
Figure 11: Choosing a view along a certain axis. Here, **c*** is chosen. The directions of reciprocal lattice vectors **a*** (red) and **b*** (green) are aligned with the horizontal and vertical axis respectively.

4. *Choose a slice:* Hold down ctrl and the left mouse button simultaneously and drag it across a row of reflections. In Figure 12 the row $0kl$ is chosen. Click "Add" in the "3D View control" to make it a separate layer of reflections. This layer can now stand out by being colored in a different color, or be isolated by hiding the other unselected reflections with a click on the eye-icon.

5. *Tilt view again and look for repeating patterns:* Change the view with the "View Along" button to view the selected slice from the last unit cell direction, along axis **a*** (which is the 0kl plane) in Figure 13. Differences in intensities reveal whether or not there exist reflection conditions in the currently viewed slice.

6. *Extend the unit cell grid:* Tick the box that says "Show Lines" and extend the hkl unit cell to a higher number to show a larger 3D cell grid.
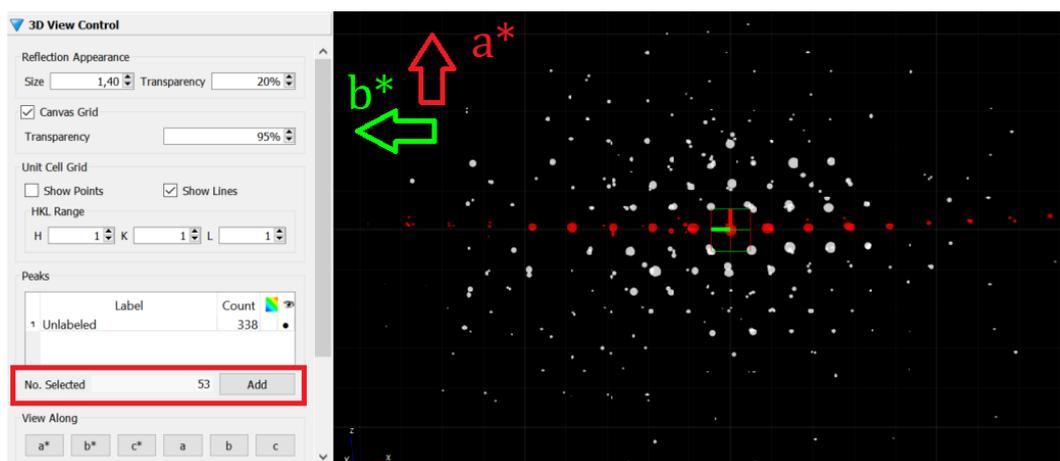
Figure 12: A slice corresponding to $0kl$ reflections were chosen and added as layer.
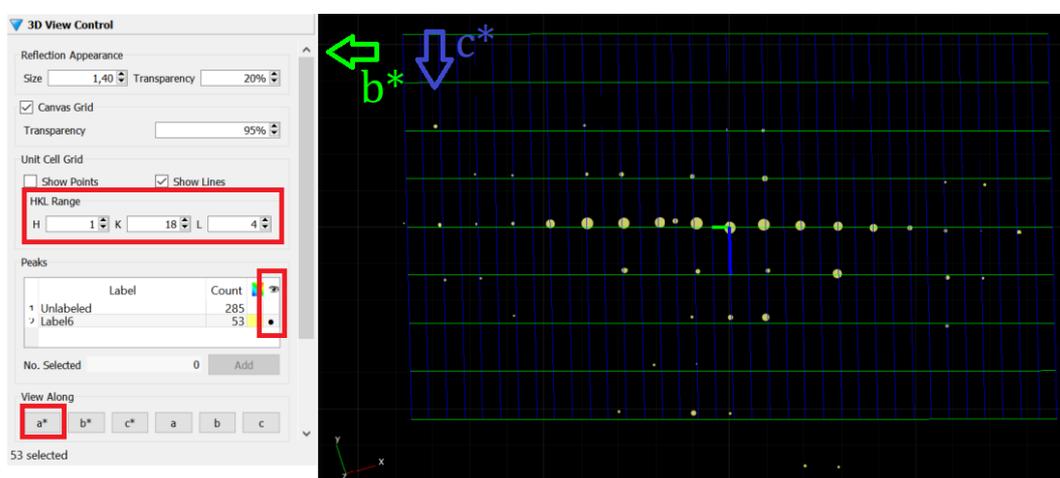

Figure 13: View along a*. Only the chosen slice is shown, the rest is hidden under the eye-button. Grid can be extended all hkl-directions to see periodicity.

7. *Find space group:* Compare the reflection conditions found with the reflection conditions listed for specific space groups in IUCr vol A tables [14]. If multiple space groups have identical reflection conditions, one would usually choose the most symmetric one. The tables are available digitally through: https://onlinelibr ary.wiley.com/iucr/itc/Ac/ch1o6v0001/sec1o6o5.pdf. An example table is shown in Figure 15.

This specific slice can be interpreted as follows: From the `cctbx` website, we know the space group Cmmm (no. 65). We can therefore look for reflection conditions for space groups of similar symmetry as this, typically located close to each other in the Tables with reflection conditions. When looking for reflection conditions, see if the diffraction spots align with the unit cell grid lines. For example, if diffraction spots only fall onto every other unit cell line, we have reflection conditions. In Figure 13, we see slice $0kl$, and observe

that there are reflection conditions in direction **b***, as diffraction spots fall onto every other blue line. In direction **c***, there are diffraction spots for every green line, hence no reflection conditions. For $0kl$ we hence have reflection condition k, which, by observing the table in 15, rules out the space groups in the four lower rows. Further inspecting new slices, like $0kl$ and $h0l$ seen in Figure 14, by following the latter described method, more reflection rules are revealed. For $h0l$, reflections rules for both l and h are observed, despite few reflections being present. For $hk0$, a new periodic pattern is revealed, indeed, the reflection conditions for $hk0$ is $h + k$. One can confirm this by observing the row of l-reflections for $k = 0$, and then see how the same line of l-reflections are shifted by one vector length for $k = 1$.
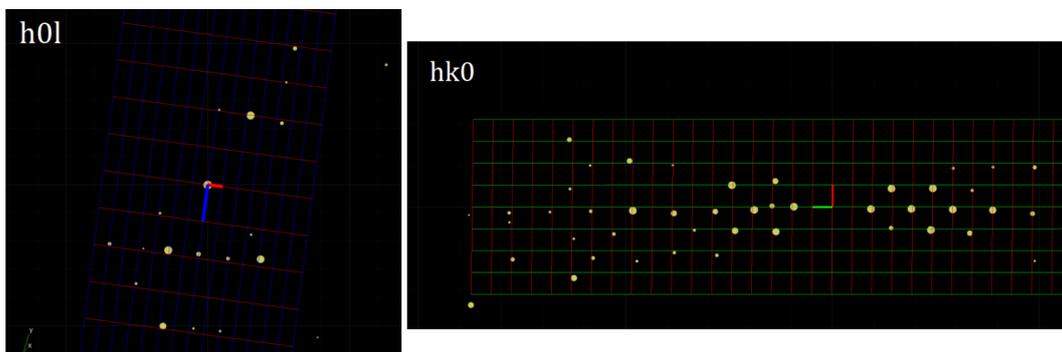

Figure 14: Slice $h0l$ and $hk0$. a* is the red grid, b* is green and c* is blue.

Observing the table in Figure 15 once again, there is only one row of space groups fitting the current observations, that is marked with a red square. In this row, there are three space groups. As mentioned previously, we choose the most symmetric one, i.e. Cmcm (no. 63). This is also the space group that will be used for further studies of MOR2 in the next programs.

Potential problems

⚠ *The expected extinction spots are not completely gone, they have some intensity.*
Because of dynamical effects, spots that should be extinct might still have some intensity. It might still be possible to distinguish between rows of high intensity versus rows of low intensities, although specific patterns where only single spots should be extinct can be tough to see. Alternatively, one can look for these extinction rules in single frames in the stack instead.

⚠ *There are only a few reflections present in my slice, I can't seem to see far enough out to determine whether or not there are repeating reflection conditions.*
There are only a few reflections present in my slice, I cannot seem to see far enough out to determine whether or not there are repeating. Try adjusting the maximum resolution range to higher values to include all reflections. This issue is especially tricky for materials with small unit cells, and hence reflections with large distances. In total, one might not

Reflection conditions							Space group	No.	Space group	No.	Space group	No.
hkl	$0kl$	$h0l$	$hk0$	$h00$	$0k0$	$00l$						
$h+k$	k	h	$h+k$	h	k		**C222** C2mm	21 38	Cmm2 Cmmm	35 65	Cm2m	38
$h+k$	k	h	$h+k$	h	k	l	**C222$_1$**	20				
$h+k$	k	h	h,k	h	k		Cm2e	39	C2me	39	Cmme	67
$h+k$	k	h,l	$h+k$	h	k	l	Cmc2$_1$	36	C2cm	40	Cmcm	63
$h+k$	k	h,l	h,k	h	k	l	C2ce	41	Cmce	64		
$h+k$	k,l	h	$h+k$	h	k	l	Ccm2$_1$	36	Cc2m	40	Ccmm	63
$h+k$	k,l	h	h,k	h	k	l	Cc2e	41	Ccme	64		
$h+k$	k,l	h,l	$h+k$	h	k	l	Ccc2	37	Cccm	66		
$h+k$	k,l	h,l	h,k	h	k	l	Ccce	68				

Figure 15: Table with reflection conditions from ITA. Found space group in red frame.

have that many reflections at all through a whole series of 3D ED. This method might not be suitable for the material then. A generally big setback of this method is that if REDp has determined an incorrect unit cell, then rotating the 3D view along any of the reciprocal unit cell axes will not make any sense.

4.2.3 How to use XDS

First thing to do in order to be able to use XDS is to get hold of a computer with a Linux operating system or use a virtual machine to access the Linux operating system on your own computer. Follow the instructions from this site `https://wiki.uni-konstanz.de/xds/index.php/Installation`.

XDS requires an instructional input file, `XDS.INP`, and each separate frame saved in `.img` format. An example `XDS.INP`-file along with a list of important parameters for this file is available in Appendix A. The frames of format `.img` do not necessarily need to be in the same folder as `XDS.INP`, as their path can be specified in the `.INP`-file. For a more intuitive visualization of the XDS results, XDSGUI can be used [8]. In XDSGUI, all output files are presented in separate tabs, and the frames can be inspected individually. This manual assumes usage of XDSGUI.

1. *Run* `xdsgui` *in terminal:* XDS is terminal-based and can be called by its name `xds` or `xdsgui` in the terminal where your `XDS.INP` input file is located. The program will read your input from the `XDS.INP`-file, it is therefore important that the input information is correct in order for the program to function correctly. XDS then runs background estimations, peak search, unit cell determination, indexing, scaling, and integration.

2. *Set up* `XDS.INP`: If Instamatic is used during data acquisition, a `XDS.INP`-file will be automatically generated. It is also possible to create a default `XDS.INP` in `xdsgui`. A single frame from the dataset can be used as input to generate this default `XDS.INP`. Either way, the file will have to be edited in order to work properly based on correct information. `XDS.INP` can be edited directly in `xdsgui` or in any text editor. The most important parameters to include are shown in Table 2. For a complete list of all XDS parameters and their functionality, see `https://xds.mr.mpg.de/html_doc/xds_parameters.html`. Edit NAME_ TEMPLATE_ OF_ DATA_ FRAMES to be the path to the folder where the `.img`-files are. Be sure to correct the wavelength and insert the calibrated detector distance. Information about the position (ORGX, ORGY parameter) and intensity (OVERLOAD parameter) of the direct beam can be found from REDp, described previously in Section 4.2.1. The rotation axis in XDS is defined differently than in REDp. If the angle of the rotation axis in REDp is given as $\theta°$, then the rotation axis parameter that is inputted in XDS is given as the vector $[\cos(\theta) - \sin(\theta)\ 0]$. Remember to click 'save' and 'run xds' when the input file is done.

3. *Inspect input frames:* In `xdsgui` one can view all the individual frames under the tab 'Frame'. Here, some of the input parameters of `XDS.INP` are visible, such as the "UNTRUSTED RECTANGLE" marked with red lines, the "RESOLUTION RANGE" marked in green and "TRUSTED REGION" marked in blue, see Figure 16. The center

Table 2: List of important parameters, their meaning and the value used in 'XDS.INP'.

Parameter	Description	Typical value	Automatic
SPACE_GROUP_NUMBER	Space group number determined by REDp.	227	no
UNIT_CELL_CONSTANT	Conventional unit cell parameters from REDp.	5.43 5.43 5.43 90 90 90	no
OSCILLATION_RANGE [°]	Rotation step (end angle - start angle)/#frames.	0.1593	yes
ROTATION_AXIS	Defined as a vector [$\cos(\theta)$, -$\sin(\theta)$, 0] where θ is the rotation axis from REDp.	[0.7986 -0.6018 0.0]	no
ORGX, ORGY [pixel]	Position of the direct beam. Often close to the center of the image.	274.79, 235.76	yes, but might need correction
DETECTOR_DISTANCE [mm]	Virtual distance between intermediate and projector lens. Needs to be calculated from the physical pixel size and calibrated camera length.	245.24 (for 120 cm) 273.40 (for 150 cm)	no
QX, QY [mm]	Physical pixel size.	0.0550, 0.0550	yes
X-RAY_WAVELENGTH [Å]	Wavelength of the accelerated electron.	0.041757 (for 80 kV)	yes, but might need correction
INCLUDE_RESOLUTION _RANGE [Å]	Lowest and highest data resolution.	10 0.7	yes, but might need correction
MAXIMUM_ERROR_OF _SPOT_POSITION	Max. deviation from the ideal diffraction spot position	8.0	no
MAXIMUM_ERROR_OF _SPINDLE_POSITION	Max. deviation from ideal spindle angle	4.0	no
DELPHI [°]	Degrees of spindle rotation, controls the number of learned profiles used in INTEGRATE	5	no

of the direct beam is marked with a green cross, and its position is specified by the parameters ORGX and ORGY in XDS.INP which can also be found in REDp during shift correction. Make sure that the untrusted rectangle is covered by the red lines, or else adjust the parameters. Turn the higher value of the trusted region or resolution range up if there is much noise around the direct beam. Then a larger radius of the inner blur or green ring should appear.
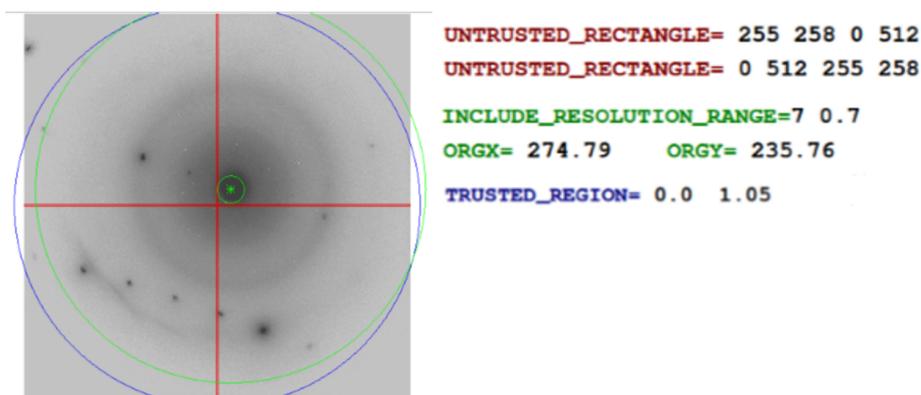


```
UNTRUSTED_RECTANGLE= 255 258 0 512
UNTRUSTED_RECTANGLE= 0 512 255 258

INCLUDE_RESOLUTION_RANGE=7 0.7
ORGX= 274.79    ORGY= 235.76

TRUSTED_REGION= 0.0  1.05
```

Figure 16: Single frame with colour-coded parameters marking the untrusted rectangle (in red), the resolution range (in green), the trusted region (in blue) and the center of the direct beam marked with a green cross.

4. *Inspect output file IDXREF:* First see if the 2D reconstruction of reciprocal space is filled with mostly red or black diffraction spots. Figure 17 shows an example where most of the spots are indexed and one where only a few of the spots are indexed. The output file also says how many spots were detected and indexed, and further what unit cell it detects, both primitive and conventional. Suggestions are given for the space group of the primitive unit cell. If there is any error during refinement or indexing of spots, this is the file that crashes. Bad data quality or crystals with very small unit cells might cause problems of too few spots being indexed.

5. *Inspect output file INTEGRATE and adjust SCALE:* In xdsgui, the INTEGRATE tab views the log-file INTEGRATE.LP on the left and graphs on the right. This log-file is very technical and describes the profile fitting process XDS does, although in most cases one does not need to bother with this. What is interesting about this log-file is the final number of reflections that are saved to INTEGRATE.HKL. The XDS.INP-parameter DELPHI[3] (default is 5) has a large effect on this result and can be tuned down if there are problems of too few reflections being detected. The graphs on the right should also be inspected. One would want smooth graphs with few or no large spikes or rapid change in values. Large spikes in the SCALE-graph,

---

[3]Defined as number of degrees of spindle rotation, controls the number of learned profiles used in INTEGRATE, see https://xds.mr.mpg.de/html_doc/xds_parameters.html
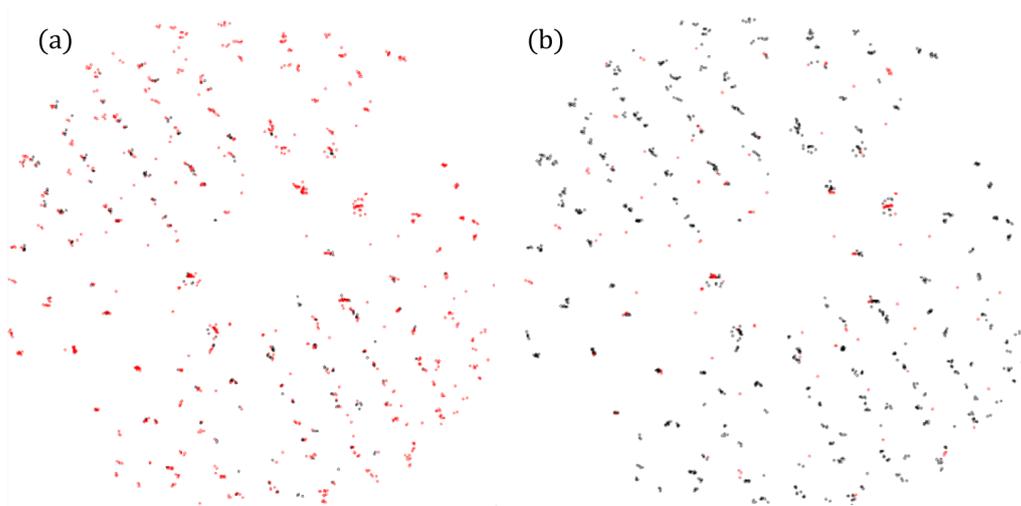
Figure 17: Visualizations of the reciprocal space in XDS of the Si data with (a) few indexed peaks, and (b) many indexed peaks. The spots turn black if indexed, red if not.

for example, can be corrected by excluding the frames that caused the spiking, see Figure 18. This often happens when the defocused frames are not properly removed.
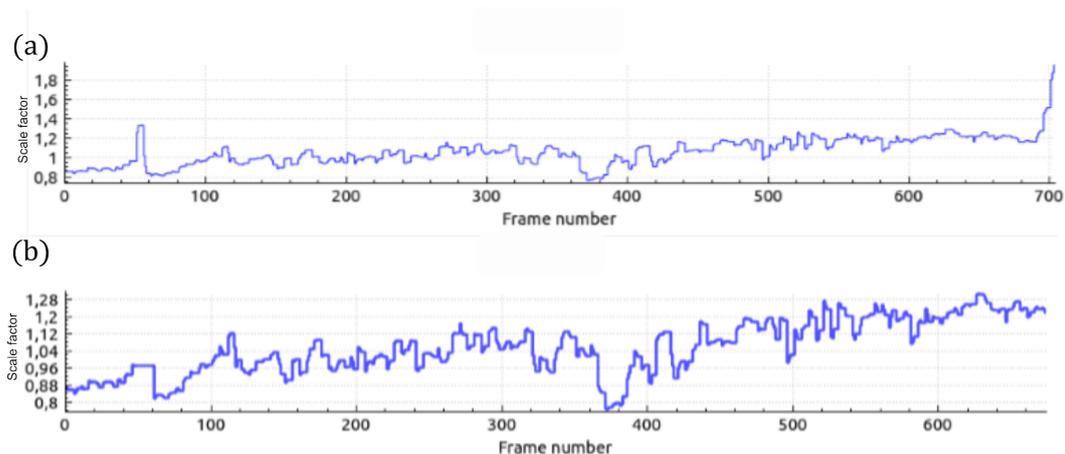

Figure 18: The SCALE from INTEGRATE (a) before exclusion of frames, (b) after exclusion of frames that caused the tall peaks in scale graph (a).

6. *Inspect output file CORRECT:* CORRECT contains a series of graphs presenting statistical data, among these are data completeness, $CC_{1/2}$, R_{meas} and I/sigma. In `xdsgui`, the CORRECT tab views the log-file CORRECT.LP on the left and the corresponding graphs on the right. This step of XDS refines the unit cell, applies empirical corrections on the intensities during integration, and scales the data. A summary of all the statistics can be found in the last table (scroll almost all the way down in the log-file) see Figure 6. High data completeness, I/SIGMA and $CC_{1/2}$, with

low R_{meas} is preferable. The asterix next to the $CC_{1/2}$ means that the correlation is significant.

```
SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION
RESOLUTION     NUMBER OF REFLECTIONS    COMPLETENESS R-FACTOR  R-FACTOR COMPARED I/SIGMA   R-meas  CC(1/2)  Anomal  SigAno    Nano
   LIMIT     OBSERVED  UNIQUE  POSSIBLE    OF DATA   observed  expected                                      Corr

    2.35        229      72        81       88.9%       8.6%      7.4%      229   14.22     10.6%   98.0*      0    0.000       0
    1.68        366     116       130       89.2%       6.8%      7.5%      362   11.75      8.4%   99.5*      0    0.000       0
    1.38        447     137       149       91.9%       6.1%      7.9%      445   11.13      7.4%   99.4*      0    0.000       0
    1.20        579     173       184       94.0%       9.2%      8.6%      578    9.86     10.9%   98.8*      0    0.000       0
    1.07        642     181       193       93.8%      13.5%     12.4%      635    7.11     16.0%   99.5*      0    0.000       0
    0.98        702     207       217       95.4%      17.0%     16.4%      695    5.61     20.2%   95.5*      0    0.000       0
    0.91        832     227       236       96.2%      19.1%     20.6%      827    4.50     22.6%   97.4*      0    0.000       0
    0.85        847     238       246       96.7%      29.4%     32.6%      837    3.32     34.7%   93.6*      0    0.000       0
    0.80        872     241       265       90.9%      25.8%     33.5%      860    2.97     30.3%   96.4*      0    0.000       0
   total       5516    1592      1701       93.6%       8.8%      8.7%     5468    6.65     10.7%   98.7*      0    0.000       0
```

Figure 19: The final table of the CORRECT step, showing a summary of the data statistics. The most important characteristics are marked with a red box.

7. *Insert unit cell details:* Input lattice parameters found by XDS in the space group found by REDp analysis (add UNIT_ CELL_ CONSTANTS parameter and SPACE_ GROUP_ NUMBER parameter) and run the CORRECT step once more.

8. *Inspect output file FRAME.cbf:* Under the tab 'tools' there is a feature called 'Show frame with predicted spots' where the output file FRAME.cbf is used to project the predicted diffraction patterns on top of the actual frames. Choose a frame number with an easily recognizable pattern or any pattern with clearly visible spots to see if the predictions fit the obtained data frames. This can be useful to determine whether or not the suggested REDp space group used as input is sufficient; see Figure 20. The wrong rotation axis can also be exposed by inspecting this pattern.

9. *Convert hkl-file to SHELX-format with XDSCONV:* When XDS has successfully run, a reflection data file (`.hkl`) will be produced as output (along with many other files). This is typically called XDS_ ASCII.HKL and is *NOT* of the ideal format for being directly plugged into SHELXT. To remake the `.hkl`-file into a useful format, click the XDSCONV-tab in `xdsgui` and change the text 'CCP4_ I+F' to 'SHELX'. Then click save and run XDSCONV. An example input file, XDSCONV.INP, is shown in Listing 1. The last line mentiones Friedel's law. The name of the desired output `.hkl`-file might also be changed if desired. The `.hkl`-file should contain five columns of the format $h\ k\ l\ F_o^2\ \sigma(F_o^2)$, the last parameters being the structure factor and the variance of the structure factor.

```
1  INPUT_FILE=XDS_ASCII.HKL
2  OUTPUT_FILE=MOR1.hkl SHELX
3  FRIEDEL'S_LAW=TRUE
```

Listing 1: Example input file XDSCONV.INP.

The third line in the listing, mentions Friedel's law. This law applies only to centro-symmetric unit cells. This law states that intensities of "opposite" indices, like (h, k
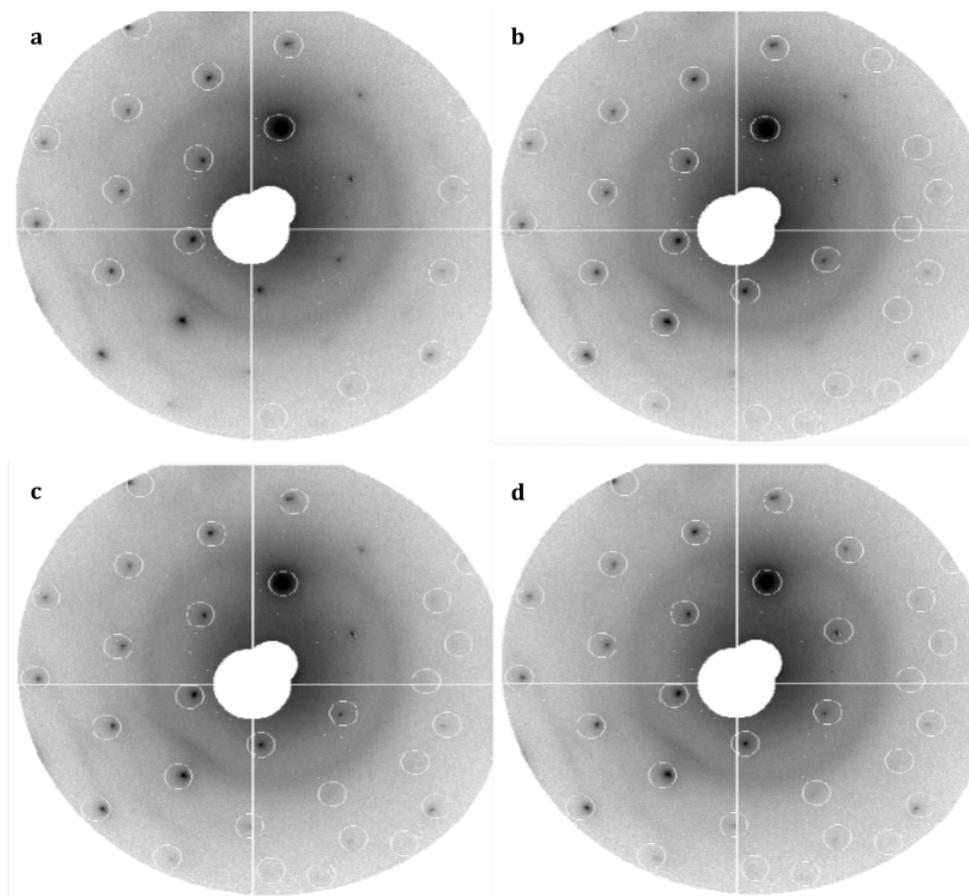
Figure 20: The effect of different input space group on the expected detected spots of a single frame from a Si data set run in XDS with FRAME.cbf. Sorted from lowest to highest space group. (a) Space group no. 139, I4/mmm. (b) Space group no. 166, R$\bar{3}$m. (c) Space group no. 225, Fm$\bar{3}$m, the highest symmetry guess from REDp. (d) The expected space group Fd$\bar{3}$m, no. 227. Visibly, only the correct space group predicts all the visible diffraction spots. The center is here masked out to avoid noise from central beam.

,l) and (-h, -k, -l) often referred to as a Friedel pair, are about equal. This can cause the real crystal to have lower symmetry than the result of a diffraction experiment would imply [15]. Friedel's law can be set to FALSE to include all reflections.

Potential problems

⚠ *The program crashed and displays an error message in the terminal.*

There are several possible errors in XDS. This manual will not cover all of them, but they are mostly covered in the troubleshooting pages of the XDS website.

See: `https://wiki.uni-konstanz.de/xds/index.php/Problems`

⚠ *Many reflections are wrongfully detected around the direct beam.*

This might not be a big issue, but the program can occasionally crash and display an error-message. In the best case, the program will simply detect the peaks, but not index them. To solve the issue, raise the lower value of the TRUSTED_ REGION to cover the area with many incorrectly detected spots. Alternatively, turn down the higher value of INCLUDE_ RESOLUTION_ RANGE to make the inner green circle larger. It should cover the whole direct beam.

⚠ *Too few reflections are indexed in XDS.*

Usually, this will be visible either in the 2D visualization of the detected spots as many red spots instead of black spots, or the percentage of indexed spots is lower than the set threshold (which by default is 50) which makes the program crash. First, double check if your input parameters are wrong, for example, did you determine the rotation axis properly? The problem can be solved by turning up parameters like MAXIMUM_ ERROR_ OF _ SPINDLE_ POSITION and MAXIMUM_ ERROR_ OF _ SPOT_ POSITION. The threshold (MINIMUM_ FRACTION_ OF_ INDEXED_ SPOTS) can also be modified to lower values to make the program run despite a low indexing percentage.

⚠ *The program runs fine until it gets to INTEGRATE/CORRECT where there are few to no reflections saved to .hkl file.*

Double check that enough reflections are being indexed, if not follow the quidelines of the previous issue. Then try turning the DELPHI parameter down (default is 5) to maybe 3 or 2.

Merge data with XSCALE (optional): If your data has low completeness, it can be beneficial to merge several datasets together to obtain higher completeness. This can be done in XSCALE, which is part of XDS and is available as a tab in XDSGUI next to XDSCONV. The datasets to be merged need to be solved in the same space group but can have different data collection parameters. Furthermore, the lattice parameters may be slightly different, but the axes *a*, *b*, and *c* should be the same in asymmetric unit cells. An example input file, XCALE.INP, is shown in Listing 2. The unit cell lengths are given in Å and the angles in degrees.

```
1  SPACE_GROUP_NUMBER= 63              !Shared space group
2  UNIT_CELL_CONSTANTS= 18.65 20.69 7.74 90.0 90.0 90.0  !Average unit
      cell constants
3
4  OUTPUT_FILE=merged.ahkl
5
6  ! List of all inputfiles to be merged
7  INPUT_FILE=../mordenite_cRED_1/SMV/XDS_ASCII.HKL
8  INPUT_FILE=../mordenite_cRED_2/SMV/XDS_ASCII.HKL
```

Listing 2: Example input file XSCALE.INP.

Here is how it is used:

1. (Optional) Copy the XDS_ ASCII.HKL file for all datasets you want merged to one folder.

2. Write the path to all XDS_ ASCII.HKL files that should be merged.

3. Input averaged unit cell parameters and the common space group.

4. Click 'Save' and 'Run XSCALE'. The program will now produce a hkl-file containing all reflections from the merged datasets, all brought to the same scale.

5. Run XDSCONV again to convert the merged hkl-file to SHELX format.

4.3 Batch dataprocessing with EDtools (optional):

EDtools is a Python-package. Therefore, one needs to open a Python-terminal (for example Anaconda Powershell) to run this program. EDtools needs to access XDS, and hence requires XDS to be installed on the same computer. The process described here is described in more detail in the README for the EDtools Github documentation https://github.com/instamatic-dev/edtools.

Before starting, copy all projects to one folder and navigate to this path in the Anaconda terminal. Further run the following commands,

1. *Autoindex:*

   ```
   edtools.autoindex
   ```

   All XDS.INP files in subdirectories will be found and run in XDS. This step might therefore take some time.

2. *Extract info:*

   ```
   edtools.extract_xds_info
   ```

   An excel file with the most important statistics from the XDS-run is produced, along with the files `cells.yaml` for usage in the next steps and `files.txt` that contain a list of files with good completeness/CC(1/2) statistics.

3. *Find cell mean and cluster:*

   ```
   edtools.find_cell cells.yaml --cluster
   ```

   Produces a dendogram, where one can choose a cutoff (visible as a straight line) by clicking in the dendogram. All datasets below the cutoff (coloured orange in the dendogram) will be kept for further analysis. The command will find the mean unit cell parameters.

4. *Make a XSCALE input file:*

   ```
   edtools.make_xscale cells.yaml -c 10.0 20.0 30.0 90.0 90.0 90.0 -s"65"
   ```

   Input mean unit cell parameters, as found in the previous step and input space group. Use the number of the spacegroup and not the letters, and remember the quotation marks ("").

5. *Run XSCALE and XDSCONV:* Use the generated XSCALE-file from the previous step and run it in XDSGUI to merge the data. Then, use XDSCONV to convert the merged hkl-file to SHELX format.

4.4 Structure solution

Structure solution methods aim to solve the crystallographic phase problem, by reconstructing the phase from diffraction information. There are several possible algorithms developed for this purpose. In this workflow, SHELXT is chosen. SHELXT is based on intrinsic phasing, a novel dual-space algorithm that accounts for the phase problem for single-crystal reflection data [9].

4.4.1 How to make input files for SHELXT

A `.hkl`-file is not enough for SHELXT to run, as it needs an additional instructional input file, namely an `.ins`-file. To prepare an input `.ins`-file for SHELXT one can either make one from scratch, use XPREP or the Python-based library EDtools.

XPREP/EDtools takes the lattice parameters and space group as input parameters, makes a SHELXT input file containing the symmetries of the specific space group, and adds information about the number of atoms in a unit cell.

XPREP method: Run `xprep NAME` in terminal, where NAME is the name of the .hkl-file produced by XDSCONV in XDSGUI. A new window will appear; see Figure 21, where the reflections will be read from the .hkl-file and one will be asked about the lattice parameters of the material.

Figure 21: Interface of XPREP.

The program will now give the user a series of questions or options. It will always provide a guess as to what the user should or will do next given in brackets []. If the answer it suggests in brackets is the right one, simply press enter. The following steps should be done.

1. Input information about space group

2. Search for higher metric symmetry (option [H] from main menu). Choose the highest symmetry or best suited one. Use the PgUp and PgDn buttons to navigate throughout the whole list of symmetries.

3. Determine or input space group (option [S] from main menu). Choose determine space group and input (again) the space group information you possess.

4. Read, modify or merge datasets (option [D] from main menu). Here, a new menu will appear. First, choose option [S] Display intensity statistics, then [A] Merge all equivalents. Exit to main menu.

5. Define unit cell contents (option [C] in the main menu).

6. Set up SHELXTL files (option [F] in the main menu).

7. (Optional) Overwrite old .hkl-file.

EDtools method: Open a python terminal and run the command:
`edtools.make_shelx -c (unit cell parameters) -s (space group) -m (composition)`
For a MOR unit cell of lattice parameters $a = 18.11$, $b = 20.53$, and $c = 5.43$ Å and $\alpha = \beta = \gamma = 90°$, the command would therefore look like:
`edtools.make_shelx -c 18.11 20.53 7.528 90 90 90 -s"63" -m O96 Si48`
Use the number of the space group and not the name. The number should be wrapped with quotation marks ("").

When an `.ins`-file is successfully produced from XPREP or EDtools, some parts of the file still need to be edited. An example `.ins`-file is shown in Figure 22. The different parameters are explained separately. The parameters highlighted in yellow need to be modified after XPREP/EDtools preparation. The wavelength is set to default values by XPREP (assuming X-rays) and EDtools (assumes 200 kV electrons). Insert the correct wavelength and for XPREP files, one needs to additionally get the electron scattering factors for each element present in the unit cell from this website https://srv.mbi.ucla.edu/faes/. The website provides information about element-specific scattering factors directly in SHELX format [16]. The LATT-parameter in the INS-file, should be what lattice type the material has. Positive numbers are centrosymmetric, while negative numbers indicate non-centrosymmetric materials. The magnitude of the LATT-number is the centering where P=1, I=2, 3 is rhombohedral obverse on the hexagonal axes, F = 4, A = 5, B = 6, and C = 7. The SYMM-parameter is the symmetry operations, given as the coordinates of the general positions as given in the International Tables of Crystallography [14].

```
TITL filename in Cmcm

          wavelength    lattice parameters
CELL 0.04176 18.11 20.53 7.528 90 90 90
number of formulas per unit cell
               standard deviation of unit cell
ZERR 1 0.0036 0.0041 0.0015 0 0 0
LATT 7  Lattice type, negative if non-centrosymmetric.
SYMM -X,-Y,0.5+Z
SYMM -X,+Y,0.5-Z     Symmetry operations
SYMM +X,-Y,-Z                         SFAC: element symbol and electron scattering factors
SFAC  O 0.143 0.305 0.51 2.268 0.937 8.262 0.392 25.665 0 0 0 0 0.63 15.999
SFAC  Si 0.363 0.428 0.974 3.557 2.721 19.39 1.766 64.333 0 0 0 0 1.16 28.086
UNIT 96 48   Number of atoms of each element in unit cell
TREF         Direct method used
HKLF 4       Standard SHELX reflection data file used
END
```

Figure 22: Example of an .ins-file for use in SHELXT. Different parameters are explained separately, whereas the parameters highlighted in yellow need to be modified after XPREP/EDtools preparation.

4.4.2 How to use SHELXT

SHELXT can be used both directly in a terminal and in Olex2. When applied directly through a terminal, the program needs to be in the same folder as the `.ins` and the `.hkl`-file it will use. These two files need to have the same name. Open a terminal in this folder and type `shelxt NAME` to launch the program, where `NAME` is the name of the `.ins` and `.hkl`-file.

To use SHELXT in Olex2, one simply has to move the program into the system folder for Olex2 (where the rest of the system files for the Olex2 download are located). Click the button Work → Solve, then choose SHELXT as the tool option to solve the structure. The interface of Olex2 is shown in Figure 23. The output of the SHELXT result will be shown
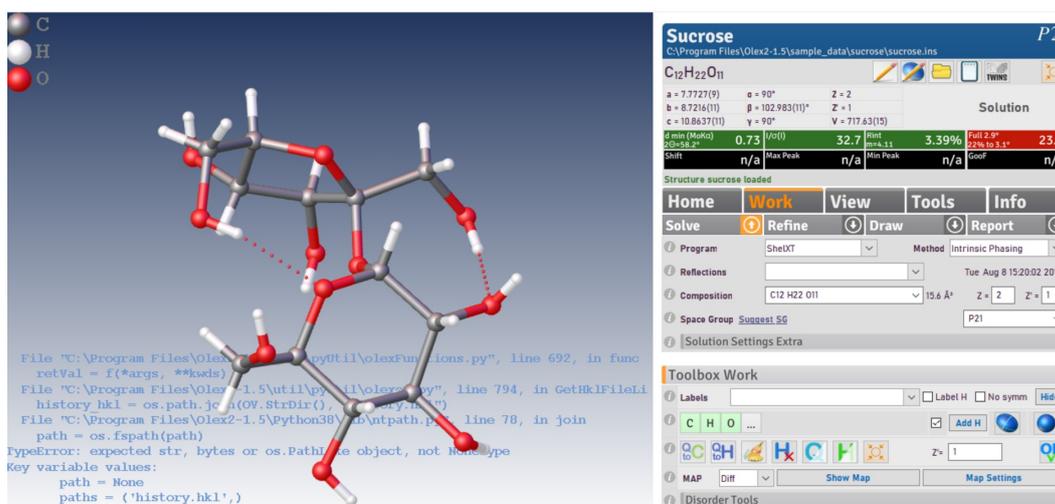

Figure 23: Olex2 GUI with test-structure sucrose.

in the terminal output to the left (blue text) and saved to a `.lxt`-file. The solution will be saved to a `.res`-file and the 3D structure will be built in the Olex2 structure viewer window.

If the structure solution fails or the wrong space group is detected, consider using the command `shelxt -s"spacegroup"`, which essentially forces a solution in the given space group.

An overview of all possible SHELXT keywords that can be used is available in `https://shelx.uni-goettingen.de/shelxt_keywords.php`.

Potential problems

⚠ *The program stops after a few seconds, with no final message of a solved structure.* If you have problems with SHELXT stopping after a few seconds, with no final message of a solved structure, then it indicates that the structure was not solved and something went wrong. Usually running SHELXT in Olex2 would show more informative error-messages than if SHELXT is run in terminal. The problem can be a bad input file or missing information in the `.ins` file. This can happen if one for example forgets to input the electron scattering factors and do not correct for the wavelength in the *.ins*-file directly produced from XPREP.

⚠ *SHELXT says I have a bad input file.*
This might happen if one forgets to use XDSCONV on the output `.hkl`-file from XDS or the SHELX-`.hkl`-file is somewhat not of the right format. One can perfectly fine run through XPREP even with an improper `.hkl`-file, SHELXT is then the first program that notifies you that something is wrong. Although there are some signs in XPREP that indicate error for example if the initial mean of intensities is 0, or the step "Determine space group" does not work.

4.5 Refine structure

4.5.1 How to use Olex2

The Olex2 software has many features and it is easy to get lost. There are buttons for nearly everything, but often it will be easier to simply type the desired command in the command-line in the lower left corner. For a complete list of all Olex2 commands and their functionality, see https://www.olexsys.org/categories/commands/. The most important commands (in my opinion) and their functionality is summarized in Table 3.

Table 3: List of important Olex2 commands and their functionality.

Command	Description
grow	Makes the structure larger, i.e. displays more atoms. Can alternatively use -s or -w.
pack cell	Fills the unit cell with atoms.
lines n	Show n lines of the terminal output. Lower n gives more space for the 3D view of molecule.
ctrl + R	Refines the structure. Can also be done by clicking the refine button.
ctrl + Q	Hide / show Q-atoms. These atoms are suggestions for extra atom positions in the refiend model.
fuse	Show only anisotropic unit.
ctrl + M	Show electron density map.
kill	Remove marked atom(s) or binding(s).
F2 / F4	Change background. F2 is mono-coloured (toggle between white / dark blue). F4 is graduated which is the default.

The Olex2 GUI is shown in Figure 23. Olex2 has a structure viewer window with a command line underneath. Here, the terminal output will be shown. The right side consists of a control panel and an overview of the input information at the top.

4.5.2 Refinement with SHELXL

SHELXT provides an initial model for the structure. Now, to validate that this result is, in fact, useful, one needs to refine the initial model structure to a (ideally) global minimum. Only then we will get a final and reliable solution.

Olex2 has many options for different refinement tools. The program has its own refinement tool called `olex2.refine`, although other tools are also available. SHELXL is another program for refinement, developed by the same creator as for SHELXT [10]. To use this in Olex2, one needs to do the same as with SHELXT previously; move the program to the system-folder for Olex2. In Olex2, the SHELXL refinement tool is found in Work → Refine (arrow down) → Program (dropdown), see Figure 24.

Figure 24: Refinement options in Olex2. Black box (a) shows the color-coded refinement statistics. (b) is the report-button that produces a list of measured quantities of the model. (c) refinement parameters EXTI and ACTA. (d) battery showing the parameter-to-reflection ratio.

Usually refinement would involve pressing the "Refine" button or `ctrl + R`, see Table 3 for more commands, until the structure shift converges to a minimum. Carefully observe the statistics in the boxes on the right, such as R_1, wR_2, $Shift$, $GooF$, $MaxPeak$, and $MinPeak$, which will decrease if the refinement proceeds successfully.

R-factors, that are used to evaluate the fit of the model to the experimental data can take much larger values for ED than for XRD. This is because the refinement tools assume kinematical diffraction, while knowing that ED can cause dynamical effects. R_1 between 10-30% is normal. Olex2 calculates and displays different R-factors calculated by SHELXT or SHELXL, see Figure 24(1). These are important for measuring different aspects of the

input data. R_{int} is defined as follows,

$$R_{int} = \frac{\sum |F_o^2 - F_o^2(mean)|}{\sum F_o^2} \tag{2}$$

while R_1 and wR_2 are defined as

$$R_1 = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \tag{3}$$

$$wR_2 = \sqrt{\frac{\sum w(F_o^2 - F_c^2)^2}{\sum (wF_o^2)^2}} \tag{4}$$

where w is the weight, F_o is the observed structure factors, and F_c is the calculated structure factors. Another quality parameter to observe is the Goodness of Fit,

$$GooF = S = \sqrt{\frac{\sum w(F_o^2 - F_c^2)^2}{n - p}} \tag{5}$$

where n is the number of reflections and p is the total number of parameters. Preferably should R_{int}, R_1, and wR_2 all be low and $GooF$ close to 1.

The boxes displaying the shift and the other statistics will change color depending on the value of the data statistics, going from red (bad) to yellow (ok) and green (acceptable). See the color coding in Figure 24(a). The ratio of the number of reflections and number of model parameters should be high, to ensure a reliable model. This ratio is illustrated with a colored battery, visible in Figure 24(d).

SHELXL and olex2.refine both refine by least-squares, and the number of cycles can be adjusted according to preference. There are several other restraints and parameters available as well for SHELXL, see Table 4. More commands and their explanations are available from https://bpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/f/266 /files/2016/11/shelxl_comlist-1mp04in.pdf.

1. *Initial refinement:* Inspect 3D view of molecule. Does it correspond to a structure from the literature? Remove or reassign atoms of wrong type. After each edit, click refine. If the edit is correct, the model should improve and the R-value drop.

2. *Q-peaks:* Refinement produces suggestions for potential new atom positions for the structural model. These are colored brown and marked with a Q in the 3D model, see Figure 25(a). The number of Q-peaks can be adjusted in the refinement panel to the right, labeled "Peaks".

3. *Reset refinement:* If a wrong move is done (i.e. an atom that should be present, was removed), it is possible to go back to the previous state of the system under Work → History. Click the bar of the version you want to go back to and run refine.

4. *Measure properties:* To know details about the structure, click View → Geometry. Here, one can measure, for example, the distance between two chosen atoms or the angle between three chosen atoms. If any of the measures appears off compared to tabulated values, SHELXL restraining parameters like DFIX and TRIA can be used.

Table 4: List of important SHELXL commands and their functionality

Command	Description
TRIA	Fix the angle between the chosen three atoms.
DFIX	Fix the interatomic bond length. Choose two atoms and type the ideal length, for example DFIX 1.6 SiO4 O007.
EXTI	Extinction parameter. EXTI-parameters can take huge values for ED compared to XRD because of dynamical effects.
ACTA	Generates a .cif-file (crystallographic information file).
RIGU	Rigid bond restraint. Maintains reasonable atomic displacement parameters.

5. *Inspect statistics:* Click Work → Report, see Figure 24(b), and a report will be produced containing information about all interatomic distances, angles, atomic displacement parameters and more.

6. *Check electron potential map:* By clicking `ctrl + M` or alternatively Toolbox Work → Show Map, the electron density map will be displayed in 3D along with the structural model, as seen in Figure 25(b). Green illustrates maxima, and red illustrates minima in the electron potential map. To validate your 3D modelled molecular structure, inspect whether or not the maxima of the electron density map correlate with the Q-peak positions of the refined crystal structure. If this is the case - assign an element to the Q-peak and refine. In Figure 25, the lower right Q-peak position is located at the same spot as the electron density maxima (green blob), indicating that an atom should be present there. After assigning the atom to the Q-peak, Figure 25(c), run a new refinement and observe the electron density map improve and Q-peak positions change as seen in Figure 25(d). Under Map Settings, one wants to inspect the "Diff" map (difference between F_{obs} and F_{calc}). The level of noise/detail in the map can be adjusted with the 'sigma level' under Tools → Maps. Set the value to the right of the bar with caption "Level /eÅ^3" to lower values if there are too few features visible in the electron potential map.

7. *Further improvements:* Use parameter ACTA to produce a cif-file. EXTI has to be refined for electron diffraction experiments. Tick the box in Figure 24(c) or type EXTI in the Olex2 terminal.
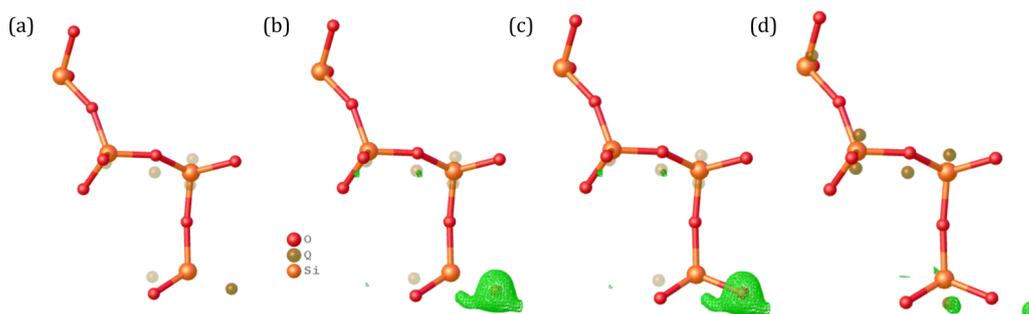
Figure 25: (a) molecule with visible Q-peaks. (b) molecule with electron density map and Q-peaks visualized simultaneously. (c) atom type 'O' assigned to Q-peak. (d) electron density map and Q-peak positions changed after new refinement.

If the R-values take very large values, click the percentage button next to the displayed R_1-value as shown in Figure 26. Then a plot showing the correlation between the observed and calculated structure factors will be displayed. Outliers where the observed and calculated structure factors do not correlate may be omitted by simply clicking on them in the plot. The normal filtered data are shown as orange data points, whereas the omitted data will turn gray.
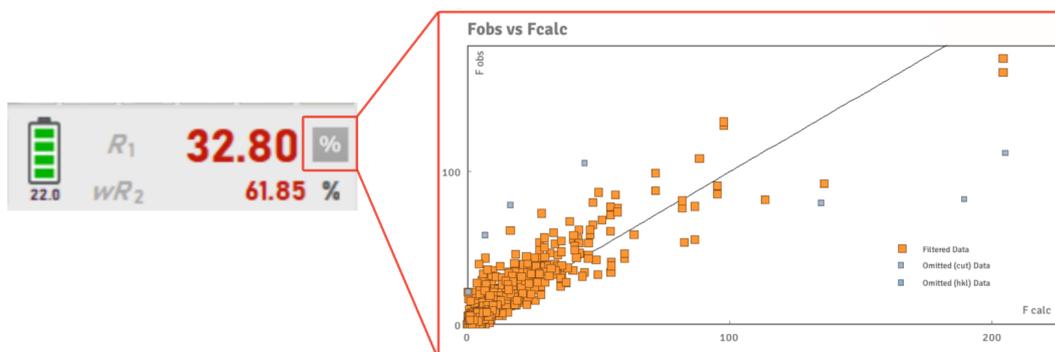

Figure 26: Plot of calculated structure factors against observed structure factors. Visible in OleX2 by clicking the percentage button next to the $R1$-value.

Potential problems

⚠ *Olex2 only shows me one atom.*

This is the default option for many structures containing only one element. To see more atoms, use the Olex2-command `grow` or alternatively `pack cell`. This will give a better overview of the full crystal structure. See more documentation at https://www.olexsys.org/olex2/docs/reference/commands-manual/changing-the-model-view/.

⚠ *How can I change colors of the atoms in the model?*

Olex2 chooses the colors of your atoms according to the default colors shown in Figure 27. If one has a molecule with elements i the same group or period they might have the same coloring, which can be unfavourable when visually inspecting the molecular structure. To change the color of the atom, right-click on it and choose Graphics → Draw style. A new window with material properties will appear where one can freely choose the color of different aspects of the graphics ("Ambient Front" being the main color). It is also possible to change the color of all atoms of the same atom type or atom name.
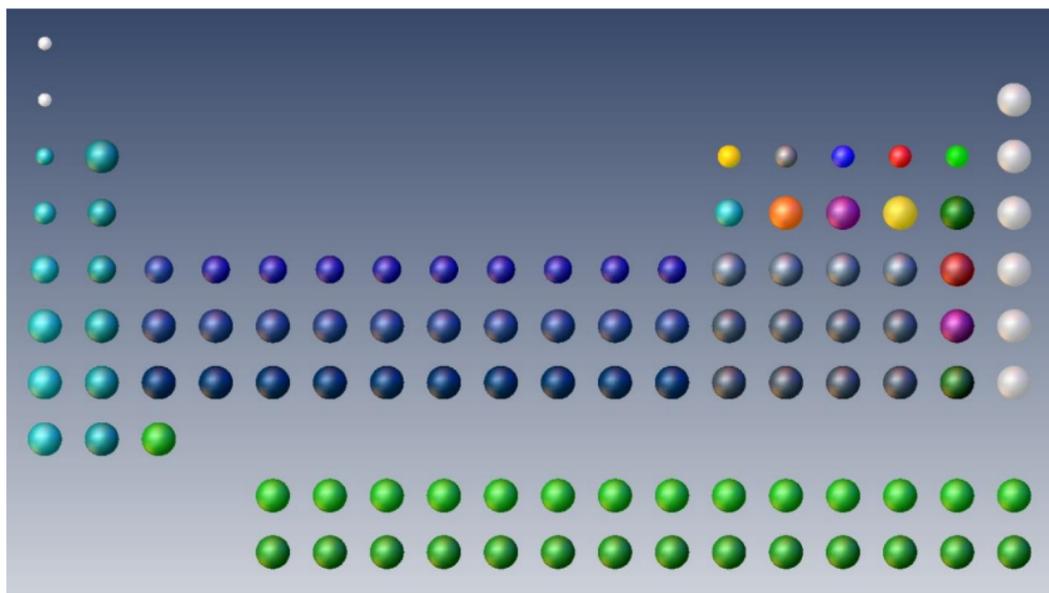

Figure 27: Default colours in Olex2 for the elements in periodic table.

⚠ *How do I improve the Hooft-value that appears when I started the refinement?*

A new statistics box with the name Hooft may appear when starting the refinement. This will only appear for non-centrosymmetric structures. A solution to this problem can be to invert the structure, even though Olex2 should be able to do this automatically. If not done automatically, it is possible to use this command `inv -f`. Omitting bad reflections with high standard deviations can also help the issue. More tips are available at https://www.olexsys.org/olex2/docs/tutorials/sucrose/tidying/.

⚠ *If I study single particles in a larger matrix, how do I proceed with refinement when the structure is very complex?*

For large complicated structures, refinement can be rather tedious. A way to simplify the problem, is to slice the structure in pieces and study single planes or parts of the structure. This is easy to do in 3D structure modelling tools like VESTA [13]. Compare this to the solved structure in Olex2. The anisotropic unit in Olex2 can be grown in specific directions by right-clicking specific atoms and then choose grow.

References

[1] M. Gemmi, E. Mugnaioli, T. E. Gorelik, U. Kolb, L. Palatinus, P. Boullay, S. Hovmöller, and J. P. Abrahams, "3d electron diffraction: The nanocrystallography revolution," *ACS Central Science*, vol. 5, no. 8, pp. 1315–1329, 2019. PMID: 31482114.

[2] P. Simoncic, E. Romeijn, E. Hovestreydt, G. Steinfeld, G. Santiso-Quiñones, and J. Merkelbach, "Electron crystallography and dedicated electron-diffraction instrumentation," *Acta Crystallographica Section E*, vol. 79, pp. 410–422, Apr 2023.

[3] Z. Huang, T. Willhammar, and X. Zou, "Three-dimensional electron diffraction for porous crystalline materials: structural determination and beyond," *Chemical Science*, vol. 12, 12 2020.

[4] T. Yang, T. Willhammar, H. Xu, X. Zou, and Z. Huang, "Single-crystal structure determination of nanosized metal-organic frameworks by three-dimensional electron diffraction," *Nature protocols*, vol. 17, p. 2389—2413, October 2022.

[5] S. Smeets, B. Wang, and E. Hogenbirk, "instamatic-dev/instamatic: 1.7.0," 2021.

[6] W. Wan, J. Sun, J. Su, S. Hovmöller, and X. Zou, "Rotation electron diffraction (red) software," Dec. 2013.

[7] W. Kabsch, "Automatic indexing of rotation diffraction patterns," *Journal of Applied Crystallography*, vol. 21, pp. 67–72, Feb 1988.

[8] W. Brehm, J. Triviño, J. M. Krahn, I. Usón, and K. Diederichs, "*XDSGUI*: a graphical user interface for *XDS*, *SHELX* and *ARCIMBOLDO*," *Journal of Applied Crystallography*, vol. 56, pp. 1585–1594, Oct 2023.

[9] G. M. Sheldrick, "*SHELXT* – Integrated space-group and crystal-structure determination," *Acta Crystallographica Section A*, vol. 71, pp. 3–8, Jan 2015.

[10] G. M. Sheldrick, "Crystal structure refinement with *SHELXL*," *Acta Crystallographica Section C*, vol. 71, pp. 3–8, Jan 2015.

[11] O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, and H. Puschmann, "*OLEX2*: a complete structure solution, refinement and analysis program," *Journal of Applied Crystallography*, vol. 42, pp. 339–341, Apr 2009.

[12] Y. Seto and M. Ohtsuka, "ReciPro: free and open-source multipurpose crystallographic software integrating a crystal model database and viewer, diffraction and microscopy simulators, and diffraction data analysis tools," *Journal of Applied Crystallography*, vol. 55, pp. 397–410, Apr 2022.

[13] K. Momma and F. Izumi, "*VESTA3* for three-dimensional visualization of crystal, volumetric and morphology data," *Journal of Applied Crystallography*, vol. 44, pp. 1272–1276, Dec 2011.

[14] IUCr, *International Tables for Crystallography, Volume A: Space Group Symmetry*. International Tables for Crystallography, Dordrecht, Boston, London: Kluwer Academic Publishers, 5. revised ed., 2002.

[15] "Iucr dictionary: Friedel's law.," 2017. https://dictionary.iucr.org/Friedel%27s_law [Accessed 2024-03-10].

[16] A. Saha, S. S. Nia, and J. A. Rodríguez, "Electron diffraction of 3d molecular crystals," *Chemical Reviews*, vol. 122, pp. 13883–13914, Sep 2022.
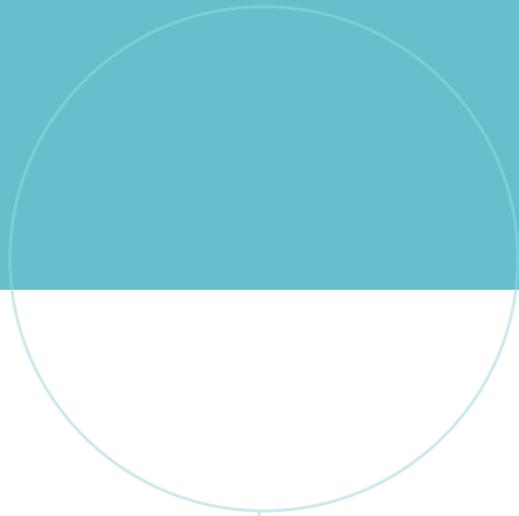
A XDS

A.1 XDS.INP example file

```
1  ! ********** Data images **********
2  NAME_TEMPLATE_OF_DATA_FRAMES= data/?????.img SMV  !EDIT the path to be
        the where the .img files are located
3  DATA_RANGE=            1 701 !(automatic from Instamatic)
4  SPOT_RANGE=            1 701 !(automatic from Instamatic)
5  BACKGROUND_RANGE=      1 701 !(automatic from Instamatic)
6  EXCLUDE_DATA_RANGE=5 6
7  !... list of all excluded frames (automatic from Instamatic)
8  EXCLUDE_DATA_RANGE=700 701
9  !If there are frames that cause spikes in the SCALE plot. Use the
        above command to remove them.
10
11 !********** Space group *********
12 SPACE_GROUP_NUMBER=227                        ! EDIT, with info from
        REDp
13 UNIT_CELL_CONSTANTS= 5.43 5.43 5.43 90 90 90 ! EDIT, with values from
        REDp
14
15 FRIEDEL'S_LAW=TRUE
16
17 !********** Angles **********
18 STARTING_ANGLE= -56.780 !(automatic from Instamatic)
19 STARTING_FRAME= 1
20
21 MAX_CELL_AXIS_ERROR= 0.05
22 MAX_CELL_ANGLE_ERROR= 2.0
23
24 TEST_RESOLUTION_RANGE=10. 1.0
25
26 !************Detector hardware ********
27 NX=512      NY=512
28 QX=0.0550   QY=0.0550
29 OVERLOAD= 5000                 !EDIT, input maximum intensity of direct
        beam
30 TRUSTED_REGION= 0.16  1.05     !OPTIONAL EDIT, enlargen the smaller
        ring if noise around the direct beam.
31 !^This is the blue rings visible in xdsgui FRAME tab, lower number
        correspond to the smaller ring.
32 DETECTOR=PILATUS
33 SENSOR_THICKNESS= 0.32
34 AIR=0.0
35
36 !************ Trusted region *************
37 UNTRUSTED_RECTANGLE= 255 258 0 512  ! EDIT from Instamatic
38 UNTRUSTED_RECTANGLE= 0 512 255 258  ! EDIT from Instamatic
```

```
39  VALUE_RANGE_FOR_TRUSTED_DETECTOR_PIXELS=10. 500000.

40

41  !********** Resolution range **********
42  INCLUDE_RESOLUTION_RANGE=7 0.7 !OPTIONAL EDIT, if much noise around
        the direct beam.
43  !This effects the green rings in xdsgui FRAME tab. Higher number
        correspond to the smaller ring.

44

45  ! ********** Detector geometry & Rotation axis **********
46  DIRECTION_OF_DETECTOR_X-AXIS=1 0 0
47  DIRECTION_OF_DETECTOR_Y-AXIS=0 1 0

48

49  ORGX= 274.79    ORGY= 235.76              !EDIT, insert beam center from
        REDp.
50  DETECTOR_DISTANCE= 245.24     !EDIT, insert calibrated value.

51

52  OSCILLATION_RANGE= 0.1593 !(automatic from Instamatic)

53

54  ROTATION_AXIS= -0.7986 0.6018 0   !EDIT, insert found rotation axis
        from REDp.

55

56  !********** Incident beam **********
57  X-RAY_WAVELENGTH= 0.041757              !EDIT, insert correct
        wavelength.
58  INCIDENT_BEAM_DIRECTION=0 0 1

59

60  ! ********** Background and peak pixels **********
61  !BACKGROUND_PIXEL= 6.0
62  STRONG_PIXEL=4               !OPTIONAL EDIT
63  MINIMUM_NUMBER_OF_PIXELS_IN_A_SPOT=4  !OPTIONAL EDIT, turn down if
        peaks are very sharp (a few pixels)
64  !SIGNAL_PIXEL=3.0

65

66  ! ********** Refinement **********
67  REFINE(IDXREF)=BEAM ORIENTATION CELL AXIS
68  REFINE(INTEGRATE)=  !POSITION BEAM ORIENTATION
69  REFINE(CORRECT)= ORIENTATION BEAM CELL AXIS

70

71  ! ********** Indexing **********
72  MINIMUM_FRACTION_OF_INDEXED_SPOTS= 0.2    !OPTIONAL EDIT, decrease
        this if IDXREF fails bc too few indexed spots.
73  MAXIMUM_ERROR_OF_SPOT_POSITION= 12.0          !OPTIONAL EDIT, increase
        if IDXREF fails bc too few indexed spots.
74  MAXIMUM_ERROR_OF_SPINDLE_POSITION=  8.0 !OPTIONAL EDIT, increase if
        IDXREF fails bc too few indexed spots.

75

76  DELPHI=2         !OPTIONAL EDIT, turn down if too few reflections
        saved to hkl in INTEGRATE.
77              !Turn up if other stuff fails in INTEGRATE (bin fail etc.)
```