UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 1451

Coherent Diffractive Imaging with X-ray Lasers

MAX FELIX HANTKE

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2016

ISSN 1651-6214
ISBN 978-91-554-9748-4
urn:nbn:se:uu:diva-306609

Dissertation presented at Uppsala University to be publicly examined in E10:1307-E10:1309, Biomedical Centre, Husargatan 3, Uppsala, Monday, 19 December 2016 at 08:30 for the degree of Doctor of Philosophy (Faculty of Theology). The examination will be conducted in English. Faculty examiner: Prof. Edgar Weckert (Hamburg University).

Abstract

Hantke, M. F. 2016. Coherent Diffractive Imaging with X-ray Lasers. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1451. 84 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9748-4.

The newly emerging technology of X-ray free-electron lasers (XFELs) has the potential to revolutionise molecular imaging. XFELs generate very intense X-ray pulses and predictions suggest that they may be used for structure determination to atomic resolution even for single molecules. XFELs produce femtosecond pulses that outrun processes of radiation damage and permit the study of structures at room temperature and of structural dynamics.

While the first demonstrations of flash X-ray diffractive imaging (FXI) on biological particles were encouraging, they also revealed technical challenges. In this work we demonstrated how some of these challenges can be overcome. We exemplified, with heterogeneous cell organelles, how tens of thousands of FXI diffraction patterns can be collected, sorted, and analysed in an automatic data processing pipeline. We improved image resolution and reduced problems with missing data. We validated, described, and deposited the experimental data in the Coherent X-ray Imaging Data Bank.

We demonstrated that aerosol injection can be used to collect FXI data at high hit ratios and with low background. We reduced problems with non-volatile sample contaminants by decreasing aerosol droplet sizes from ~1000 nm to ~150 nm. We achieved this by adapting an electrospray aerosoliser to the Uppsala sample injector. Mie scattering imaging was used as a diagnostic tool to measure positions, sizes, and velocities of individual injected particles.

XFEL experiments generate large amounts of data at high rates. Preparation, execution, and data analysis of these experiments benefits from specialised software. In this work we present new open-source software tools that facilitates prediction, online-monitoring, display, and pre-processing of XFEL diffraction data.

We hope that this work is a valuable contribution in the quest of transitioning FXI from its first experimental demonstration into a technique that fulfills its potentials.

Keywords: coherent diffractive X-ray imaging, lensless imaging, coherent X-ray diffractive imaging, flash diffractive imaging, single particle imaging, aerosol injection, electrospray injection, substrate-free sample delivery, carboxysome, phase retrieval, X-ray diffraction software, X-ray free-electron laser, XFEL, FEL, CXI, CDI, CXDI, FXI

Max Felix Hantke, Department of Cell and Molecular Biology, Molecular biophysics, Box 596, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-9748-4
urn:nbn:se:uu:diva-306609 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-306609)

Dedicated to Angelika, Klaus, Leonie, and Valerie

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **M. F. Hantke**, D. Hasse *et al.*
High-throughput imaging of heterogeneous cell organelles with an X-ray laser
Nature Photonics 8 943-949 (2014)

II **M. F. Hantke**, D. Hasse *et al.*
A data set from flash X-ray imaging of carboxysomes
Scientific Data 3:160061 (2016)

III **M. F. Hantke**, T. Ekeberg, and F. R. N. C. Maia
Condor: a simulation tool for flash X-ray imaging
Journal of Applied Crystalography 49 1356-1362 (2016)

IV B. J. Daurer, **M. F. Hantke**, C. Nettelblad, and F. R. N. C. Maia
Hummingbird : monitoring and analyzing flash X-ray imaging experiments in real time
Journal of Applied Crystallography 49 1042-1047 (2016)

V A. Barty, and R. A. Kirian, F. R. N. C. Maia, **M. Hantke**, C. H. Yoon, T. A. White, and H. Chapman
Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data
Journal of Applied Crystallography 47 1118-1131 (2014)

Reprints were made with permission from the publishers.

List of additional papers

VI A. Munke *et al.* Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source
Scientific Data 3 160064 (2016)

VII G. van der Schot *et al.* Open data set of live cyanobacterial cells imaged using an X-ray laser
Scientific Data 3 160058 (2016)

VIII T. E. Ekeberg *et al.* Three-Dimensional Reconstruction of the Giant Mimivirus Particle with an X-Ray Free-Electron Laser
Physical Review Letters 114:9 098102 (2015)

IX G. van der Schot *et al.* Imaging single cells in a beam of live cyanobacteria with an X-ray laser
Nature Communications 6 5704 (2015)

X A. D. Rath *et al.* Explosion dynamics of sucrose nanospheres monitored by time of flight spectrometry and coherent diffractive imaging at the split-and-delay beam line of the FLASH soft X-ray laser
Optics Express 22:23 28914-28925 (2014)

XI J. Andreasson *et al.* Automated identification and classification of single particle serial femtosecond X-ray diffraction data
Optics Express 22:3 2497-2510 (2014)

XII H. J. Park *et al.* Toward unsupervised single-shot diffractive imaging of heterogeneous particles using X-ray free-electron lasers
Optics express 21:23 28729-28742 (2013)

XIII E. Pedersoli *et al.* Mesoscale morphology of airborne core-shell nanoparticle clusters: X-ray laser coherent diffraction imaging
Journal of Physics B 46:16 SI 164033 (2013)

XIV N. D. Loh *et al.* Sensing the wavefront of X-ray free-electron lasers using aerosol spheres
Optics Express 21:10 12385-12394 (2013)

XV A. V. Martin *et al.* Noise-robust coherent diffractive imaging with a single diffraction pattern
Optics Express 20:15 16650-16661 (2012)

XVI N. D. Loh *et al.* Fractal morphology, imaging and mass spectrometry of single aerosol particles in flight
Nature 486:7404 513-517 (2012)

XVII M. M. Seibert, Ekeberg, T., Maia, F. R. N. C. *et al.* Single mimivirus particles intercepted and imaged with an X-ray laser
Nature 470:7332 78-81 (2011)

Contents

Abbreviations

AMO	Atomic, molecular and optical science
CDI	Coherent diffractive imaging
CXI	Coherent X-ray imaging
CXIDB	Coherent X-ray Imaging Data Bank
DFT	Discrete Fourier transform
DM	Difference map algorithm
DMA	Differential mobility analyser
cryo-EM	Cryo-electron microscopy
ER	Error reduction algorithm
ESI	Electrospray ionisation
EXFEL	European X-ray Free-Electron Laser
FF	Gas dynamic flow-focussing
FFT	Fast Fourier transform
FLASH	Free-electron LAser in Hamburg
FXI	Flash X-ray diffractive imaging
GDVN	Gas dynamic virtual nozzle
GPU	Graphics processing unit
HDF5	Hierarchical Data Format version 5
HIO	Hybrid input-output algorithm
IMMS	Ion mobility-mass spectrometry
LCLS	LINAC Coherent Light Source
MS	Mass spectrometry
MSI	Mie scattering imaging
NA	Numerical aperture
NESI	Nano electrospray ionisation
NMR	Nuclear magnetic resonance spectroscopy
NTA	Nanoparticle tracking analysis
PRTF	Phase retrieval transfer function
RAAR	Relaxed averaged alternating reflections algorithm
SFX	Serial femtosecond crystallography
SPI	Single Particle Imaging Initiative
SVD	Singular value decomposition
XFEL	X-ray free-electron laser

Part I:
Motivation

Motivation

Out of the 92 naturally occuring atomic elements, biology, as we know it, uses only 25. Even more striking 96.5% of the biological cell's mass is attributed to just four elements: carbon, hydrogen, nitrogen, and oxygen [1]. We know the underlying architectural patterns that connect atoms to biological molecules such as proteins, nucleic acids, lipids, and carbohydrates. Yet, while this knowledge is essential, it alone is insufficient for understanding how life functions. A complex network of specific molecular interactions underpins biological function. These interactions take place on atomic length scales (distances of ca. $1\,\text{Å} = 10^{-10}\,\text{m}$) and on time scales that range from years ($1\,\text{year} \approx 10^{7}\,\text{s}$) down to femtoseconds ($1\,\text{fs} = 10^{-15}\,\text{s}$) [43]. Technologies that allow measuring biological structures with high spatial and temporal resolution are key for developing a deep understanding of life at fundamental levels.

X-ray crystallography delivers atomic resolution for samples that can be crystallised. Yet, large molecular complexes and heterogeneous structures are often difficult or impossible to crystallise. Nuclear magnetic resonance spectroscopy (NMR) is a viable technique to study solution structures and structural heterogeneity but is restricted to relatively small molecules (up to ca. 100 kDa) and requires relatively large volumes of pure samples [52]. With the recent advent of single-electron detection cameras, cryo-electron microscopy (cryo-EM) has reached below 3 Å resolution for rather large biological macromolecules with lower requirements on purity and sample volume without the need of crystallisation [5]. However, cryo-EM has fundamental limitations, which are associated with the requirement for sample fixation by cryo-freezing, long exposure times, limited detectability of small particles (currently the minimum particle mass required is around 100 kDa [5]), and the short penetration depth of electrons in matter.

With the newly emerging technology of X-ray free-electron lasers (XFELs) [63] we have gained a type of radiation source that has the potential to revolutionise molecular imaging. XFELs produce very bright femtosecond X-ray pulses (currently up to ca. $10^{12}\,\text{photons}/\mu\text{m}^{2}$ and about 70 fs pulse duration) with wavelengths as short as 1 Å. The short wavelength permits in principle imaging to atomic resolution. The first lasing of an X-ray free-electron laser that can produce this kind of radiation was achieved in 2009 [27]. Over the past five years, X-ray lasers have made remarkable advances in physics, chemistry, materials science, and biology.

Femtosecond pulse durations of XFELs have the right time scale to capture fast biological processes. Moreover, the short pulses outrun processes of radiation damage and give rise to X-ray diffraction before the pulse obliterates the sample [75]. This allows to determine structures at room temperature [17]. The requirement for a sample support can be eliminated by injecting sample particles as an aerosol into the focus of the XFEL [11, 84]. XFELs generate vey intense X-ray pulses and predictions suggest that they may be used for structure determination to atomic resolution even for single molecules [75, 17]. Yet, this goal has not yet been reached. Larger particles give rise to brighter diffraction patterns and 2D projection images can be reconstructed utilising iterative phase retrieval algorithms and *a priori* information, such as the extent of the particle [15, 84]. 3D structures can be obtained by merging diffraction patterns from sequential exposures of identical copies of the particle aligned in different orientations [25]. Even for the extremely faint patterns of single proteins, 3D structures were successfully reproduced from simulated diffraction data employing an iterative 3D alignment algorithm and aggressive signal averaging [60]. Technological advances at modern XFELs promise more rapid repetition rates (reaching 27 000 Hz) and increased photon fluxes [83]. These improvements will be key for imaging smaller structural entities at higher resolution and give room for sampling conformational space to study structural heterogeneity.

In 2006, the concept of flash X-ray diffractive imaging (FXI) was experimentally demonstrated with an artificial test sample at the soft X-ray Free-electron LAser in Hamburg (FLASH) (formerly known as the VUV-FEL) [15]. In 2011, FXI with injected biological samples succeeded on Mimivirus particles at the LINAC Coherent Light Source (LCLS) with a higher repetition rate, higher photon flux, and harder X-rays [84]. While the experiment on Mimivirus was encouraging as a proof-of-concept, it also demonstrated technical challenges associated with FXI, such as difficulties with low hit ratios and reconstruction artefacts due to saturated and obscured regions of the detector. Furthermore, Mimivirus particles with a diameter of 450 nm and a mass of 28 GDa represent one of the biggest virus species known and produce a lot of scattering signal. But, nevertheless, the resolution achieved was limited to 32 nm. Attempts of injecting significantly smaller biological particles with the same aerosolisation technique as used for Mimivirus failed because of the predominant formation of aggregates instead of isolated particles [51, 23]. These difficulties have to be overcome somehow. Also, it became obvious that open experimental data and specialised software tools for data prediction, online monitoring, data pre-processing, and for automated analysis were needed but unavailable [66, 67]. These problems indicated that significant technological development and improvement were needed to advance FXI to its true potential [3]. This work addresses many of these challenges and contributes with new methodology and software to overcome them.

Part II:
Concept

1. Coherent diffractive imaging with X-ray lasers

1.1 Diffraction before destruction

For achieving the goal of imaging biological structres at atomic length scales, one needs to consider the damage created by the probing radiation, which is detrimental and ultimately limits resolution. A photon interacts with an atom through either elastic scattering, photon absorption, (inelastic) Compton scattering, photonuclear absorption, or pair production. For photon energies well below 1 MeV that are of relevance here photonuclear absorption and pair production are extremely rare and can be neglected [78]. Elastic scattering denotes scattering without energy transfer. Elastic scattering is the most "useful" form of interaction for many structural techniques with X-rays as it provides a way of harvesting structural information without causing structural damage. Unavoidably, a large fraction of X-rays are not elastically scattered and can cause radiation damage through photoabsorption and Compton scattering. The transfer of energy from the photons to matter results in electron ejection followed by dislocation of entire atoms and radicals [17]. Such electronic and structural changes lead to the degradation of the scattering signal and may result in lower resolution or artefacts in the determined structure.

The "tolerable" radiation dose-limit quantifies for a given resolution the minimum energy deposited in the sample at which structural damage is detected [77, 15]. In electron microscopy and crystallography, plunge-freezing into liquid nitrogen or ethane is used to fixate the sample for subsequent measurement at cryogenic temperatures [40]. This procedure reduces effects of radiation damage and increases tolerable radiation doses in X-ray crystalography by about two orders of magnitude [77, 76]. In this work we employ a different approach to reduce effects of radiation-induced damage on image quality. The approach makes use of the fact that the time scale of the process of X-ray diffraction is much shorter than the process of radiation-induced sample degradation [87]. The concept of *diffraction before destruction*, suggested in 2000 [75] and experimentally demonstrated in 2006 [15], exploits the short pulse durations of X-ray lasers to outrun key processes of radiation damage. XFELs [63] can produce femtosecond pulses with peak spectral brightnesses up to eight orders of magnitude higher than third-generation synchrotron sources. In the micron-sized focus of an XFEL, power densities of 10^{16} W/cm^2 and above are reached in a single pulse. These conditions lead to the vaporisation of the sample, but given the pulse is sufficiently short, significant structural changes do not occur before the pulse has passed the sample [75, 17]. Under these circumstances the X-rays probe the nearly intact structure and no cryofixation is needed. Measurements can be performed at room temperature, at which the structure better resembles its native state [59].

The bombardment of matter by the trillions of photons in a femtosecond X-ray pulse is a violent process. In FXI experiments, doses of more than 1 GJ/kg can be reached, which correspond to more than one ionisation event per atom on average over the duration of the exposure [17]. Because of the high energy of X-ray photons, absorption of a photon by an atomically bound electron is usually followed by its ejection from the atom. For X-rays the core shell electrons have the largest cross sections and are ejected first [75]. Atoms with electron core holes ("hollow atoms") are more transparent and less likely to absorb more photons [101]. The electron core holes are repopulated within femtoseconds by Auger decay (modulated by shake-up or shake-off effects [79]) and fluorescence. Auger decay leads to the generation of additional free electrons. Free electrons created by Auger decay and absorption, unless they escape from the sample on a direct path, initiate cascades of secondary impacts and ionisations and reach thermalisation within 10-100 femtoseconds [103, 95, 39]. Finally, as a consequence of increasing temperature and successive ionisation the sample turns into a plasma and deteriorates entirely by thermal hydrodynamic expansion or Coulomb explosion, depending on pulse parameters, sample size, and atomic composition [75, 9]. It has been suggested that pulses with durations below 5 femtoseconds would be ideal for imaging as they would outrun Auger decay and damage that accumulates by secondary electron collisions [38].

A lot of experimental evidence suggests the validity of the concept "diffraction before destruction". Serial femtosecond crystallography (SFX) [18] has demonstrated that with this concept at least three orders of magnitude higher radiation doses are acceptable compared to conventional protein crystallography at cryogenic temperatures [17]. In SFX, X-ray damage breaks the periodicity in the crystal lattice and as a consequence damage may terminate Bragg diffraction before the pulse has finished passing through the crystal [7]. This effect allows SFX to reach atomic resolution with pulses as long as 50-100 fs [7]. Recently, a similar effect was predicted for single particles and pulse durations of 30-50 fs [69]. Current limitations for resolution in FXI are not related to radiation damage but mostly to low signal-to-noise ratios in the diffraction data [3] and it remains to be seen in practice what effects will ultimately limit resolution in FXI.

In addition to the reduction in observable radiation damage, the ultra-short pulses produced by XFELs have other advantages for imaging. For example, short pulses allow imaging of radiation-sensitive structures [53, 90] and anomalous signal due to dose-dependent bleaching of heavy atoms can be used for *de novo* phase retrieval [33]. Also, femtosecond XFEL pulses at high repetition rates could make it possible to study dynamics by combining snapshots of thermal fluctuation of the structure or by tracking structural changes subsequent to an external trigger (e.g. a pump pulse, which initiates a process that is then probed by the X-ray pulse) [53, 55]. These opportunities open the way for creating “molecular movies”.

1.2 Lensless X-ray imaging

1.2.1 Diffractive imaging

Traditional microscropes use lenses for magnification and image formation. For X-rays, however, refractive indices are very close to unity and consequently X-ray lenses perform poorly. Therefore most structural X-ray methods dispense of image forming lenses and instead retrieve structural information from the free-space propagated diffraction pattern in combination with additional informaton, such as the priorly known sample size, for phase reconstruction.

This work deals with "plane-wave coherent diffractive imaging (CDI)". The geometry for this lensless X-ray imaging method is minimalistic and simple (Fig. 1.1a). A coherent plane wave with wave vector $\mathbf{k}_0$ illuminates a scattering volume of finite extent. The scattered wave field propagates along wave vectors $\mathbf{k}_1$ and the diffraction pattern, from which the object is reconstructed, is measured with an area detector.

For elastic scattering the wave vectors $\mathbf{k}_0$ for the incoming wave and $\mathbf{k}_1$ for the outgoing wave have the same length and all scattering vectors $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0$ lie on the surface of a sphere in diffraction space. This sphere is called *Ewald sphere*. The concept of the Ewald sphere is discussed in more detail in ch. 1.2.6.

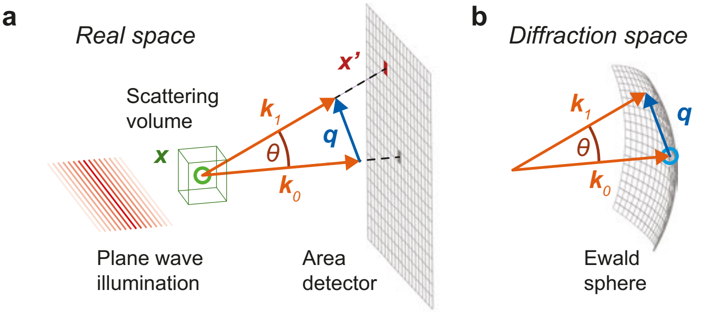


Figure 1.1. Geometry for plane-wave CDI in real space (**a**) and diffraction space (**b**). We assume a plane wave illuminates a set of scatterers at positions $\mathbf{x}$ in a finite scattering volume. The scattered wave is measured with an area detector at pixel positions $\mathbf{x}'$. $\mathbf{k}_0$ is the wave vector of the incoming wave, $\mathbf{k}_1$ the wave vector of the outgoing wave, and $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0$ denotes the scattering vector. All scattering vectors $\mathbf{q}$ lie on the Ewald sphere, which is centered at $-\mathbf{k}_0$ in diffraction space and has the radius k. (Fig. 1 in **Paper III**)

1.2.2 The time-independent scalar wave equation

The coupling of the electromagnetic field of the X-ray wave with matter and its propagation in space determines how structural information is encoded in the diffracted wave field. In electrodynamics the coupling with matter is parameterised by two material-specific quantities: the electric permittivity $\varepsilon(\mathbf{x})$ and the magnetic permeability $\mu(\mathbf{x})$. For biological samples we may safely assume that the medium represents a linear and isotropic dielectric[1] and is not magnetic ($\mu(\mathbf{x}) \approx \mu_0$). Further, we do not consider electric permanent dipoles, electric currents, and electric charge densities to be present in the medium[2]. Under these assumptions the spatial propagation of an electro-magnetic oscillating field of frequency ω with the electric field $\mathbf{E}(\mathbf{x})$ and the magnetic field $\mathbf{H}(\mathbf{x})$ is goverend by *Maxwell's equations* in the form

$$\nabla \cdot [\varepsilon(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x})] = 0 \tag{1.1}$$

$$\nabla \cdot \mathbf{H}(\mathbf{x}) = 0 \tag{1.2}$$

$$\nabla \times \mathbf{E}(\mathbf{x}) = \mathrm{i}\,\omega\,\mu_0\,\mathbf{H}(\mathbf{x}) \tag{1.3}$$

$$\nabla \times \mathbf{H}(\mathbf{x}) = -\mathrm{i}\,\omega\,\varepsilon(\mathbf{x})\,\mathbf{E}(\mathbf{x}) \tag{1.4}$$

(for reference see for example ref. 78 or 12). For spatial variations of $\varepsilon(\mathbf{x})$ that are on a length scale larger than the wavlength the time-independent wave equations

$$\nabla^2 \mathbf{E}(\mathbf{x}) + \varepsilon(\mathbf{x})\mu_0\omega^2\,\mathbf{E}(\mathbf{x}) = 0 \text{ and} \tag{1.5}$$

$$\nabla^2 \mathbf{H}(\mathbf{x}) + \varepsilon(\mathbf{x})\mu_0\omega^2\,\mathbf{H}(\mathbf{x}) = 0 \tag{1.6}$$

can be obtained from (1.1) to (1.4).

For all following considerations in this work we can neglect the fact that the electromagnetism is a vector-field theory and describe the electro-magnetic wave by the complex-valued scalar wave function $\Psi(\mathbf{x}) = A(\mathbf{x}) \cdot \exp(\mathrm{i}\,\phi(\mathbf{x})$ that obeys the scalar *Helmholtz equation*

$$\nabla^2 \Psi(\mathbf{x}) + \varepsilon(\mathbf{x})\mu_0\omega^2\,\Psi(\mathbf{x}) = 0\,. \tag{1.7}$$

The complex nature of the wave $\Psi(\mathbf{x})$ has the interpretation that $\phi(\mathbf{x})$ represents the phase and $A(\mathbf{x})$ the magnitude of the electromagnetic wave. By identifying in (1.7) the wave number with the relation $k = \omega/c$, where c is the speed of light in vacuum, and the refractive index with $n(\mathbf{x}) = c\,\sqrt{\mu_0\varepsilon(\mathbf{x})}$ one obtains the Helmholtz equation in its usual form

$$\nabla^2 \Psi(\mathbf{x}) + (kn(\mathbf{x}))^2\,\Psi(\mathbf{x}) = 0\,. \tag{1.8}$$

---

[1] A dielectric is linear if the electric displacement **D** is proportional to the electric field **E** with the proportionality constant ε. The dielectric is isotropic if ε is invariant with respect to the direction of the electromagnetic field vectors.

[2] As a consequence of radiation damage processes this assumption does no longer hold.

1.2.3 The single-scattering approximation

X-rays are known for their weak interaction with matter and their penetrating property. This not only permits to study relatively thick samples with X-rays but it also allows in many applications to model the scattering process in the single-scattering approximation[3]. This means that we can neglect multiple-scattering events and the decrease of the illumination intensity due to absorption or scattering by the sample.

To motivate the diffraction formula in this approximation we refer back to the scattering scenario for plane-wave CDI as depicted in Fig. 1.1a. Let the sample volume be represented by infinitesimal point scatterers located at positions $\mathbf{x}$ within the sample volume. In this picture the scattered wave $\Psi(\mathbf{x}')$ can be subdivided into the incoming plane wave $\Psi^{(0)}$ and the scattering term $\Psi^{(1)}$, which accounts for the superposition of the many spherical waves that emmanate from point scatterers in the sample volume and we may write[4]

$$\begin{aligned} \Psi(\mathbf{x}') &= \Psi^{(0)}(\mathbf{x}') \qquad &&+\Psi^{(1)}(\mathbf{x}') && (1.9) \\ &= \Psi_0 \exp(\mathrm{i}\mathbf{k}_0\mathbf{x}') &&+\Psi_0 \iiint \varphi(\mathbf{x}) \exp(i\mathbf{k}_0\mathbf{x}) \frac{\exp(-ik|\mathbf{x}' - \mathbf{x}|)}{|\mathbf{x}' - \mathbf{x}|} \mathrm{d}\mathbf{x}\,, && (1.10) \end{aligned}$$

with the incoming wave amplitude Ψ_0 and the scattering potential $\varphi(\mathbf{x})$ defined as

$$\varphi(\mathbf{x}) = \frac{k^2}{4\pi} \left[1 - n^2(\mathbf{x})\right] \,. \tag{1.11}$$

1.2.4 Projection approximation and optically thin objects

The limits of the applicability of (1.10) become more concrete when solving the Helmholtz equation with the ansatz

$$\Psi(x, y, z) = \psi(x, y, z) \cdot \exp(ikz)\,. \tag{1.12}$$

It describes a plane wave $\exp(ikz)$ travelling in z with an envelope $\psi(x, y, z)$. By inserting (1.12) into (1.8) we obtain with some algebra[5]

$$\left(\partial_x^2 + \partial_y^2 + \partial_z^2 + 2\mathrm{i}k\partial_z - k^2 \left[1 - n^2(x, y, z)\right]\right) \psi(x, y, z) = 0 \tag{1.13}$$

Here we are interested in slowly varying envelopes $\psi(x, y, z)$ that are beam-like. Second derivatives ∂_x^2, ∂_y^2, ∂_z^2 of $\psi(x, y, z)$ are therefore small and can be neglected in (1.13) because they do not scale with k in the case of X-rays. It follows

$$\partial_z\psi(x, y, z) = \frac{k}{2\mathrm{i}} \left[1 - n^2(x, y, z)\right] \psi(x, y, z)\,. \tag{1.14}$$

---

[3] Depending on the context the single-scattering approximation is also referred to by the terms *kinematic approximation* or *Born approximation*.

[4] For a derivation of (1.10) from (1.8) for small $\varphi(\mathbf{x})$ see for example ref. 78 chapter 2.3.

[5] See for example ref. 78.

For a sample volume confined by the planes $z = 0$ and $z = z_0$ the differential equation (1.14) can be solved by functions

$$\psi(x, y, z_0) \propto \exp\left(\frac{k}{2\mathrm{i}} \int_0^{z_0} \left[1 - n^2(x, y, z)\right] \, \mathrm{d}z\right) . \tag{1.15}$$

Finally, for deriving (1.10) from (1.15) we must be able to linearise the exponential function in (1.15), which requires a small exponential argument. For X-rays this can be realised by small deviations of n from unity (i.e. weak X-ray-matter interaction) and small z_0 (i.e. thin objects). Conceptually, we call objects *optically thin* if they fulfill the requirements for linearisation of the exponential term in (1.15).

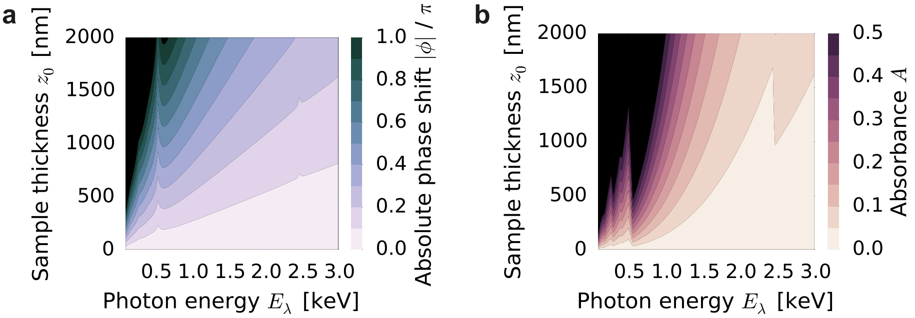


Figure 1.2. Absolute phase shift (**a**) and absorbance (**b**) of a wet cell. Absolute phase shift $|\Delta\phi|$ and absorbance A as a functions of photon energy E_λ (horizontal axis) and sample thickness z_0 (vertical axis) for the sample model of a wet biological cell. For the calculation of δ and β atomic composition and mass density for a wet cell were defined as in ref. 9 and atomic scattering factors were taken from the Henke tables [42].

In X-ray physics the refractive index n is defined as a complex-valued quantity, which deviates only slightly from unity. It is often expressed as

$$n = 1 - \delta + \mathrm{i}\,\beta\,, \tag{1.16}$$

where δ and β are small real numbers such that we may safely approximate $1 - n^2 \approx 2(1 - n)$. On the basis of this approximation and (1.15) we may calculate for any material characterised by $\delta(x, y, z)$ and $\beta(x, y, z)$ the phase shift $\Delta\phi$ and the absorbance A as

$$\Delta\phi(x, y, z_0) = -\,k \int_0^{z_0} \delta(x, y, z) \, \mathrm{d}z \text{ and} \tag{1.17}$$

$$A(x, y, z_0) = 1 - \exp\left(-2k \int_0^{z_0} \beta(x, y, z) \, \mathrm{d}z\right) . \tag{1.18}$$

For optically thin objects $|\Delta\phi(x,y,z_0)|/\pi$ and $A(x,y,z_0)$ must be significantly smaller than unity. For example a cell organelle[6] with a thickness of 115 nm imaged at a photon energy of 1.1 keV has $|\Delta\phi| = 0.04\pi$ and $A = 0.03$ and thus fullfills the requirements for an optically thin object (see Fig. 1.2). Consequently, diffraction from this sample can be described using the single-scattering approximation (1.10) and the exit-wave at $z = z_0$ may be interpreted in terms of a projection.

1.2.5 Fraunhofer far-field

Scattering potential φ and scattered wave $\Psi^{(1)}$ are linked by a Fourier transform if the detector screen is placed at a far distance. More explicitly, the detector distance must be long enough such that propagation distances $r = |\mathbf{x}' - \mathbf{x}|$ are much larger than the extent of the sample. The term *Fraunhofer far-field* denotes this regime for which we may simplify (1.10) to[7]

$$\Psi^{(1)}(\mathbf{q}) = \frac{\Psi_0}{r} \iiint \varphi(\mathbf{x}) \exp(-i\mathbf{q}\mathbf{x}) \mathrm{d}\mathbf{x} \,, \qquad (1.19)$$

where $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0$ denotes the scattering vector (or vector of momentum transfer). By using the definition for the continuous Fourier transform for any well-behaved function $h(\mathbf{x})$ in l Euclidian dimensions

$$\mathscr{F}\left[\, h(\mathbf{x}) \,\right](\mathbf{q}) = \tilde{h}(\mathbf{q}) = (2\pi)^{-l/2} \int_{\mathbb{R}^l} h(\mathbf{x}) \exp(-i\mathbf{q}\mathbf{x}) \mathrm{d}\mathbf{x} \qquad (1.20)$$

we may substitute the integral in (1.19) by (1.20) and obtain for $l = 3$

$$\Psi^{(1)}(\mathbf{q}) = \frac{\Psi_0}{r} \cdot (2\pi)^{3/2} \cdot \mathscr{F}\left[\, \varphi(\mathbf{x}) \,\right](\mathbf{q}) \,. \qquad (1.21)$$

Equation (1.21) predicts diffraction based on a known structure. The inverse process for retrieving the structure $\varphi(\mathbf{x})$ from a given function $\Psi^{(1)}$ can be mathematically derived by applying the inverse Fourier transform $\mathscr{F}^{-1}$ to (1.21). We define $\mathscr{F}^{-1}$ consistently with (1.20) as

$$\mathscr{F}^{-1}\left[\, \tilde{h}(\mathbf{q}) \,\right](\mathbf{x}) = h(\mathbf{x}) = (2\pi)^{-l/2} \int_{\mathbb{R}^l} h(\mathbf{q}) \exp(i\mathbf{q}\mathbf{x}) \mathrm{d}\mathbf{q} \qquad (1.22)$$

and obtain from (1.19) the aspired relation

$$\varphi(\mathbf{x}) = \frac{r}{\Psi_0} \cdot (2\pi)^{-3/2} \cdot \mathscr{F}^{-1}\left[\, \Psi^{(1)}(\mathbf{q}) \,\right](\mathbf{x}) \,. \qquad (1.23)$$

---

[6] For this rough estimate we may assume that the atomic composition and mass density of the cell organelle are sufficiently similar to the values for a wet cell as defined in ref. 9. Scattering factors are obtained from the Henke tables [42].

[7] Intensity measurements only capture the absolute value of $\Psi^{(0)}$, which is the reason why we omit from here on the phase factor $\exp(ikr)$.

For short propagation distances r the Fraunhofer approximation does no longer apply and instead the theory of *Fresnel diffraction* provides an adequate description for the propagated wave field. In this work we only deal with diffraction data that was collected in the far-field. For details on Fresnel diffraction we refer the reader to the literature (see for example ref. 78).

1.2.6 The Ewald sphere

It follows from (1.21) that $\Psi^{(1)}$ in diffraction space, i.e. space of scattering vectors **q**, is a representation of the Fourier transform of the scattering potential $\varphi(\mathbf{x})$. For solving a structure in 3D we want to obtain $\varphi(\mathbf{x})$ through execution of the Fourier integral in (1.23). This is only possible if we know amplitude and phase of $\Psi^{(1)}$ for all scattering vectors **q** up to a given resolution. As illustrated in Fig. 1.1b, a single diffraction pattern samples $\Psi^{(1)}(\mathbf{q})$ at points **q** that lie on the surface of the Ewald sphere.

While a single diffraction pattern only provides partial 3D information collecting multiple diffraction patterns of the sample in different orientations enables us to collect in principle full 3D information up to the diffraction limit. The resolution at the diffraction limit is given by the diameter $2k$ of the Ewald sphere and is $(2\pi)/(2k) = \lambda/2$.

1.2.7 2D imaging

A single diffraction pattern does not generally provide sufficient information for recovering the full 3D structure of an unknown object. But a 2D image of the exit wave front can be obtained if the Ewald sphere lift-off (i.e. the departure of the Ewald sphere from the plane orthogonal to $\mathbf{k}_0$) is sufficiently small. Without loss of generality we assume as in ch. 1.2.4 that the wave propagates along z and the sample slab is confined between the two planes $z = 0$ and $z = z_0$. The Ewald sphere lift-off is negligible if $\forall z \in [0, z_0]$: $\exp(iq_z z) \approx 1$, which allows in (1.19) to execute the integration in z. The result is

$$\Psi_{\perp}^{(1)}(q_x, q_y) = \frac{\Psi_0}{r} \iint \varphi_{\perp}(x, y)\ \exp\left(-\mathrm{i}(q_x x + q_y y)\right)\ \mathrm{d}x\, \mathrm{d}y \qquad (1.24)$$

with

$$\varphi_{\perp}(x, y) = \int \varphi(x, y, z)\ \exp(-\mathrm{i}q_z z)\mathrm{d}z\,. \qquad (1.25)$$

With the same argument as for the 3-dimensional case (1.24) can be inverted by applying (1.22). This results in

$$\varphi_{\perp}(x, y) = \frac{r}{\Psi_0} \cdot (2\pi)^{-1/2} \mathscr{F}^{-1}\left[\, \Psi^{(1)}(q_x, q_y)\,\right](x, y)\,, \qquad (1.26)$$

which gives the exit wave field ($z = z_0$) as a function proportional to the inverse Fourier transform of the diffracted wave field.

1.2.8 Scattering strength

We define the scattering intensity I_p as the expectation value for the number of scattered photons measured by a detector pixel that covers the solid angle $\Delta\Omega$. I_p can be expressed as the product of $\Delta\Omega$, the incident photon flux $F_0 = |\Psi_0|^2$, and the particle's differential cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$:

$$I_p = F_0 \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \Delta\Omega \tag{1.27}$$

(for reference see for example [2] ch. 1.2). For scattering objects smaller than the wavelength the differential cross-section is determined by the product of the polarisation factor P and the modulus square of Φ, which denotes the volume integrated particle's scattering potential φ

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = P\,|\Phi|^2 \text{ with } \Phi = \int \varphi(\mathbf{x})\,\mathrm{d}\mathbf{x}\,. \tag{1.28}$$

For larger scattering objects intereference must be taken into account and for a small sensitive detector area $A_p = r^2\Delta\Omega$ the diffraction intensity I_p may be expressed as

$$I_p = P\,|\Psi^{(1)}(\mathbf{q})|^2 A_p\,. \tag{1.29}$$

P accounts for the angular dependence of the scattering with respect to the polarisation of the incident wave. For linearly polarised radiation, such as undulator radiation from XFELs and synchrotrons, the polarisation factor is

$$P = \cos^2(\chi)\,, \tag{1.30}$$

where χ denotes the angle between the axis of observation and the plane orthogonal to the polarisation direction. We note that for small-angle scattering geometries the angular dependence of the polarisation factor may be neglected as $P \approx 1$.

X-rays interact primarily with the electrons in matter. Φ for a (quasi) free electron equals the classical electron radius

$$r_0 = \frac{e^2}{4\pi\varepsilon_0 m_e c^2} \approx 2.817\,940 \times 10^{-15}\,\mathrm{m}\,, \tag{1.31}$$

where e denotes the electron charge, m_e the mass of an electron, and c the speed of light in vacuum. By integrating the differential cross section of a single electron in (1.28) over the entire solid angle 4π we obtain the total cross section σ_t. For a free electron $\sigma_t = (8\pi/3)\cdot r_0^2 \approx 0.665 \times 10^{-28}\,\mathrm{m}^2 = 0.665\,\mathrm{barn}$. For illustrating the weakness of this interaction we make a gedanken experiment. Let us position a free electron into the focus of an XFEL beam. Even with a photon flux of $F_0 = 10^{14}\,\mathrm{photons}/\mu\mathrm{m}^2$ we would expect scattering with a probability of only $F_0\sigma_t = 0.665\,\%$. This number illustrates the challenge that single molecule imaging faces at photon fluxes currently obtainable with XFELs.

For a spatial distribution of free electrons at positions $\mathbf{x}_i$ the scattering potential $\varphi(\mathbf{x})$ can be written as a product of r_0 and the sum over Kronecker deltas[8] $\delta(\mathbf{x}_i)$

$$\varphi(\mathbf{x}) = r_0 \cdot \sum_i \delta(\mathbf{x}_i)\,. \tag{1.32}$$

For atomically bound electrons the scattering length differs from r_0. Atomic scattering factors $f_a^{(\lambda)}$ are tabulated values, which specify the relative scattering length (relative to a free electron) for atom species a and wavelength λ [42]. By using the formalism of atomic scattering factors we may express $\varphi(\mathbf{x})$ in analogy to (1.32) as a sum over all atoms i of element species a_i located at the positions $\mathbf{x}_i$. This can be written as

$$\varphi(\mathbf{x}) = r_0 \cdot \sum_i \delta(\mathbf{x}_i)\, f_{a_i}^{(\lambda)}\,. \tag{1.33}$$

It should be noted that the shape of atomic orbitals accounts for an angular dependence of $f_a^{(\lambda)}$, which we neglect here but must be considered as soon as soon as resolutions close to the scale of a single atom are approached. By combining (1.33) and (1.11) we find that the refractive index for a material of known atomic composition of atom species a with number densities ρ_a can be calculated with the formula

$$n = 1 - \frac{2\pi}{k^2} \sum_a \rho_a f_a^{(\lambda)} r_0\,. \tag{1.34}$$

This formula bridges the concept of refraction by a continous optical medium and the concept of scattering by discrete particles. These are two ways of describing X-ray-matter-interaction and depending on the resolution in a specific application one or the other may be chosen.

1.2.9 Intensity measurements

Even in the ideal case of no noise and no signal loss, X-ray intensitiy measurements are a less than exact representation of I_p. The reason lies in the quantum-mechanical nature of X-rays. More specifically, measuring I_p is a Poisson process for detecting a discrete numbers of photons of defined energy

$$E_\lambda = \frac{hc}{\lambda}\,, \tag{1.35}$$

where h denotes the Planck constant. The probability $p_N^{(I_p)}$ for detecting $N \in \mathbb{N}_0$ photons follows the Poissonian distribution

$$p_k^{(\Lambda)} = \frac{\Lambda^k \exp(-\Lambda)}{k!} \tag{1.36}$$

---

[8] We define the unit of $\delta(\mathbf{x})$ as $1\,\mathrm{m}^{-3}$.

with expectation value $\Lambda = I_p$ and the number of photon counts $k = N$.

It is a courious fact that Poisson noise in I_p leads to a constant standard deviation for $|\Psi^{(1)}|$ independent of the value of I_p. The reason for it is that the standard deviation of Poisson distributed values generally equals the square root of the expectation value[9]. By error propagation for $|\Psi^{(1)}| \propto \sqrt{I_p}$ it follows that the standard deviation of $|\Psi^{(1)}|$ is the constant $1/2$.

---

[9] We assume $\Lambda > 0$.

1.3 Phase retrieval

1.3.1 The phase problem

The electromagnetic wave fields of X-rays have oscillation periods of the order of 10^{-20} to 10^{-16} seconds. To date no detector has the temporal resolution to directly measure the time evolution of these fields. Instead X-ray detectors measure time-integrated intensities I_p, which deliver for

$$\Psi^{(1)} = |\Psi^{(1)}| \cdot \exp(-\mathrm{i}\phi) \qquad (1.37)$$

the amplitude $|\Psi^{(1)}|$ but not the phase ϕ. Phasing methods compensate the lack of phases in the measured data by exploiting additional *a priori* information. For X-ray diffraction the first phasing methods were developed in crystallography. The most widely used phasing method is the *molecular replacement method* [81]. It generates phases on the basis of a known protein structure with high sequence identity. *De novo* phases for crystal structures are determined by methods such as single-wavelength and multiple-wavelength anomalous diffraction (SAD/MAD) [41, 37], which exploit the relatively large imaginary part of the scattering factors of heavy elements. In SFX bleaching of heavy atoms due to radiation damage may be used in a similar fashion [88]. For a detailled review on phasing methods in crystallography see for example ref. 91.

Despite the fact that CDI is a relatively young method, many different phasing strategies have been developed (for a comprehensive review see fore example ref. 19). Some methods make use of the coherent interference with a reference wave to retrieve phase information. In *holography* [71] for example a small object (ideally a point source) is placed in the proximity of the sample and creates the reference wave. *Fresnel CDI* [99] instead uses as a reference a curved wave that illuminates both the sample and the detector. Very differently, the scanning method *ptychography* [29] harvests phase information from redundancy in diffraction data collected at consciously overlapping scan positions. In this work we employ *plane-wave CDI*, which is probably the most minimalistic phasing method of CDI. *Plane-wave CDI* on isolated single objects [82, 72] does not use a reference wave of multiple exposures to obtain additional phase information. Phases are derived directly from the object's far-field intensity pattern using additional information through *oversampling* of the intensities. Oversampling relies solely on the finite and relatively small size of the object. For making the reader familiar with this concept we provide in the following sections the necessary foundation of sampling theory, its application to plane-wave CDI, and an introduction into iterative phasing algorithms.

1.3.2 Discrete sampling

In the preceding sections all equations have been formulated under the assumption of continuous sampling. Yet, in reality we measure diffraction patterns with pixelated detectors that generate discrete sequences of values

$$I_p = \frac{\int_{A_p} I(\mathbf{x}')\mathrm{d}A}{A_p}\,, \tag{1.38}$$

each representing the measured intensity integrated over the sensitive area A_p of the respective pixel. Given that A_p is sufficiently small such that $I_p \approx I(\mathbf{x}')$ we may substitute in all equations above continuous Fourier transforms by discrete Fourier transforms (DFTs). For the definition of the DFT let the functions $h(\mathbf{x})$ and $\tilde{h}(\mathbf{q})$ be sampled at M^l regularly spaced Euclidian grid positions $\mathbf{X}_i \Delta x$ in direct space and $\mathbf{Q}_i \Delta q$ in Fourier space with $\mathbf{Q}_i,\ \mathbf{X}_i \in \mathbb{N}_0^l$. For the sampled representations of $h(\mathbf{x})$ and $\tilde{h}(\mathbf{q})$ by the vectors (or "arrays") **h** and $\tilde{\mathbf{h}}$, respectively, we formulate the direct DFT as

$$\tilde{h}_{\mathbf{Q}_i} = \mathrm{DFT}\big[\,\mathbf{h}\,\big]_{\mathbf{Q}_i} \tag{1.39}$$

$$= \left(\frac{1}{\sqrt{M}}\right)^l \sum_{X_1=0}^{M-1} \sum_{X_2=0}^{M-1} \dots \sum_{X_l=0}^{M-1} h_{\mathbf{X}} \exp\left(-2\pi\mathrm{i}\,\frac{\mathbf{Q}_i\mathbf{X}}{M}\right) \tag{1.40}$$

and the inverse DFT as

$$h_{\mathbf{X}_i} = \mathrm{IDFT}\big[\,\tilde{\mathbf{h}}\,\big]_{\mathbf{X}_i} \tag{1.41}$$

$$= \left(\frac{1}{\sqrt{M}}\right)^l \sum_{Q_1=0}^{M-1} \sum_{Q_2=0}^{M-1} \dots \sum_{Q_l=0}^{M-1} \tilde{h}_{\mathbf{Q}} \exp\left(2\pi\mathrm{i}\,\frac{\mathbf{Q}\mathbf{X}_i}{M}\right)\,. \tag{1.42}$$

As can be seen from (1.40) and (1.42) the calculation of the DFT for an entire array with $J = M^l$ the computational complexity is $O(J^2)$. The fast Fourier transform (FFT) algorithm expresses the DFT as a product of sparse matrices reduce the computational complextiy to $O(J\log(J))$ [21].

Let Δx and Δq denote the increments in direct and inverse space, respectively. Δx and Δq are related by the identity

$$\Delta x \cdot \Delta q = \frac{2\pi}{M}\,. \tag{1.43}$$

When substituting the Fourier integrals by discrete Fourier transforms (1.40) and (1.42) we must add the prefactors $(\Delta x)^l$ and $(\Delta q)^l$ respectively to obtain correct scaling units.

An upper limit for the choice of the sampling increment Δq for minimal information loss is well defined through Shannon's sampling theorem [85]. Given a function $\tilde{h}(\mathbf{q})$ that is band-limited[10] within $[0, s]$, i.e. $h(|\mathbf{x}| > s) = 0$,

---

[10] The term "band-limited" means that the signal of the function is bound to a finite interval of frequencies.

the discrete Fourier transform (1.40) gives the identical result as the continuous Fourier transform (1.20) if $\Delta q \leq \Delta q_c$ with

$$\Delta q_c = \frac{2\pi}{s}\,. \tag{1.44}$$

We conclude that for recovering $\varphi(\mathbf{x})$ from $\Psi^{(1)}(\mathbf{q})$ we require $\Delta q \leq \Delta q_c$ with s being the largest dimension of the sample. For following considerations we define the linear sampling ratio

$$\kappa = \frac{\Delta q_c}{\Delta q}\,. \tag{1.45}$$

and note that $\kappa = 1$ at the Shannon limit. We denote this special case "critical sampling".

1.3.3 Oversampling

In crystallography, the regular arrangement of unit cells in the crystal lattice gives rise to constructive interference of the diffraction signal in Bragg peaks. In 1952, Sayre realised[11] that sampling the intensities at the positions of the Bragg peaks is equivalent to sampling the modulus of the Fourier transform of the intensities (i.e. the Patterson function) at the Shannon limit. He proposed that phase information is contained in the diffraction pattern itself if the intensities between the Bragg peaks could be measured [82]. The diffraction pattern from a crystal is typically too weak to be interpretable between the Bragg peaks. But in plane-wave CDI the signal is continuous and phase information can be recovered from the *oversampled* (i.e. $\kappa > 1$) diffraction pattern. By adjusting wavelength and diffraction geometry κ can be relatively freely adjusted. For example an increase of detector distance or wavelength or a decrese of particle size or detector pixel size results generally in an increase of κ. For example for the small-angle far-field diffraction pattern from an isolated particle of size s it is

$$\kappa = \frac{r\lambda}{ps}\,, \tag{1.46}$$

where r denotes the detector distance and p the pixel size. The lower sampling limit for phase retrieval from the oversampled intensity pattern[12] is $\kappa = \kappa_o \geq 2$. This may be intuitive due to the simple fact that a sinus wave doubles its

---

[11] This realisation was not entirely new. Already in 1938, Bernal *et al.* [10] came to a similar conclusion during studies on Haemoglobin.

[12] To avoid unnecessary complexity we describe oversampling here only for the 1D case. The arguments can be extended to higher dimensions. For example for a square object with edge length s it is $\kappa_o \geq \sqrt{2}$.

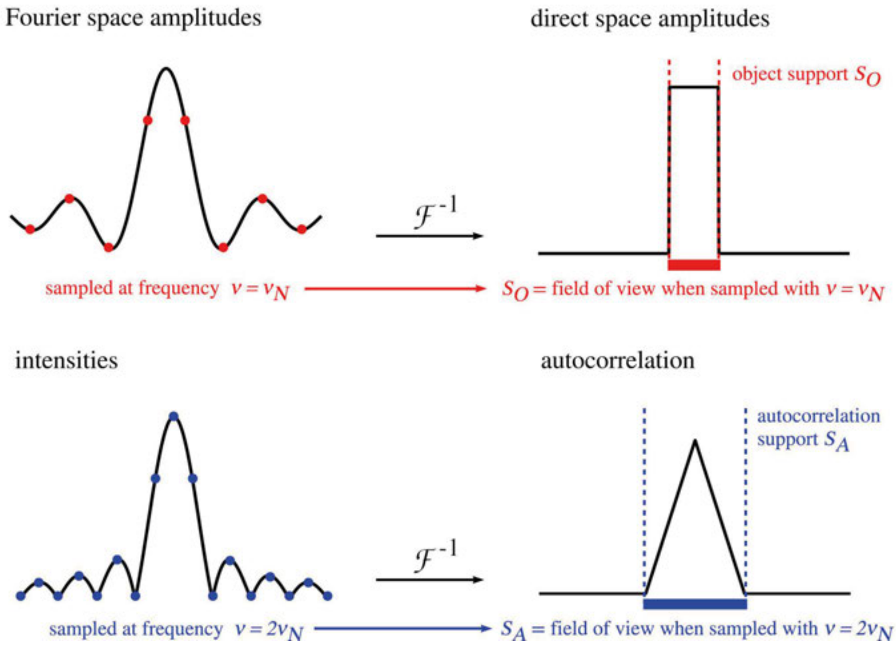


Figure 1.3. Shannon band-limits for the spectra of a top-hat function and its autocorrelation.

wave number by squaring

$$\sin^2(kx) = \frac{1}{2}(1 + \cos(2kx)) \tag{1.47}$$

and therefore critical sampling of intensities requires twice as many sampling points as critical sampling of amplitudes. A more rigorous argument is based on the realisation that the phase problem is equivalent to the problem of knowing the object's autocorrelation $R(\mathbf{d})$ because of the mathematical identity

$$R(\mathbf{d}) = \int h(\mathbf{x} + \mathbf{d})\bar{h}(\mathbf{x})\mathrm{d}^3\mathbf{x} = \mathscr{F}^{-1}\left[\left|\tilde{h}(\mathbf{x})\right|^2\right]. \tag{1.48}$$

This argument is exemplified in Fig. 1.3 for a top-hat function in one dimension. The extent of the autocorrelation $R(d)$ is for any object $h(x)$ at most twice as large as the object itself. This means that $\left|\tilde{h}(q)\right|^2$ is band-limited within $[0, 2s]$ and hence for critical sampling of the spectrum of the autocorrelation we reach to the same conlusion as above $\kappa_o \geq 2$.

Another argument can be made with basic algebra. The phase problem can be written as M^l equations

$$\left|\tilde{h}_{\mathbf{Q}_i}\right| = \left|\mathrm{DFT}\left[\,\mathbf{h}\,\right]_{\mathbf{Q}_i}\right|. \tag{1.49}$$

for $2 \cdot M^l$ unknown real and imaginary elements of the vector **h**, which has the lenght M^l. For this problem it was shown that the possibility of equations being linear combinations of others is in practice and for 2D and higher dimensions rare [6]. But still, many solutions may exist. Fortunately, they represent the same object because they are equivalent up to a constant phase offset, a translation in real space or a centrosymmetric inversion through the origin. For identifying at least one of these solutions we must constrain the problem by reducing the number of unknowns by at least M^l. By defining the *support* S as the set of points inside the known boundary of the object we formulate the *support constraint* as a new set of equations

$$\forall \mathbf{x} \notin S : h(\mathbf{x}) = 0 \,. \tag{1.50}$$

The support constraint (1.50) enforces "padding" of the object domain with zeros, which results in oversampling of $|\tilde{h}(\mathbf{q})|$ with $\kappa = (2 \cdot M^l)/M^l = 2$ and with the argument of Shannon sampling we reach to the same conclusion as above $\kappa_o \geq 2$.

1.3.4 Iterative phase retrieval algorithms

In 1972 Gerchberg and Saxton established the foundation for the most successful iterative phase retrieval algorithms in CDI until today [36]. Their strategy was to solve a related but slightly different problem by progressively approaching the solution of the phase problem in iterations of Fourier transformations back and forth between the object and the Fourier domain while repeatedly applying the respective constraints. Gerchberg and Saxton demonstrated that their algorithm was successful in the case of known amplitudes not only in the Fourier domain but also in the real space domain.

In 1978 Fienup applied the Gerchberg-Saxton scheme to the phase problem under the conditions that we face in CDI, i.e. known amplitudes in the Fourier domain and known support in the object domain [30]. He found the algorithm to be equivalent to the steepest-descent gradient search method, which monotonously decreases the error in every step, and hence named his algorithm error reduction algorithm (ER) [31]. Fienup was able to prove that ER converges [31] but as illustrated in Fig. 1.4, to succeed with ER from any starting point in the search space the constraint sets must be convex. The support constraint set is indeed convex but the intensity constraint set is not, which makes ER in our application prone to getting trapped in local minima and never reaching the solution. For evading this problem Fienup proposed the hybrid input-output algorithm (HIO) [31]. The iterate of HIO has a negative feedback term leading to repulsion from shallow minima.Convergence could not be proven for HIO, but the algorithm has been demonstrated successful in many applications [32]. Nevertheless, HIO struggles if the support is not precisely known or if the object wave is complex-valued. In those cases modifications of the HIO can prove more robust. A popular modification is the shrink-wrap algorithm [68], which progressively refines the support by thresholding in the object domain. Popular variations of the phasing iterate include for example the relaxed averaged alternating reflections algorithm (RAAR) [61] and the difference map algorithm (DM) [26].

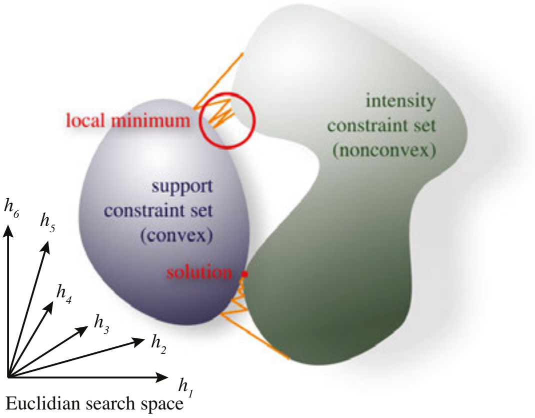


Figure 1.4. The iterative phasing scheme in Euclidian search space with twice as many dimensions as number of pixels. During an iterative phase search support constraint and amplitude constraint are applied by projecting to the closest point of the respective constratint set in Euclidian space. As the intensity constraint set is nonconvex algorithms such as ER that reduce the error in every step can be trapped in local minima depending on the starting point of the phase search.

1.3.5 Additional constraints

With iterative phase retrieval algorithms there is no guarantee of finding the correct solution. Usually, many phase searches from different starting points are carried out to identify and validate the solution. Robustness can be improved by applying additional object-specific constraints. Commonly used constraints are the *reality constraint*

$$\mathrm{Im}(h_{\mathbf{X}}) = 0\,, \tag{1.51}$$

which constrains the phase of the object to be either 0 or π and the *positivity constraint*

$$\mathrm{Re}(h_{\mathbf{X}}) \geq 0 \text{ and } \mathrm{Im}(h_{\mathbf{X}}) \geq 0, \tag{1.52}$$

which restricts the phase to lie within $[0, \pi/2]$. These constraints require that the maximum phase shift and the absorbance of the object of study are sufficiently small (see formulas (1.17) and (1.18), respectively). Note that due to the fact that the reality constraint implies Friedel symmetry and the positivity constraint conflicts with large phase ramps in the object domain, these constraints put the additional requirement of an accurately known center position of the diffraction pattern.

1.3.6 Missing data

Often it cannot be avoided that beam stops, gaps between detector tiles, or faulty or saturated areas of the detector obscure parts of the diffraction pattern. Such missing information is fundamentally problematic and can result in ambiguities and artefacts in the reconstruction [94, 84].

The impact of missing data can be estimated mathematically on the basis of the set of obscured pixels D in the Fourier domain and the set of pixels S within the support in the object domain [94, 84]. For the sake of simplicity, we restrict ourselves here to the one-dimensional case, which can be generalised to more dimensions.

We will make use of the fact that the Fourier transform is a linear transformation and can be expressed as a matrix multiplication. Our goal is to employ singular value decomposition (SVD) to identify the least constrained components of the Fourier transform, which have most power in unconstrained sets of pixels S and D. SVD factorizes any matrix $\underline{\mathbf{M}}$ into the unitary matrix $\underline{\mathbf{U}}$, the diagonal matrix $\underline{\Sigma}$, and the unitary matrix $\underline{\mathbf{V}}^H$ such that the columns $\mathbf{u}_i$ of $\underline{\mathbf{U}}$ are the orthogonal eigenvectors of $\underline{\mathbf{M}}\underline{\mathbf{M}}^H$, the columns $\mathbf{v}_i$ of $\underline{\mathbf{V}}$ are the orthogonal eigenvectors of $\underline{\mathbf{M}}^H\underline{\mathbf{M}}$, and the diagonal entries $s_i = \Sigma_{i,i}$ denote the singular values. s_i represent the squared eigenvalues of both $\underline{\mathbf{M}}\underline{\mathbf{M}}^H$ and $\underline{\mathbf{M}}^H\underline{\mathbf{M}}$ and follow descending order.

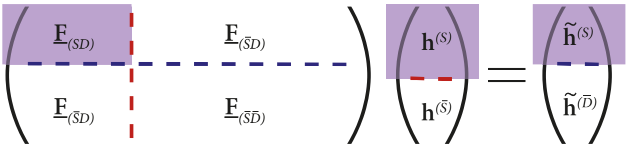
We want to apply SVD to the matrix $\underline{\mathbf{F}}_{(SD)}$, which describes the transformation of unconstrained values in the object domain $\mathbf{h}^{(S)}$ to unconstrained values $\tilde{\mathbf{h}}^{(D)}$ in the Fourier domain.

$$\underline{\mathbf{F}}_{(SD)}\,\mathbf{h}^{(S)} = \tilde{\mathbf{h}}^{(D)}\,. \tag{1.53}$$

For constructing $\underline{\mathbf{F}}^{(SD)}$ we make use of the DFT in its matrix notation

$$\underline{\mathbf{F}}_{i,j} = \left(\frac{\mathrm{e}^{2\pi\mathrm{i}\frac{ij}{M}}}{\sqrt{M}}\right)_{i,j=0,\dots,M-1}\,. \tag{1.54}$$

The order of the vector coordinates in $\mathbf{h}$ and $\tilde{\mathbf{h}}$ is arbitrary and may be altered if the column and the row numbers of the Fourier matrix are manipulated accordingly. If we arrange $\mathbf{h}$ and $\tilde{\mathbf{h}}$ such that the pixels in S and in D come first we obtain an expression for (1.53) with the structure



We may omit the matrix entries $\bar{S} \to D$ because $\mathbf{h}^{(\bar{S})}$ denotes the vector of values in the object domain outside the support and they are all zero. The desired expression for (1.53) can be obtained by truncating the Fourier matrix such the the elements highlighted in violet remain. The matrix that is left is $\underline{\mathbf{F}}^{(SD)}$, which is typically sufficiently small to be decomposed by SVD routines on a standard computer.

The singular values s_i for $\underline{\mathbf{F}}_{(SD)}$ allow us to estimate how unconstrained each mode is (**Paper II**) and from the eigenvectors we can predict artefacts in reconstructions that are likely associated with missing-data [94, 84].

We acknowledge that this analysis implicitly assumes perfect knowledge of phases outside the missing data region and this assumption is not always met. It is possible, but not guaranteed, that phases outside the missing data region are recovered successfully by iterative phasing methods. Importantly, missing data can interfere with the iterative phase retrieval in many different ways. For example if large amounts of missing data are present at low resolution it can be difficult to recover the support [84].

In conclusion, iterative phase retrieval is effected by missing data in different ways. Some of the effects can be estimated. For obtaining reliable reconstruction results the extent of missing data must be minimised and reconstructions must be tested with rigorous validation methods.

1.3.7 Validation

There is no guarantee that an iterative phase search from a single set of random starting phases will lead to the correct solution. Yet, typically failed searches can be identified by their high Fourier error

$$E_F = \sqrt{\frac{\sum_i \left(\left|\tilde{h}_i\right| - \sqrt{I}\right)^2}{\sum_i \left(\sqrt{I}\right)^2}} \tag{1.55}$$

and discarded. The real space error metric

$$E_R = \sqrt{\frac{\sum_{i \notin S} |h_i|^2}{\sum_{i \in S \cup \bar{S}} |h_i|^2}}\,. \tag{1.56}$$

is a measure for the integrated density outside of the support and can be used to detect an incorrect support constraint. High real space errors often indicate an incorrect support or even more dramatically the disintegration of the support.

Residual fluctuations of reconstruction results can be associated to the common problem that the "best" solution, i.e. the solution that satisfies best the given constraints, is not the "true" solution. This missmatch is attributed to noise in the Fourier domain and non-zero density ("background") outside the support. To guard against missinterpretations of such structural fluctuations and to obtain the most representative reconstruction the average image is computed from the results of many successful (i.e. with low E_F) phase searches from different random phase seeds. By quantifying the reproducibility of the individual reconstructions we define an upper limit for the reliability of the phase-retrieval result. For the phase of every pixel i the phase retrieval transfer function (PRTF) [15, 16] can be calculated from J individual reconstructions j as

$$\mathrm{PRTF}_i = \frac{1}{J} \sum_{j=0}^{J-1} \frac{\tilde{h}_i^{(j)}}{\left|\tilde{h}_i^{(j)}\right|} \,. \tag{1.57}$$

If the phases of pixel i from J individual reconstructions match well, the PRTF approaches 1. In contrary, if the phases are entirely random the expectation value for the PRTF is $1/\sqrt{J}$. The resolution criterion where the radially averaged PRTF drops below e^{-1} is commonly used [15]. This resolution criterion is most meaningful if J is large such that $1/\sqrt{J} \approx 0$.

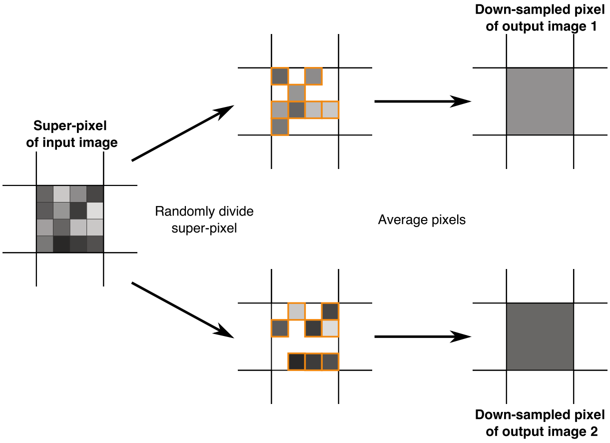


Figure 1.5. Image splitting for Fourier ring correlation. (Supp. Fig. 2 in **Paper I**)

A high PRTF is an indicator for reproducible phasing. Yet, the reprodubility may be just the result of overfitting to noise. We show in **Paper I** how the FRC, which is typically used to validate EM reconstructions [97], can be applied to FXI to guard against overfitting. Given a sufficiently oversampled diffraction pattern, two binned diffraction patterns A and B can be obtained that differ due to noise in the data. As illustrated in Fig. 1.5, for each 4×4 "super-pixel" pixel values are split into two random sets. For the patterns A and B values $I_i^{(A)}$ and $I_i^{(B)}$ are obtained for every super-pixel i by separately averaging the pixel values of each set. We then carry out iterative phase retrieval independently from these two patterns and compare the results. The degree of agreement of the reconstructed images $\tilde{\mathbf{h}}^{(A)}$ and $\tilde{\mathbf{h}}^{(B)}$ is scored as a function of resolution by using the formula for the FRC given in ref. 97

$$\mathrm{FRC}(q) = \frac{\sum\limits_{i|q_i \in q} \left(\tilde{h}_i^{(A)} \cdot \tilde{h}_i^{(B)*} \right)}{\sqrt{\sum\limits_{i|q_i \in q} \left| \tilde{h}_i^{(A)} \right|^2 \sum\limits_{i|q_i \in q} \left| \tilde{h}_i^{(B)} \right|^2}} \quad . \tag{1.58}$$

The FRC signals the risk of overfitting if it drops below the threshold of 0.5 [97] at lower resolution than the resolution estimated on the basis of the PRTF.

2. Creation and injection of aerosols from suspensions of bioparticles

2.1 Substrate-free sample delivery

Substrate-based sample delivery of biological samples for FXI has been tested with moderate success [62, 54]. The presence of a sample container or substrate is associated with background scatter and any source of background noise is to be avoided for reaching atomic resolution[1]. This is especially important for imaging small objects such as single proteins that produce very faint diffraction patterns. Consider also that contact to the substrate typically affects structure and orientation. Finally, for taking full advantage of the rapid repetition rates of modern XFELs sample exchanges within less than a microsecond must be reached and this seems impractical with substrate-based techniques.

Sample injection techniques ([11, 84] and **Paper I**) lift the requirement for a particle substrate and circumvent most associated problems by delivering the sample as a narrow and dense stream of aerosolised particles. The following chapters introduce basic concepts and implications of the technique.

2.2 Particle aersolisation

2.2.1 Droplet formation

As a means of transfering sample particles from the liquid to the gas phase our first objective is the massive creation of small droplets by atomisation. The term *atomisation* denotes the process of disintegrating liquids into airborne droplets [44]. More specifically *jet atomisation* is the name for techniques that produce a steady stream of droplets at the tip of a laminar liquid jet that originates from an orifice [44]. The liquid jet breaks up into droplets due to an effect called the Rayleigh instability [57]. For droplets thus formed the characteristic diameter d_d is proportional[2] to the jet diameter d_j, which normally

---

[1] In crystallography the background from the substrate is not as much a concern as here because scattering signal from every unit cell is amplified by the crystal lattice. For single particle techniques, such as FXI and also cryo-EM, the signal from the structure under investigation is not amplified and contributes equally as the signal from the substrate.

[2] The ratio $d_d/d_j \approx 1.9$ (classical Rayleigh break-up) matches well observations for low viscosity systems like water [20]. A comprehensive and detailed discourse about the proportionality between d_j and d_d in relation to the dimensionless Weber number can be found in ref. 35.

roughly equals the inner diameter of the jet-feeding nozzle. By applying vibrations near the optimal break-up frequency it is possible to synchronise droplet break-up and improve monodispersity [57, 24].

For suspensions of non-interacting evenly-distributed particles of concentration c_d that are atomised into monodisperse populations of droplets of volume V_d the droplet occupancy k_d follows a Poisson distribution (1.36) with $k = k_d$ and $\Lambda = c_d V_d$ [44, 56].

$$p_{k_d} = \frac{(c_d V_d)^{k_d} \exp(-(c_d V_d))}{k_d!} \tag{2.1}$$

It follows that decreasing the jet diameter increases the production of singlet droplets ($k_d = 1$). Yet, in practice shrinking the jet by shrinking the nozzle comes at the price of the increased risk of clogging. We can overcome this problem by replacing the firm walls of the nozzle by a smooth potential ("virtual nozzle") that tapers the wide liquid stream into a narrow cone-jet. In Fig. 2.1 two realisations of this concept are shown. Fig. 2.1a shows the functional principle of gas dynamic flow-focussing (FF) [34, 24] that employs a sheath gas to taper the liquid stream. The gas flow generates a pressure gradient that fuels the inertia of the jet. The principle of cone-jet electrospray ionisation (ESI) [92, 100] is depicted in Fig. 2.1b. Here an electric field with voltage ΔV is applied to the liquid and associated Maxwell surface-stress shapes the cone and drives the jet.

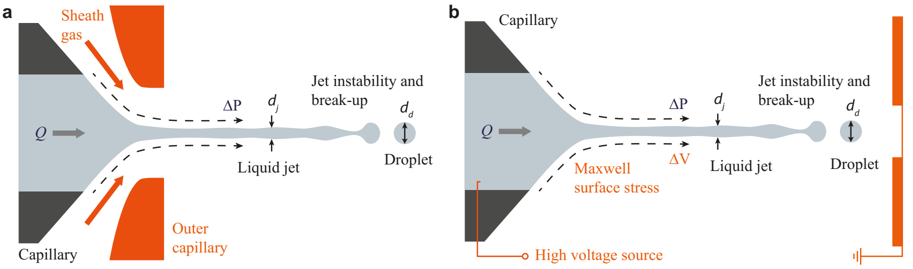


Figure 2.1. Steady tip streaming from a cone jet realised by gas dynamic flow-focussing (FF) (**a**) and electrospray ionisation (ESI) (**a**) through a pressure drop ΔP. Instability at the tip leads to the break-up of the liquid jet into droplets of diameter d_d proportional to the jet diameter d_j.

For both techniques the jet diameter d_j can be estimated on the basis of energy conservation. Thus obtained expressions for d_j can be formulated as functions of an effective pressure drop ΔP [35]. These formulas exhibit good resemblance with experimental data and indicate that the mechanisms of droplet formation for FF and ESI are closely related despite the fact that the nature of the potentials that drive the jet are fundamentally different [35].

For deriving the formulas for d_j let Q denote the flow rate of the liquid, ρ its density, and U the characteristic jet velocity. We assume that practically all energy $Q\Delta P$ is transformed into kinetic energy $Q\rho U^2$ and for FF. Under this assumption it can be shown [35] that

$$d_j^{(FF)} = 2 \cdot \sqrt{Q \cdot \left(\frac{\rho}{2\pi^2 \Delta P}\right)^{1/2}} . \tag{2.2}$$

From (2.2) follows that $d_j^{(FF)}$ can be reduced by decreases of the flow rate Q and increases of the pressure drop ΔP (observe footnote 2).

For deriving the jet diameter $d_j^{(ESI)}$ for cone-jet ESI we make use of Taylor's electrostatic solution of the cone [92], which delivers a relationship for the electric field along the jet direction as

$$E_j \sim \left(\frac{\sigma}{\varepsilon_0 L}\right)^{1/2} , \tag{2.3}$$

where σ denotes the surface tension of the liquid, ε_0 the vacuum permittivity, and L an axial characteristic length[3]. We assume that the voltage drop happens over L such that $\Delta V \sim E_s L$ and the electrical current intensity in the jet is $I \sim E_s K (d_j/2)^2$, where K is the electrical conductivity. As above we neglect any energy losses and assume instead that all electric energy $I\Delta V$ is converted into kinetic energy $Q\rho U^2$. It can be shown [35] that

$$d_j^{(ESI)} \sim 2 \cdot \sqrt{Q \cdot \left(\frac{\rho\varepsilon_0}{\sigma K}\right)^{1/3}} . \tag{2.4}$$

From (2.4) follows that decreases of Q and increases of K represent a useful route for decreasing $d_j^{(ESI)}$.

As mentioned above, $d_j^{(ESI)}$ can be expressed equivalently to FF as a function of ΔP. By combining (2.2) and (2.4) we obtain

$$\Delta P = k_p \left(\frac{\sigma^2 K^2 \rho}{\varepsilon_0^2}\right)^{1/3} , \tag{2.5}$$

where k_p is a constant of order unity.

Obviously, (2.2) and (2.4) only apply under the premise of the stablility of the cone-jet. The cone-jet breaks down when the kinetic energy becomes comparable to either the energy of viscous dissipation or the reduction of free surface by transformation into a blunt meniscus tip. The stability limits can be derived with similar energy scaling arguments as used for (2.2). For liquids

---

[3] L is larger than d_j and smaller than the feeding capillary orifice.

characterised by viscosity μ, surface tension σ, and mass density ρ one obatains at the respective stability limits

$$Q_\mu \sim \sqrt{\frac{\mu^4}{\rho^3 \Delta P}} \text{ and} \tag{2.6}$$

$$Q_\sigma \sim \sqrt{\frac{\sigma^4}{\rho (\Delta P)^3}} \tag{2.7}$$

[35]. We notice that Q_μ and Q_σ are not only functions of the material-specific parameters but also of ΔP. This means that not only the droplet size but also the stability limits are tunable by nozzle design and the applied sheath gas pressure (Fig. 2.1a).

Charged droplets exhibit an additional atomisation mechanism in train of evaporation and electrostatic repulsion [20]. The term *droplet fission* denotes this phenomenon. Droplet fission is driven by the increasing Coulomb repulsion of surface ions in shrinking droplets. At the point when surface tension is overcome by the electrostatic forces the droplet undergoes uneven droplet fission by ejecting a stream of small droplets. Cycles of evaporation and fission events result in the formation of smaller and smaller droplets.

2.2.2 Particle desolvation

After droplet creation evaporation removes excess solvent from the particle and ideally results in the formation of the bare sample particle in its native conformation. Results from mass spectrometry (MS) [8] and molecular dynamics simulations [96] indeed indicate that aerosolised biological particles can maintain intact despite almost full deprivation of their solvent (mainly structural waters remain). However, it was shown by ion mobility-mass spectrometry (IMMS) that high amounts of charges on the sample can induce conformational bias [49]. Yet in contrast to MS the charge is not needed for FXI and may be removed. Oppositely charged air molecules, which can be produced with ionising radiation sources such as ^{210}Po or ^{85}Kr, can be used to discharge and ultimately neutralise the charged aerosol particles [44].

2.3 Particle focussing

2.3.1 Hit ratios

For making efficient use of the sample the particle beam is ideally matched to the extent of the X-ray focus. In FXI experiments the hit ratio, which is the fraction of X-ray pulses that result in measurable diffraction signal above background, is proportional to the density of particles c_x in the given interaction volume V_x. In most cases we can safely assume that interactions between particles are negligible. In such cases we expect that the probability for illuminating k_x particles with an X-ray laser pulse follows a Poisson distribution (1.36) with $k = k_x$ and $\Lambda = c_x V_x$.

$$p_{k_x} = \frac{(c_x V_x)^{k_x} \exp(-c_x V_x)}{k_x!} \tag{2.8}$$

The interaction volume can be estimated as $V_X = (d_x)^2 d_p$, where d_x and d_p denote the diameters of the X-ray beam and the particle beam, respectively. The density of particles $c_x = F/(v_p(d_p)^2)$, where F denotes the particle flux and v_p the particle velocity. It follows (**Paper I**)

$$\Lambda = c_x V_x = \frac{F(d_x)^2}{v_p d_p}. \tag{2.9}$$

It must be noted that, in practice, d_x depends on the particle size because the diffraction signal and the detectability increases with particle size. Hence, large particles effectively sample a seemingly bigger interaction volume than small particles and are in practice overrepresented in FXI data sets (**Paper I**).

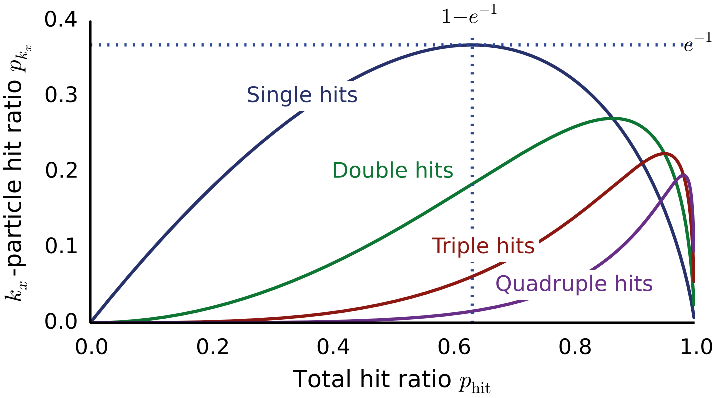


Figure 2.2. Hit ratios p_{k_x} for $k_x = 1$ (single hits), 2 (double hits), 3 (triple hits), and 4 (quadruple hits) as functions of the total hit ratio p_{hit}. (Figure adopted from Fig. 2d in **Paper I**)

For real data, where it is often not immediately obvious from a diffraction pattern how many particles were hit simultaneously, we define the "total hit ratio" p_{hit} as the fraction of X-ray pulses that hit one or more particles. In this case p_{hit} can be calculated as

$$p_{\text{hit}} = \sum_{n=1}^{\infty} \frac{\lambda^k e^{-\Lambda}}{n!} = e^{-\Lambda}\left(\sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} - 1\right) = 1 - e^{-\Lambda} = 1 - e^{-c_x V_x}. \tag{2.10}$$

By substiution of (2.10) in (2.8) we obtain the equation (**Paper I**)

$$p_{k_x} = \frac{\ln(\frac{1}{1-p_{\text{hit}}})^{k_x} \cdot (1 - p_{\text{hit}})}{k_x!}, \tag{2.11}$$

which expresses the hit ratio p_{k_x} for k_x particles in terms of the total hit ratio p_{hit} (Fig. 2.2). We find that the theoretical maximum for the single-particle hit ratio is $p_1^* = 1/e \approx 36.8\,\%$ and it is reached at the total hit ratio of $p_{\text{hit}}^* = 1 - e^{-1} \approx 63.2\,\%$ (**Paper I**).

2.3.2 Aerodynamic lenses

As illustrated in Fig. 2.3, by guiding an aerosol through a serial axisymmetrical arrangement of cylindrical cavities, which are inter-connected by co-aligned orifices, particles can be collimated to a narrow particle beam [74]. Through the repeated confinement of the gas flow through the orifices a pressure profile is generated that effectively drags particles towards the central axis. Particle beams created by aerodynamic lenses can be very narrow with beam widths, which are solely limited by the effects of Brownian motion and the lift forces that act on non-spherical particles [58].

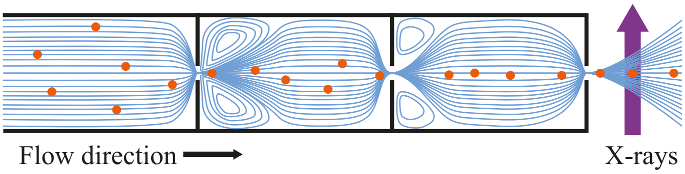


Figure 2.3. Schematic representation of the generation of a narrow particle beam from an aerosol by a series of orifices. The orange circles represent particles and the blue lines isobars (i.e. lines of constant pressure).

Hit ratios can be optimised by tuning operational parameters such as input and output pressure of the aerodynamic lens stack. These parameters influence particle transmission, particle speed and the width of the particle beam focus.

Part III:
Realisation

3. Proof of concept

In 2005, the VUV-FEL (later known as FLASH) started lasing. Already in the following year researchers succeeded in demonstrating the concept of "diffraction before destruction" at this facility [15]. From a single-shot exposure a diffraction pattern of a silicon nitride test structure was obtained and despite the fact that the shot entirely obliterated the sample, an image of the intact test structure could be reconstructed from the intensities. In both experiments, a graded multilayer mirror reflected the diffraction pattern onto an orthogonally positioned CCD detector. The graded multilayer mirror separated the diffracted beam from the direct beam, which then passed through an apotised hole in the centre of the mirror without damaging the detector. A couple of years later the first image of a dry biological cell deposited on a membrane was obtained in a similar set-up at the same facility [62]. With gas dynamic virtual nozzles (GDVNs), which were described for the first time in 2008, it was possible to efficiently atomise sample suspensions into micron to submicron droplets without charging the sample [24]. In 2009, the construction of the LCLS was completed and provided high repetition rates (120 Hz), harder X-rays (photon energies between 0.3 to 9.6 keV) and more intense pulses (currently ca. 1×10^{12} photons in a micron to sub-micron focus). The same year, researchers coupled a GDVN to the Uppsala sample injector and succeeded in transfering airborne Mimivirus particles into the pulse train of the LCLS. This lead to the first publication demonstrating imaging of single injected biological entities with an XFEL [84]. The experiment on Mimivirus was a breakthrough but challenges became apparent such as low hit ratios, relatively low resolution (32 nm), and problems in the reconstruction due to about 8 to 12 missing low-resolution modes [84].

4. Experiment on carboxysomes

4.1 Carboxysomes

For the study described in **Paper I**, we used carboxysomes as our test sample. The polyhedral micro-compartments facilitate fixation of carbon dioxide in cyanobacteria and in certain proteobacteria [86, 80]. Carboxysomes are fragile organelles and appear icosahedral in the electron microscope [46]. They show a broad range of sizes (between 90-500 nm), depending on the source organism. In the present study, we used carboxysomes from *Halothiobacillus neapolitanus*, and these carboxysomes have a mean diameter of about 115 nm (Fig. 1b in **Paper I**).

Variations in their size hinder crystallisation. A large body of information indicates that they have an outer protein shell, consisting of 3 to 5 different proteins (for reviews see refs. 80 and 28). The outer shell encapsulates a large number of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) molecules (11 nm diameter), which may form paracrystalline arrays inside the organelle [80]. Many functional and structural details of these organelles remain unknown.

4.2 Data collection

The experiment described in **Paper I** was carried out at the atomic, molecular and optical science (AMO) end station [14, 13] of the LCLS [27] and Fig. 4.1 shows the arrangement of the experiment. Far-field diffraction patterns were recorded on a pair of fast pnCCD detectors [89] at the 120 Hz repetition rate of the LCLS. The intense primary beam passed through a narrow gap between the two detector halves and was absorbed in a beam dump at a distance behind the detectors.

The X-ray pulses contained 2.2×10^{12} photons/pulse on average at a photon energy of 1096 eV (1.131 nm wavelength). Pulses were focused to a spot of about 5 µm in diameter (full width at half maximum) at the interaction point, giving 6.8×10^{10} photons/μm^2 in the centre of the beam, assuming a Gaussian beam profile. The pulse length was 120 fs long (full duration at half maximum). The detector was placed 741 mm downstream from the interaction point. The maximum full-period resolution at the edge of the detector was 21.5 nm, and it reached 15.2 nm at the corners. In this configuration, the linear oversampling ratio κ for a 100-nm object is 50-fold. This allowed us to

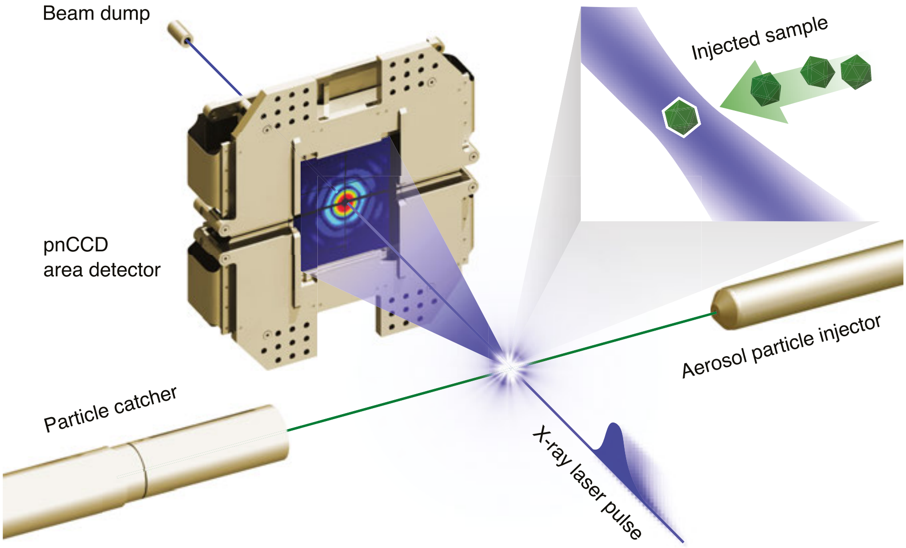


Figure 4.1. Experimental set-up. A beam of particles is created with the aerosol particle injector and intersected with the X-ray laser beam. Particles that are exposed to the X-ray laser beam give rise to a diffraction pattern that is recorded with a pnCCD X-ray area detector. The direct X-ray beam passes through a gap between the two detector halves and is absobed by a beam dump downstream without damaging the detector. (Figure adapted from **Paper I**.)

perform direct phase retrieval from the measured diffraction intensities. For particles smaller than 210 nm, low-resolution modes were well constrained by the diffraction data.

A total of 86945 exposures were recorded in 12 minutes, containing 68810 hits at 79 % hit ratio.

4.3 Data analysis

In silico sample purification

We automated the sorting procedure of diffraction patterns of single particles by exploiting the fact that at very low resolution a spheroid can be used to approximate the shape of most particles. The central speckle in the measured diffraction patterns was fitted with an equivalent speckle from a spheroid, and the goodness of the fit was assessed. This step separated spheroid-like objects from more complex morphologies. A total of 13419 hits out of the 21000 strongest hits were adequately modelled by spheroids and passed this filter. Analysis of the filtered hits shows two sub-classes: hits on isolated single particles (10563 exposures), and hits on spatially-separated multiple particles with similar sizes (2856 exposures, see center panel of Fig. 3 in **Paper I**). Patterns that could not be fitted by a spheroid contained hits primarily on aggregates and hits on multiple particles with very different diameters (7581 exposures, see right panel of Fig. 3 in **Paper I**). These latter classes were not analysed further.

The photon flux experienced by a particle varies from shot to shot. We derived the photon flux at the particle from the number of scattered photons in the diffraction pattern and from the size and shape of the spheroid fitted to the pattern. We assumed a particle density of 1 g/cm^3 and an overall composition of $H_{23}C_3NO_{10}S$ (similar to that of a cell [9]). The size distribution that we obtained from strong hits on single particles is in good agreement with the independently measured size distribution of carboxysomes in solution (black curve 3 in Fig. 4b of **Paper I**).

Image reconstruction

The automatic parameterisation of diffraction patterns sorted hits according to the size and shape of the objects. Fig. 4.2a shows reconstructed projection images for three different size ranges. Phases for these reconstructions were retrieved reproducibly to at least 21.5 nm full-period resolution, which is the resolution at the edge of the detector. The reconstructions were carried out automatically with the Hawk software package [65], using a combination of the iterative phase retrieval algorithms shrink-wrap [68], RAAR [61], and ER [31].

The diffraction pattern of a single carboxysome and the reconstruction of the projection image is shown in Fig. 4.3a and compared to a simulated data for an uniform icosahedron in matching orientation shown in Fig. 4.3b. The good match between the two indicates the success of delivering carboxysomes into the XFEL beam without altering the stucture significantly up to the resolution attained here. Fig. 4.3a and Fig. 4.3b illustrate that at 18.1 nm resolution, the injected carboxysome can be well approximated by a solid icosahedron of uniform density. The diameter of a Rubisco molecule is about 11 nm, and these molecules cannot be resolved at 18.1 nm resolution.

The PRTF in Fig. 4.3c (red curve) shows phases were reproducibly retrieved to 18.1 nm full-period resolution in the reconstruction. The Fourier ring correlation (blue curve in Fig. 4.3c) does not drop below 0.5 over the entire range of scattering angles and thus shows that the good reproduciblility of the reconstruction indicated by the PRTF is not attributed to overfitting to noise.

Objects with straight edges were found in the size domain expected for our carboxysomes (100-130 nm, Fig. 4.2a (centre), and Fig. 4.2b). Objects outside this size domain were predominantly spherical. Carboxysomes were purified on the basis of their size and density from a cell culture, and such impurities

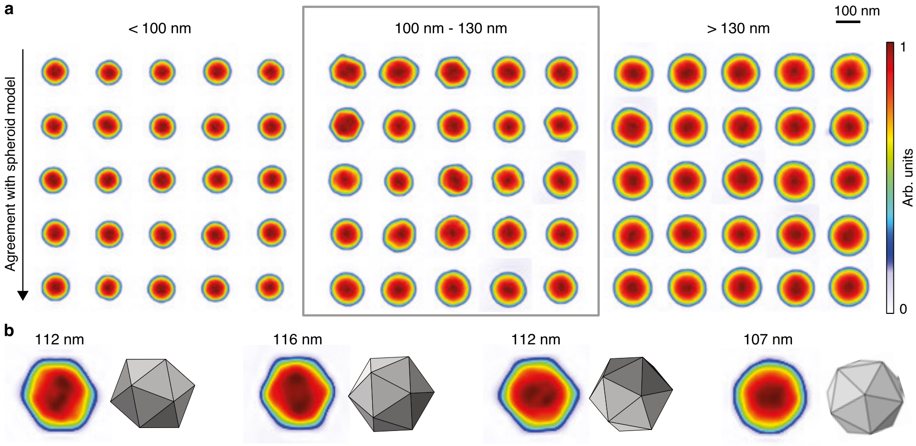


Figure 4.2. Computationally sorted particles according to size and shape. **a** Automatically reconstructed projection images of single particles grouped by particle size and sorted according to shape (top to bottom). The figure shows a representative selection of 75 reconstructions from a total of 10,563 strong hits on single particles. **b** Comparison of selected reconstructions with images of icosahedra rotated to match the projections. (Figure adopted from **Paper I**.)

are not completely unexpected. The results show that we can separate contaminants from carboxysomes.

4.4 Results

The diffraction data were purified *in silico* and analysed automatically, and we achieved higher resolution reconstructions than we previously achieved on Mimivirus particles [62] even though carboxysomes have two-orders of magnitude smaller scattering power. This improvement is only partially due to the increased photon flux on the sample. There are three main factors at play: (i) our diffraction measurements constrain all low-resolution modes, (ii) the measurements have negligible noise, and (iii) the objects are truly isolated in the beam. These factors lead to robust phase retrieval without the usual artefacts, and establish flash-diffractive imaging as a reliable imaging method. We also achieved significantly higher hit ratios. This achievement is attibuted in parts to the improved injector design, alignment, and operation. But the increase in hit ratio may be also an indication for residue particles from non-volatile sample contaminants (this is discussed in more detail in ch. 5).

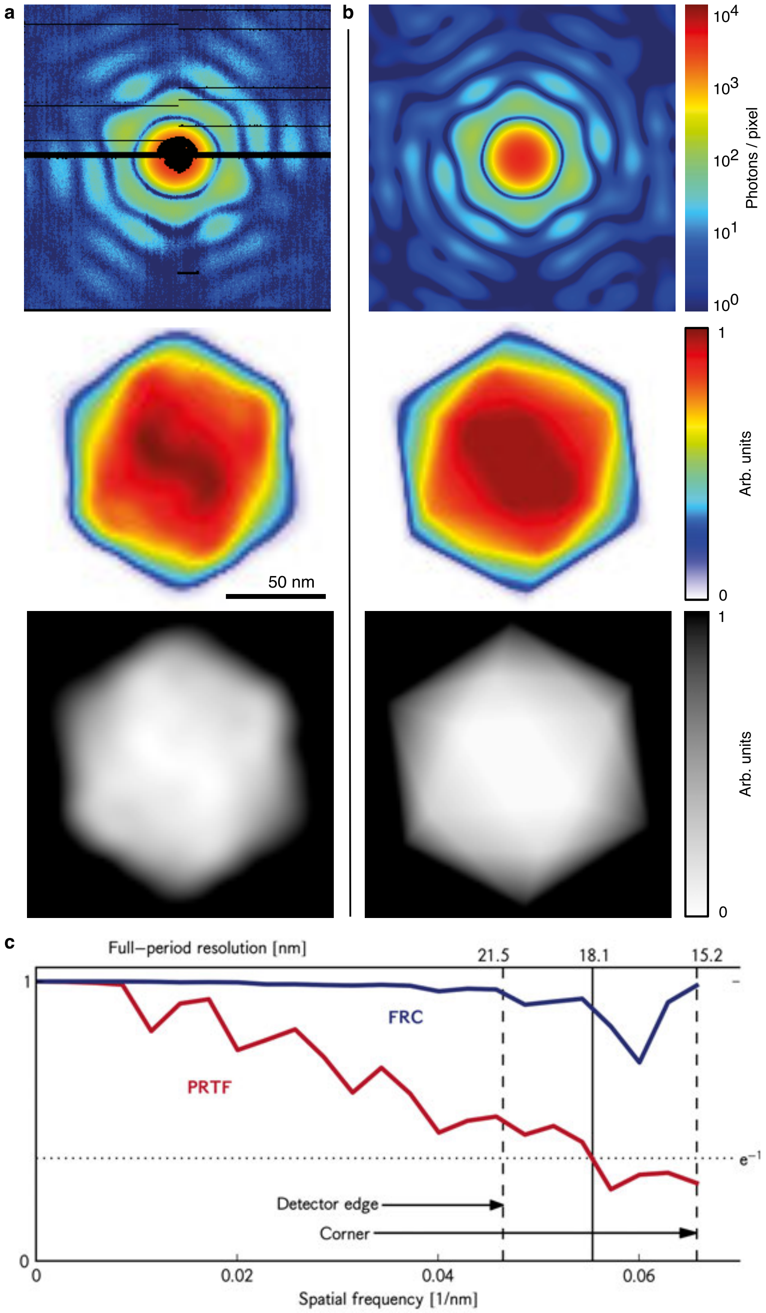


Figure 4.3. Projection image of a carboxysome particle reconstructed from a single diffraction pattern. The measured diffraction intensities (**a** top) and the reconstructed projection image (**a** bottom) are well approximated by the simulated diffraction pattern (**b** top) and the projection image (**b** bottom)for a uniform icosahedron. The resolution is estimated to 18.1 nm where the radially averaged PRTF (red curve in **c**) drops below $1/e$. The FRC does not indicate problems with overfitting as all values of the FRC are well above 0.5. (Figure adapted from Fig. 6 in **Paper I**.)

4.5 Data deposition

XFEL beam time is scarce and therefore many researchers lack access to experimental FXI data. By publishing data descriptors and depositing experimental data online [64] FXI data can be made accessible to a wide community. Making FXI data available aids making more use of the collected data, helps to make published results more reproducible, and is essential for developers who want to test the rebustness of their algorithms against real data. We deposited both raw and pre-processed data of the the data set on carboxysomes on the Coherent X-ray Imaging Data Bank (CXIDB) (http://www.cxidb.org/id-25.html) and published a data descriptor (**Paper II**) that includes further experimental details, the precise structure of the deposited files, and data validation.

5. Improvements on sample injection

5.1 The Uppsala sample injector

The Uppsala sample injector (Fig. 5.1) was developed to introduce free-flying particles into the pulse train of X-ray lasers at a reduced pressure. Purified particles are transferred into a volatile buffer (for carboxysomes 20 mM ammonium acetate, pH 7.5) and aerosolized with helium gas, using a GDVN. The sample consumption for carboxysomes was 2-4 µl/min from a solution of 1.2×10^{11} particles/mL. The aerosol enters the injector via an inlet nozzle coupled to a skimmer. Excess nebulising gas is pumped away at this stage. The concentrated aerosol passes through a relaxation chamber from where the adiabatically cooled particles enter an aerodynamic lens.

In the data on carboxysomes (**Paper I** and **II**), most particles were round and this is unexpected if the sample had been perfectly pure. Carboxysomes from *Halothiobacillus neapolitanus* have straight edges and an approximately icosahedral appearance [46]. A majority of round particles of unknown origin were also observed in other FXI experiments [51, 23]. In all these experiments GDVNs were employed for aerosolisation and the sample particle was smaller than 150 nm. We hypothesise that that relatively large droplets from GDVNs (droplet diameters of about 1-2 µm) may contain sufficient quantities of impurities to cause this effect.

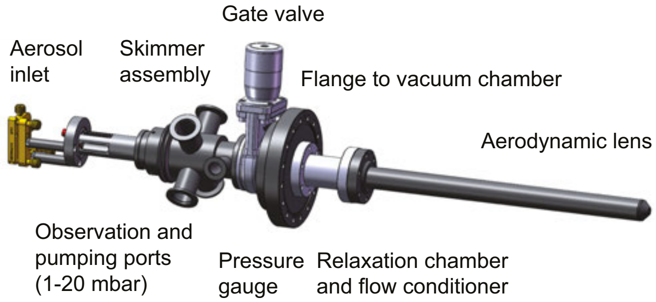


Figure 5.1. The Uppsala sample injector. (Figure adopted from Fig. 2a in **Paper I**.)

5.2 Droplet evaporation

Droplets are created by jet atomisation in the nozzle box, which is a small chamber that is coupled to the aerosol inlet. In the nozzle box prevail room temperature and pressures of 100-1000 mbar, depending on settings. The observation that dusting spots at the exit of the injector appear about one to a few seconds after initiation of the GDVN jet and the fact that particles are accelerated by the aerodynamic lens to high velocities[1] give us reason to believe that before particles are sucked into injector segments with lower pressure, droplets have a time window of the order of seconds to evaporate. The time evolution of steady-state droplet evaporation is described by the mass transport equation

$$\frac{\mathrm{d}(d_d)}{\mathrm{d}t} = \frac{4DM}{R\rho_d d_d}\left(\frac{p_\infty}{T_\infty} - \frac{p_d}{T_d}\right) \cdot \phi(d_d)\,, \tag{5.1}$$

where d_d is the diameter of the droplet, D is the diffusion coefficient, M the molar mass, ρ is the mass density of the liquid, p_∞ and p_d the vapor pressures and T_∞ and T_d the temperatures far away and at the surface of the droplet, respectively [44, 48]. $\phi(d_d)$ is the Fuchs factor, which compensates for kinetic diffusion effects at the droplet boundary [44]. The curvature of the liquid surface elevates p_d with respect to the saturation vapor pressure (called Kelvin effect [44]) and drives in combination with kinetic effects the evapoartion of submicron droplets. Temperature depression of micron-droplets as a consequence of evaporative cooling slows down evaporation and may under certain conditions (e.g. vacuum) fully stall droplet evaporation [98].

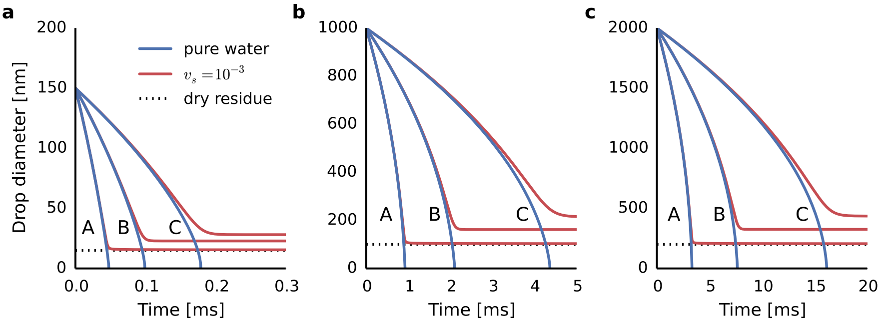


Figure 5.2. Evaporation time of small water droplets. Evolution of the droplet diameter d_d over time with initial diameters 150 nm (**a**), 1000 nm (**b**), and 2000 nm (**c**), respectively. Predictions were made for pure water (blue lines) and for water with a volume fraction of $v_s = 10^{-3}$) of non-volatile fully soluble contaminants (solid red lines). Calculations were performed for air atmosphere at standard conditions and relative humidities with 0% (A), 50% (B), and 75% (C). The horizontal dotted line indicates the diameter of the solid dry residue that would remain after complete evaporation of all solvent and under the assumption that the mass density of the contaminant equals the mass density of the solvent.

---

[1] The particle velocity depends on injector settings and the particle size. For Mimivirus particle velocities of ca. 100 m/s are anticipated [84]. We have recently started measurements of particle velocities by Mie scattering imaging (see ch. 5.5).

Non-volatile impurities (including the sample particle itself) lead to a reduction of evaporation speed. This effect can be roughly estimated for fully soluble non-volatile contaminants of mass fraction m_s in the initial droplet and initial droplet diameter $d_d^{(0)}$ as a reduction of p_d by the factor $\exp[-m_s/((d_d^{(0)}/d_d)^3 - m_s)]$ [44, 48].

In Fig. 5.2 we show the evolution of the droplet diameter obtained from numerically integrating (5.1) at standard conditions for initial droplets diameters $d_d^{(0)} = 150\,\mathrm{nm}$, $1000\,\mathrm{nm}$, and $2000\,\mathrm{nm}$ (Fig. 5.2a, b, and c). We assumed a steady-state temperature depression of the droplet surface by using for T the formula given by ref. 44 (p. 278). The data in Fig. 5.2 indicate evaporation times of 0.05-15 ms depending on humidity and initial droplet size. In theory, droplets can be considered evaporated when they arrive seconds after their creation at the interaction point. This is also in line with the observation that no particles emerge at the exit of the injector when running pure water (**Paper I**). This conclusion presumes reasonably low humidities and relatively low contaminant mass fractions m_s (see red curves in Fig. 5.2).

5.3 Contaminant residues

Samples may contain non-volatile contaminants, the amount of which may change, depedning on the purification procedure and on processes of sample particle disintegration. The bigger the droplet around the sample particle, the more contaminantns will be delivered with the sample. As the droplet shrinks and evaporates, the contaminants settle on the surface of the sample and this is refered to as “caking”.

For estimating the size of contaminant residues after evaporation of a droplet let us consider a spherical droplet of initial diameter $d_d^{(0)}$. If the droplet contains no sample particle and only contaminant particles the diameter of the residue d_s can be calculated as [102]

$$d_s = (v_s)^{1/3}\, d_d^{(0)}\,, \tag{5.2}$$

where v_s denotes the contaminant’s volume fraction in the initial droplet. If one sample particle of diameter d_p occupies the droplet the contaminants will, after full evaporation of the solvent, coat the particle with a layer that has an approximate thickness

$$\Delta_s = \frac{1}{2}\left[\left(d_p^3 + ((d_d^{(0)})^3 - d_p^3)\, v_s\right)^{1/3} - d_p\right]\,. \tag{5.3}$$

The variable v_s can be expressed in quantities that are more commonly used when characterising substances in solution. For example

$$v_s = \frac{c_s}{\rho_s}\,, \tag{5.4}$$

where c_s denotes the mass concentration of the contaminant in the solution and ρ_s denotes the mass density of the dry contaminant.

Values for d_s and Δ_s, predicted with (5.2) and (5.3), respectively, are plotted in Fig. 5.3 for two initial droplet diameters (1000-2000 nm and 150 nm) as functions of the contaminant volume fraction v_s. Δ_s is calculated for the sample particle diameters $d_p = 115$ nm (solid lines) and $d_p = 35$ nm (dashed lines).

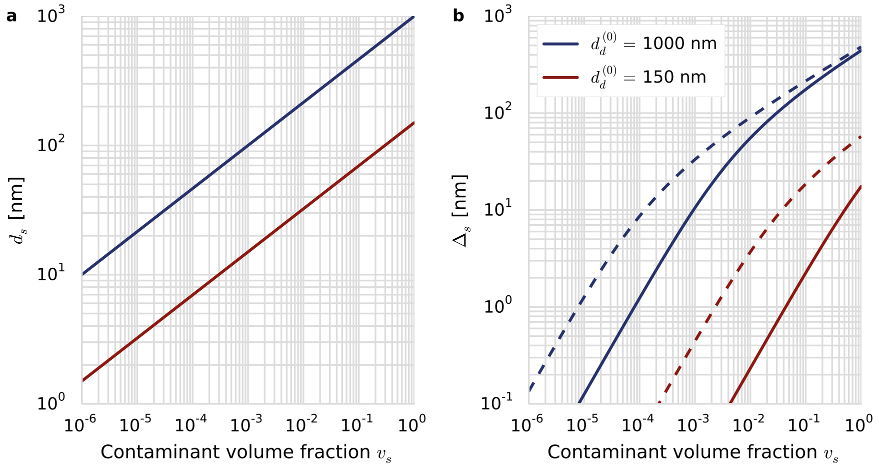


Figure 5.3. Contaminant residues from evaporated droplets with initial diameters $d_d^{(0)} = 1000$ nm *(blue lines) and* 150 nm *(red lines).* **a** the diameter d_s of non-volatile contaminant residues from droplets that do not contain a sample particle (5.2) is plotted as a function of contaminant volume fraction v_s. **b** For droplets that contain one spherical particle of diameter $d_p = 115$ nm (solid lines) or diameter $d_p = 35$ nm (dashed lines), the thickness Δ_s of a uniform layer of contaminants "caked" on top of the particle is plotted as a function of contaminant volume fraction v_s.

For a given sample, the volumetric fraction of non-volatile impurities can be measured with a differential mobility analyser (DMA). Measuring volume fractions[2] v_s for contaminants of around 10^{-3} were measured for sample suspensions that appeared pure in the electron microscope and in nanoparticle tracking analysis (NTA). For initial droplet sizes of 1000 nm equation (5.2) predicts ca. 100-nm-sized contaminant residues from droplets that do not contain a sample particle (Fig. 5.3a). For droplets occupied by one sample particle of 115 nm diameter the impurities would account for a coating layer of ca. 7 nm

---

[2] Contaminant may consist of proteinaceous fragments from broken particles ($\rho_s = 1.35\,\mathrm{g/cm^3}$), of sucrose as residuent from density gradient centrifugation ($\rho_s = 1.59\,\mathrm{g/cm^3}$), or salt from the buffer solution (for sodium chloride it is $\rho_s = 2.16\,\mathrm{g/cm^3}$). According to (5.3) we would expect for these contaminant substances and $v_s \approx 10^{-3}$ contamint concentrations c_s between ca. 0.5 and 1 mg/ml.

thickness (Fig. 5.3b). These predictions roughly match the observations described in **Paper I** and could also explain the more severe problems described for experiments with smaller particles (ref. 51 and 23). Note however, that droplets created with GDVNs are not monodisperse (see Fig. 5.5 following chapter). Diffraction patterns and reconstructions that matches expectations for shape and size of carboxysomes (Fig. 4.3a and 4.2b) could be explained by small droplets in the size distribution of droplets generated by a GDVN.

If these conclusions are correct a reduction in droplet size would improve data quality significantly. The amount of contaminant residue per particle would be negligible far beyond current limits of resolution.

5.4 Smaller droplets

Cone-jet ESI is a viable technique for producing sub-micron droplets and the huge success of MS and IMMS on proteins, and protein complexes demonstrates the strength and maturity of the technique as a method for transfering biological particles into the gas phase [8]. The reduction of capillary diameters from about 500 µm down to about 1 µm made it possible to reduce flow rates to only 1-100 nl/min [100] and obtain initial droplet diameters d_d of 150-200 nm routinely [8]. nano electrospray ionisation (NESI) is compatible with samples that are available only in small quantities. Furthermore, the drastic decrease in droplet size made the method "softer" and also more tolerant towards contamination [50].

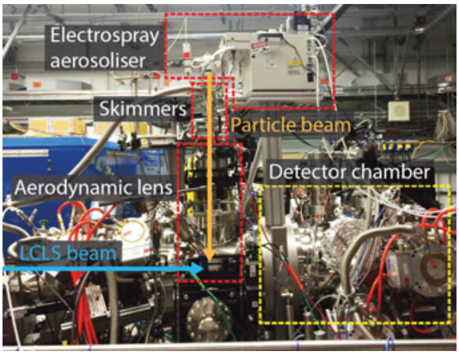


Figure 5.4. Realisation of electrospary injection at the AMO end station at the LCLS.

We replaced the GDVNs with an adapted ESI source (capillary size of 20 µm, flow rate 50-100 nl/min) to the Uppsala sample injector (see Fig. 5.4). We were able to show experimentally a reduction in droplet size from up to 2000 nm for injection with a GDVN to about 150 nm for injection by ESI (see Fig. 5.5). In June 2016, we received beam time at the AMO end station of the LCLS and were able to reach with ESI aerosolisation similar hit ratios as previously achieved with GDVNs. We injected samples in the size range between 12 nm

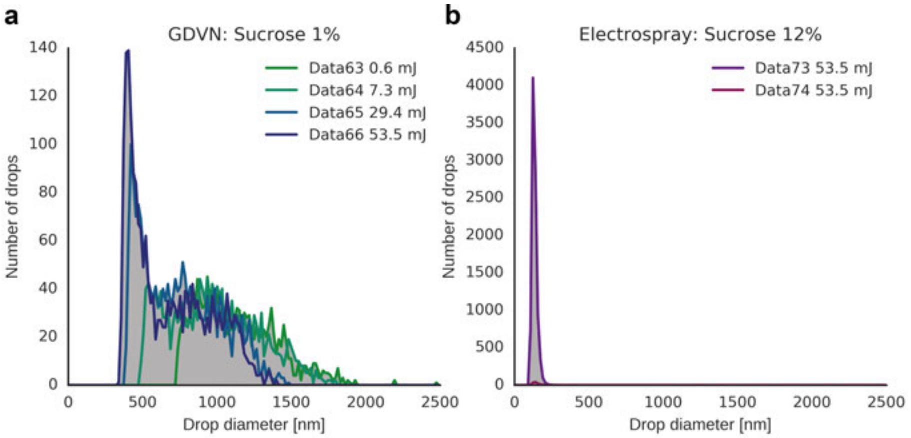


Figure 5.5. Droplet size distributions aerosolised by injection with a ***a)*** *GDVN and* ***b)*** *ESI* . The data were obtained by MSI on injected sucrose solution residues. The sucrose in a droplet of sucrose solution is not volatile and by measuring the size of the surcrose residue, which is left after droplet evaporation, the initial droplet size was determined. The measurements were carried out with sucrose volume fractions v_s of 1 % (**a**) and 12 % (**b**). Each curve represents a data set measured at a constant pulse energy, which allows detection while avoiding saturation for a particular size range of sucrose residues. For more details on the method of MSI see ch. 5.5.

and 100 nm (see Fig. 5.6). Collected diffraction data on small virus particles (35 nm in diameter) resembles favourably simulated data (Fig. 5.6a and b) and the size distribution almost exactly matches expectations. Most autocorrelations functions that we obtained from diffraction patterns on cartboxysomes indicate straight edges and thus indicate the success in reducing problems with non-volatile contaminants. The diffraction data (top row in Fig. 5.6d) extends to higher resolution than reached previously (**Paper I** and **II**). Analysis of these data is ongoing.

5.5 Injection diagnostics

For accelerating injector development and for gaining more precise control over injection, precise lab-based diagnostic tools that do not require an XFEL are indispensable. Rayleigh and Mie scattering analysis are standard techniques for determining size distributions of particles in two-phase systems such as sprays and aerosols [47]. Light scattering analysis is used in a wide range of applications such as research in atmospheric science [70] and studying the physics of droplets [45, 48].

A recent publication [4] shows the application of Mie scattering imaging to determining the positions and velocities of individual micron-sized particles injected with the Uppsala aerosol injector. We have designed a similar set-up

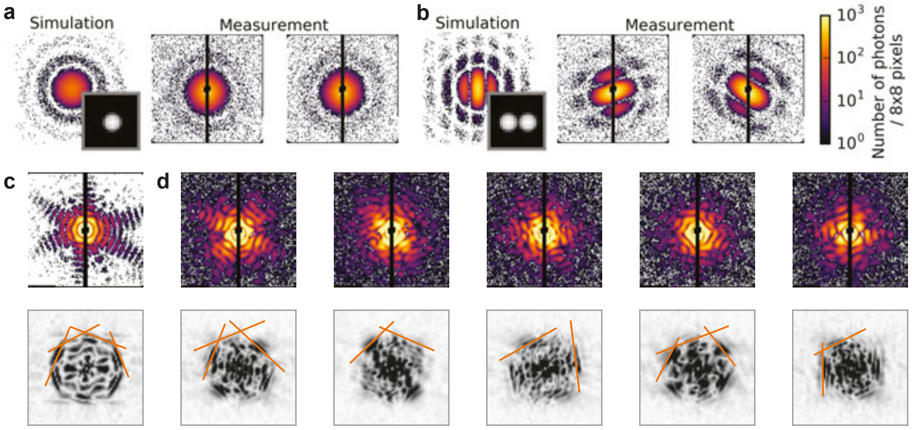


Figure 5.6. Realisation of electrospary injection at the AMO end station at the LCLS in June 2016 (experiment *LL34*). The data indicate that both size and shape for small virus particles and carboxysomes are maintained during injection by electrospray. Simulated and measured diffraction patterns for single hits (**a**) and a double hits (**b**) of spherical 35-nm-sized virus particles. Measured and simulated diffraction data match well indicating successful electrospray injection of these small virus particles. The insets show projection images of the sample models that were used for the simulations. **c** Simulated diffraction pattern (top) and autocorrelation (bottom) for a 100 nm-sized uniform icosahedral carboxysome model. **e** Recorded diffraction images (top) and autocorrelation (bottom) of *Halothiobacillus neapolitanus* carboxysomes. Autocorrelation images shown here were high-pass filtered to reduce artefacts from missing data. Orange lines highlight positions and orientations of straight edges.

with the ambition to image even smaller particles and assess besides positions and velocities also the particle sizes and intial droplet sizes[3]. Fig. 5.7a shows the geometry of our Mie scattering imaging set-up. The particle aerosol is illuminated by a vertically poliarised double-pulsed green laser ($\lambda = 532$ nm) with pulse energies of 50 mJ and with down to sub-microsecond inter-pulse delays. Dark field images are collected at the 25 Hz repetition rate with a microscope (0.055 numerical aperture (NA)) coupled to a CMOS area detector with 82% quantum efficiency. The Mie scattering signal of particles is the strongest in forward direction and falls off with increasing scattering angle (left panel in Fig. 5.7b). We were able to reduce in our set-up the scattering angle to 20°, which allows us to detect polystyrene spheres down to 40 nm. At this scattering angle we are able to estimate particle sizes unambiguosly from the particle brightness for particle sizes to up to ca. 1000 nm (Fig. 5.7b). In the Rayleigh scattering regime (i.e. if $d_p \ll \lambda$) the scattering signal scales with the sixth power of d_p and therefore permits very accurate assessments of relative parti-

---

[3] By atomisation of sucrose solution of known concentration evaporation of the droplet's water results in the formation of residues of defined size. By measuring the size of the residue the size of the original droplet can be infered with (5.2) [102].

cle size differences (see right panel in Fig. 5.7b), which is important to detect the nanometer-thin impurity layers anticipated from non-volatitle impurities (see equation (5.3)). In-house results for carboxysomes are shown in Fig. 5.7c and d. The beam profile and the size histogram that we obtained match expectations. The results demonstrate the viabilitiy of the method for sub-micron-sized biological samples. A manuscript about this work is in preparation.

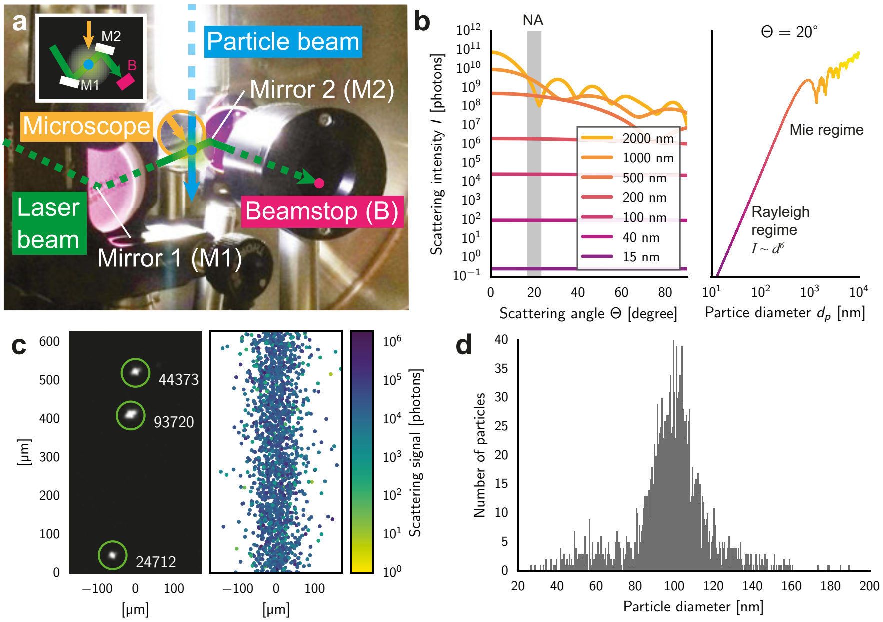


Figure 5.7. Mie scattering imaging as a diagnostics tool for sample injection. **a** An optical laser beam is intersected with the particle beam that exits the aerosol particle injector. Light emanating from the illuminated particles due to Mie scattering is imaged with a microscope at a scattering angle of 20°. After the direct beam has passed the interaction region it is reflected away from the microscope lens and absorbed by a beamstop. **b** Estimated brightness for single cell organelles imaged with this set-up as a function of scattering angle and diameter. The gray area spans the interval of scattering angles covered by the aperture of the microscope. In the right panel scattering intensities averaged over this interval of scattering angles are shown as a function of paricle diameter. For a particle size of about 1000 nm this function reaches a maximum indicating an upper size limit for unambiguos size determination in this configuration. **c** The left panel shows a single dark field image recorded on injected carboxysome particles. The numbers denote the integrated measured signal in number of scattered photons. The right panel shows from 500 images the particle positions and their respective (color-coded) intensities. **d** Size histogram of injected particles. Sizes were extracted from the same data that are shown in **c** using the scaling law between scattering signal and particle size shown in **b**.

6. Software

In a significant part of this project, I worked on software development and implementation of new computational tools to enable successful FXI experiments, including the planning and execution of the experiments, analysis of the data, and interpretation of the results.

6.1 *Condor*: Data prediction

For advancing the method to its current technological limits meticulous planning of beam times and optimisation of the set-up is required. The software package *Condor* (**Paper III**, https://github.com/FXIhub/condor) predicts FXI data under realistic conditions and is used to estimate the feasibility of FXI experiments, to optimise parameters for imaging experiments, to execute model-guided data analysis, and to test and validate reconstruction and classification algorithms and monitoring software. The software exploits fast implementations of discrete Fourier transform algorithms, and can be used from an online platform (http://lmb.icm.uu.se/condor) without the requirement of a local installation. Particle models can be created from structues determined by electron microscopy or X-ray crystallography or on the basis of uniformly filled bodies of common shapes (e.g. icosahedrons, spheroids, cubes, etc.).

6.2 *Hummingbird*: Online monitoring

For FXI experiments at XFELs procedures must be established that make it possible to judge data quality not after the beam time but in real time. This enables researchers to adjust critical parameters (e.g. sample concentration, injection pressures, injector position, detector settings etc.) for reaching most favorable experimental conditions early in the beam time. The software package *Hummingbird* (**Paper IV**, https://github.com/FXIhub/hummingbird) was developed for online monitoring of FXI experiments at XFEL end stations. Besides pre-processing the raw readout from detectors for display the software can be used to analyse data from the raw data stream and generate for example particle size histograms, hit scores, or autocorrelation images that can be used to judge data quality and characteristics while data are being collected. The user configures the program by providing a single *Python* file. This permits easy adjustment of the functionality to specific requirements of different experiments. The software has a modular architecture and is subdivided into backend and frontend processes. The backend analyses the raw data stream and makes the output data stream accessible from the network. Independent frontend processes can subscribe to output data streams for displaying results. *Hummingbird* has been used at experiments at the LCLS at the AMO end station and the coherent X-ray imaging (CXI) beamline [73] and was tested recently also at FLASH.

6.3 *Cheetah*: Data pre-processing

During XFEL experiments data rates of 100 Hz and more are reached and large data volumes of many tens of terrabytes are generated during a beam time. Due to the large size of these data sets data analysis poses computational challenges and benefits from specialised and efficient pre-processing software. We have developed the open-source software package *Cheetah* (**Paper V**, https://github.com/antonbarty/cheetah), which can rapidly identify hits, pre-process diffraction images, and generate meta data such as histograms of pixel values, noise maps, and virtual powder patterns. Outputs are written to files in the Hierarchical Data Format version 5 (HDF5) [93]. The output file structure complies with the guidelinges for the CXIDB [64]. This standardisation faciliates data deposition and makes it easy to interface *Cheetah* with other FXI/SFX software.

6.4 *Owl*: Data visualisation

After data pre-proceessing the first step of data analysis involves skimming through large amounts of data, inspecting diffraction patterns and selecting promising data subsets. We develeoped the open-source software *Owl* (https://github.com/FXIhub/owl), which provides a graphical user interface to browse, inspect, tag, sort, and filter diffraction patterns. *Owl* can read HDF5 files that comply with the CXIDB file structure [64].

6.5 *Hawk*: Phase retrieval

Image reconstruction by iterative phase-retieval algorithms involves many DFTs and is therefore computationally demanding. For **Paper I** the open-source software package *Hawk* (https://github.com/FXIhub/hawk) with the underlying library *spimage* (https://github.com/FXIhub/libspimage) [65] was used and extended. *Hawk* provides rapid implementations of iterative phasing algorithms that can exploit the high performance of FFTs on graphics processing units (GPUs) [22].

Part IV:
Summary and outlook

Achievements

In the course of this work we have collected, analysed, and deposited large amounts of FXI data and developed new methodology and specialised software. We hope that these are valuable contributions for the quest of transitioning FXI from its first experimental demonstration into a technique that fulfills its potentials.

In **Paper I** we exemplified with heterogeneous cell organelles how tens of thousands of FXI diffraction patterns can be collected, sorted, and analysed in an automatic data processing pipeline. Compared to reconstructions on Mimivirus particles [84], we improved resolution and were able to practically overcome issues with missing data. We validated, described, and deposited the experimental data in the Coherent X-ray Imaging Data Bank (Paper II).

We demonstrated that aerosol injection can be used to collect FXI data at high hit ratios, with low background, and with negligible missing data (**Paper I**). Yet, problems with GDVN injection have been identified for particles smaller than ca. 100 nm [51, 23]. We suspect that these problems are associated with residue formation from non-volatile impurities in sample suspensions. The load of non-volatile residue per aerosolised particle can be effectively reduced by decreasing the droplet size. In a recent experiment at the LCLS we successfully reduced droplet sizes from ca. 1-2 µm to about 150 nm by replacing the GDVN with a new ESI aerosoliser. We reached similar hit ratios as with GDVNs, and initial results show no signs of residues from non-volatile impurities even for small particles of 35 nm in diameter. With this improvement the breakthrough of imaging particles as small as single proteins may be imminent and may be possible even from the data already collected on single Rubsico proteins (11 nm in diameter). Yet, for reaching atomic resolution with FXI, the technical prerequisites in terms of photon flux and background noise have to be improved at XFEL beam lines. Fortunately, higher photon fluxes are anticipated with the advent of the European X-ray Free-Electron Laser (EXFEL) [83]. To improve conditions at the LCLS, the Single Particle Imaging Initiative was set up, bringing together many collaborative institutes [3].

To test, characterise, and improve injector design and operation in the lab, a diagnostic tool became necessary that does not require an XFEL beam. In our lab we set up Mie scattering imaging, which enables us to measure positions, sizes, and velocities of individual injected particles in flight without the requirement of an XFEL beam. Our MSI implementation primarily requires a powerful pulsed optical laser and a microscope and can be integrated into existing beam lines. MSI is a powerful tool for the real-time monitoring of sample injection during FXI experiments.

XFEL experiments generate large amounts of data at high rates. Preparation, execution, and data analysis of these experiments benefits from specialised software. We have developed open-source software tools that facilitates prediction (**Paper III**), online-monitoring (**Paper IV**), display (*Owl* software), and pre-processing (**Paper V**) of XFEL diffraction data.

Challenges

Efforts of improving data quality in FXI are hampered by the fact that XFEL facilities are highly overbooked and provide limited opportunities for data collection. Furthermore, no single-purpose beamline exists, where the design can be optimised solely for the specific requirements of FXI (i.e. low background, low noise, small missing data, high flux, no extreme vacuum requirements). To improve conditions at LCLS, the Single Particle Imaging Initiative was set up, bringing together many collaborative institutes [3].

Currently, maybe the most obvious challenge is the limit in resolution due to the low signal-to-noise ratios at current XFEL beam lines (**Paper I**). Improvements are anticipated with the advent of the EXFEL [83] and with improvements by the Single Particle Imaging Initiative at the LCLS [3].

For reaching the goal of atomic resolution on heterogenous populations of structures we require large data sets, improved signal-to-noise ratios, more advanced *in silico* purification routines, and clever 3D alignment algorithms that can handle the large amounts of data.

Prospects

The rapid repetition rates anticipated for the EXFEL [83] will allow to collect diffraction patterns of many individual particles in short time. These large amounts of diffraction data sample conformational space very densely, which may permit us, if computational challenges can be overcome, reaching atomic resolution even from very heterogeneous structures. Furthermore, the rapid data rates may enable us to study the structural response of reactions subsequent to an external trigger (e.g. a pump pulse).

Employing aerosol injection for FXI implicates the potential for a number of promising future applications. One distinguishing feature is that the sample can be injected directly from solution and has the natural compatibility with numerous techniques for sample separation and sample mixing. Techniques such as liquid chromatography or capillary electrophoresis could be employed in conjunction with *in silico* purification (**Paper I**) for solving structures of inherently heterogeneous molecular species at high resolution. Microfluidic devices allow mixing of components in defined ratios could be applied to study the assembly process of viral capsids, DNA packing, protein folding, vesicle formation and more. Furthermore, the fact that the sample solution is atomised ("compartmentalised") into droplets may be used to trap and image transient complexes. Evaporation of the solvent gradually increases the effective concentration of the components, which may promote complex formation despite low affinity. Evaporation traps the complex structure and makes it amenable for investigation by FXI. Transient interactions are ubiquitous in biology and are challenging to study with existing techniques.

Results presented in this thesis open the way to overcome present limitations for FXI and to transition this imaging concept into a method that offers new research opportunities to study structure, heterogeneity, and dynamics in biology and elsewhere.

Sammanfattning på svenska

Biologisk funktion bygger på ett komplext nätverk av specifika interaktioner mellan molekyler. Dessa interaktioner sker på avstånd lika små som storleken av atomer (ca $1\,\text{Å} = 10^{-10}\,\text{m}$) och på tidsskalor som sträcker sig från år (1 år $= 10^7$ sekunder) ner till femtosekunder (1 femtosecond $= 10^{-15}$ sekunder) [43]. Tekniker för att studera biologiska strukturer med hög tids- och rumsupplösning är nyckeln till en djup förståelse av livet.

Med den nya, och snabbt växande, röntgenfrielektronlaser (X-ray free-electron laser, XFEL)-tekniken [63] har vi fått en strålkälla som har potential att revolutionera hur vi avbildar molekyler. XFELs producerar mycket starka och korta röntgenpulser (för närvarande upp till ca 10^{12} fotoner$/\mu\text{m}^2$ och 70 fs lång) med våglängder ned till en Å. Den korta våglängden tillåter i princip avbildning till atomär upplösning. Den första demonstrationen av laserverkan från en röntgenfrielektronlaser med Å-våglängder skedde 2009 [27]. Under de senaste fem åren har röntgenlasertekniken lett till anmärkningsvärda framsteg inom fysik, kemi, materialvetenskap och biologi.

Pulslängder på femtosekunder har rätt tidsskala för att fånga snabba biologiska processer. Dessutom är pulserna korta nog att ge upphov till röntgendiffraktion innan de orsakar strålskador som förstör provet [75]. Detta gör det möjligt att samla in data vid rumstemperatur utan cryo-fixering [17]. Vi kan undvika att använda en provhållare genom att injicera provet som en aerosol i fokus av strålen [11, 84]. XFEL-pulser är mycket intensiva och det förväntas att diffraktion i princip kommer att kunna mätas till atomär upplösning även från en partikel så liten som en enda molekyl [75, 17]. Detta mål har dock ännu inte nåtts. Större partiklar ger upphov till starkare diffraktion, från vilken 2D-projektionsbilder kan rekonstrueras med hjälp av iterativa fasbestämningsalgoritmer och förhandsinformation, som till exempel storleken på partikeln [15, 84]. 3D-strukturer kan åstadkommas genom en sammanslagning av diffraktionsdata från identiska kopior av partikeln tagna från olika riktningar, där datan sedan orienteras för att matcha [25]. Även för extremt svag diffraktionsdata från enskilda proteiner har 3D-strukturer åstadkommits från simulerad diffraktionsdata med hjälp av en iterativ 3D-orienteringsalgoritm och aggressiv signalbehandling [60].

År 2006 demonstrerades flash diffraktions-röntgenavbildning (flash diffractive X-ray imaging, FXI) experimentellt för första gången. Här användes ett konstgjort prov vid en frielektronlaser i Hamburg (FLASH, tidigare känd som VUV-FEL) [15]. År 2011 publicerades resultaten från ett experiment där biologiska partiklar injicerades. Provet var Mimivirus-partiklar och experimentet använde röntgenstrålar med högre energi och större fotonflöde vid LINAC

Coherent Light Source (LCLS) i USA [84]. Medan experimentet på Mimivirus var uppmuntrande i och med att det visade att konceptet fungerar, visade det också på tekniska utmaningar som är förknippade med FXI, såsom svårigheter med låg träffprocent och rekonstruktions-artefakter på grund av överlagrade och skymda regioner på röntgendetektorn. Dessutom är Mimiviruspartikeln en av de största kända virusarterna med en diameter på 450 nm och en molekylvikt av 28 GDa och ger upphov till en stark spridningssignal. Trots detta uppnåddes endast en upplösning på 32 nm. Försök att injicera betydligt mindre biologiska partiklar med samma aerosolteknik som användes för Mimivirus misslyckades på grund av att partiklarna bildade aggregat i stället för att injiceras som enstaka partiklar [51, 23]. Dessa svårigheter måste övervinnas på något sätt. Dessutom blev det uppenbart att ett antal nödvändiga hjälpmedel saknades, såsom allmänt tillänglig experimentell data och specialiserad mjukvara för datasimulering, realtidsövervakning av insamlad data, förbehandling av data och automatisk analys [67, 66]. Dessa problem innebar ett betydande behov av teknisk utveckling och förbättring för att FXI ska kunna nå sin fulla potential [3]. Denna avhandling tar upp många av dessa utmaningar och bidrar med nya metoder och ny programvara för att övervinna dem.

Under detta arbete har forskargruppen samlat in, analyserat, och tillgängligjort stora mängder XFEL-data och utvecklat nya metoder och specialiserad mjukvara. Det är min förhoppning att detta är ett värdefullt bidrag i vår strävan att utveckla metoden från den första experimentella demonstrationen till en teknik som utnyttjar dess potential till fullo.

I Artikel I har vi visat expempel på hur tiotusentals diffraktionsmönster från olikformade cellorganeller kan parametreras och analyseras med en automatiserad löpande band princip. Jämfört med tidigare rekonstruktioner av Mimivirus-partiklar [84] förbättrade vi upplösningen och kunde praktiskt taget övervinna de problem som uppstår när data är ofullständigt. Vi har tillgängligjort experimentell data i “Coherent X-ray Diffraktion Data Bank” och i Artikel II beskriver vi data, tekniska detaljer och visar resultat från datavalidering.

Vi har visat att aerosol-injektion kan användas för att samla FXI data med hög träffprocent, låg bakgrundssignal och en försumbar andel data som saknas (Artikel I). Trots detta har problem med aerosolbildning genom dynamiska virtuella gasmunstycken (gas-dynamic virtual nozzles, GDVNs) påträffats för partiklar mindre än ca. 100 nm [51, 23]. Vi misstänker att dessa problem är förknippade med utfällnining av icke-flyktiga föroreningar i vätskan runt provet. Mängden förorening per partikel kan minskas effektivt genom att minska droppstorleken. I senare experiment vid LCLS har vi framgångsrikt minskat droppstorleken från ca. 1000 nm till ca 150 nm genom att ersätta GDVN med en ny elektrojoniseringsaerosoliser. Vi nådde liknande träffprocent som med GDVN och de första resultaten visar inga tecken på rester från föroreningar även för små partiklar med en diameter på 35 nm. Denna förbättring kan innebära att ett genombrott för avbildning av mycket små partiklar är nära förestående, och kanske kan uppnås redan med de data som nyligen samlats på enskilda proteiner (Rubisco, 11 nm i diameter). Men för att nå atomär upplösning måste tekniska förutsättningar med avseende på fotonflöde och bakgrundsbrus förbättras vid XFEL. Lyckligtvis förväntas högre fotonflöden i och med öppnandet av "European XFEL" (EXFEL) [83]. Genom ett gemensamt initiativ från LCLS och användare har en grupp bildats, "the Single Particle Imaging Initiative", som arbetar med att förbättra villkoren för FXI vid LCLS genom att utnyttja den kombinerade kompentensen i gruppen [3].

För att testa, karaktärisera och förbättra injektordesign och drift i labortoriet behövs diagnosverktyg, som inte behöver tillgång till en XFEL. Vi har byggt upp en station för Mie diffraktionsavbildning (Mie scattering imaging, MSI), som gör det möjligt för oss att mäta position, storlek och hastighet hos enskilda injicerade partiklar under injicering utan användning av XFEL röntgenstrålning. Vid en XFEL kan MSI installeras som ett kraftfullt verktyg för att övervaka provinjektion under ett pågående FXI experiment.

XFEL experiment genererar stora mängder data på kort tid. Förberedelse, genomförande och analys av data från dessa experiment underlättas av specialiserad programvara. Vi har utvecklat mjukvara med öppen källkod som underlättar simulering (Artikel III), experiment-övervakning (Artikel IV), visualisering (*Owl* programvara), och förbearbetning (Artikel V) av röntgendiffraktionsdata.

Den höga pulsfrekvensen som förväntas för EXFEL [83] kommer att möjliggöra insamling av diffraktionsdata från ett stort antal enskilda partiklar på kort tid. Dessa stora mängder diffraktionsdata täcker konformationsrymden mycket tätt, vilket kan hjälpa oss att nå atomär upplösning även för mycket heterogena strukturer, förutsatt att de beräkningsmässiga problemen kan lösas. Vidare kan snabba datahastigheter möjliggöra studier av konformationsrymden för strukturella fluktuationer vid rumstemperatur eller som svar på reaktioner orsakade av en extern trigger (t ex en pump-puls).

Utnyttjande av aerosolinjektion för FXI ger möjlighet till flera lovande framtida tillämpningar. Ett utmärkande drag är att provet kan injiceras direkt i lösning och kan därmed lätt kombineras med många tekniker för provseparering och provblandning. Tekniker såsom vätskekromatografi eller kapillärelektrofores skulle kunna användas i kombination med in silico rening (Artikel I) för att lösa strukturer av inneboende heterogena molekylkomplex till hög upplösning. Mikrofluidiktekniker som möjliggör blandning av komponenter i definierade förhållanden kan tillämpas för att studera montering av virala kapsider, DNA-packning, proteinveckning, vesikelbildning med mera. Vidare kan det faktum att provlösningen finfördelas ("kompartmentaliseras") till små droppar användas för att fånga och avbilda transienta komplex. Indunstning av lösningsmedlet ökar gradvis den effektiva koncentrationen av ingående komponenter, och kan främja komplexbildning trots låg affinitet. Avdunstning fångar den komplexa strukturen och gör den möjlig att studera med FXI. Transienta interaktioner förekommer överallt i biologiska system och är utmanande att studera med existerande tekniker. XFEL-tekniken och de utvecklingar av databehandling och mjukvara som beskrivits i denna avhandling ger oss nya verktyg för sådana studier.

Acknowledgments

This work was carried out in an inspiring interdisciplinary team and benefited greatly from the support by colleagues and friends in Sweden and abroad. Thank you everyone!

In particular I thank: **Filipe Maia** and **Janos Hajdu** who have been such brilliant, supportive, and encouraging supervisors. Thank you for sharing your expertise and for being patient with me. From when I started in the group as an exchange student until now, **Tomas Ekeberg** has been a mentor, role model, and friend. Discussions with Tomas involved often iterative promenades to diffraction space and back. This has been educating, inspiring and always fun. Thank you! I am indebted to **Inger Andersson**, **Nicusor Timneanu**, and **Martin Svenda** who always had an open door when I needed help with various problems that were difficult for me and fortunately often not for them. I thank **Johan Bielecki** and **Daniel Westphal** without whom the great progress in injector development and characterisation would likely not have happened. Thank you also for valuable discussions with you in the hot tub. Thank you **Benedikt Daurer** for being my friend, sailing mate, wonderful colleague, and cave-food buddy. Also, many thanks to **Jonas Selleberg**, **Gijs van der Schot**, **Carl Nettelblad**, **Alberto Pietrini**, **Ida Lundholm**, **Jing Liu**, **Olof Jönsson**, and **Christofer Östlin** for inspiring discussions, bug reports, and exciting volleyball matches between the night shifts of beam times. Also the help of my colleagues **Jakob Andreasson**, **Bianca Iwan**, **Marvin Seibert**, and **Kerstin Mühlig** is gratefully acknowledged. Thank you **Laura Gunn**, **Dirk Hasse**, **Gunilla Carlsson**, **Anna Larsson**, **Anna Munke**, **Daniel Larsson**, **Kenta Okamoto**, **Hemanth Kumar**, **Karin Valegård**, and **Margareta Ingelman** for your biological samples, contageous enthusiasm for your biological research, and for patiently answering my frequent questions. Thank you **Duane Loh**, **Kartik Ayyer**, **Andrew Martin**, and **Anton Barty** for discussions during beam times and our weekly Skype meetings that broadened my horizons significantly. Also, I am much obliged to my collaborators and strong supporters **Arwen Pearsson** and **Klaus Gieweckemeyer** in Hamburg.

For indirect but vital contribution to this work I thank my parents **Angelika** and **Klaus**, my sisters **Leonie** and **Valerie**, and my friends **Preeti Baskaran**, **Maja Malmberg**, **Gustaf Boström**, **Friderike Beyer**, **Vincenz Reith**, **Elisabeth Miller**, **Cissi Lefverman**, **Johann Schmiedinger**, **Daniel Jones**, **Jonas Sottmann**, **Nicole Schaffer**, **Magdalena Muckenthaler**, **Katharina Rall**, **Jakob Funk**, **Igor Rocha**, **Ana Oliveira**, **Terese Bergfors**, **Katrin Beilharz**, **Showgy Maayeh**, **Alexis Boucharin**, **Prune Leroy**, and **Jennah Dharamshi**.

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, New York, USA, 4th edition, 2002.

[2] J. Als-Nielsen and D. McMorrow. *Elements of modern x-ray physics*. Wiley, New York, USA, 2001.

[3] A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. DePonte, P. Drell, S. Doniach, K. H. Downing, T. Earnest, H. Elmlund, V. Elser, M. G??hr, J. Hajdu, J. Hastings, S. P. Hau-Riege, Z. Huang, E. E. Lattman, F. R N C Maia, S. Marchesini, A. Ourmazd, C. Pellegrini, R. Santra, I. Schlichting, C. Schroer, J. C H Spence, I. A. Vartanyants, S. Wakatsuki, W. I. Weis, and G. J. Williams. The linac coherent light source single particle imaging road map. *Structural Dynamics*, 2(4), 2015.

[4] S. Awel, R. A. Kirian, N. Eckerskorn, M. Wiedorn, D. A. Horke, A. V Rode, and J. Küpper. Visualizing aerosol-particle injection for diffractive-imaging experiments. 24(6):6507–6521, 2016.

[5] X.-C. Bai, G. McMullan, and S. H. W. Scheres. How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences*, 40(1):49–57, 2015.

[6] R. Barakat and G. Newsam. Necessary conditions for a unique solution to two-dimensional phase recovery. *J. Math. Phys.*, 25(11):3190–3193, 1984.

[7] A. Barty, C. Caleman, A. Aquila, N. Timneanu, L. Lomb, T. a White, J. Andreasson, D. Arnlund, S. Bajt, T. R. M. Barends, M. Barthelmess, M. J. Bogan, C. Bostedt, J. D. Bozek, R. Coffee, N. Coppola, J. Davidsson, D. P. DePonte, R. B. Doak, T. Ekeberg, V. Elser, S. W. Epp, B. Erk, H. Fleckenstein, L. Foucar, P. Fromme, H. Graafsma, L. Gumprecht, J. Hajdu, C. Y. Hampton, R. Hartmann, A. Hartmann, G. Hauser, H. Hirsemann, P. Holl, M. S. Hunter, L. Johansson, S. Kassemeyer, N. Kimmel, R. a Kirian, M. Liang, F. R. N. C. Maia, E. Malmerberg, S. Marchesini, A. V. Martin, K. Nass, R. Neutze, C. Reich, D. Rolles, B. Rudek, A. Rudenko, H. Scott, I. Schlichting, J. Schulz, M. M. Seibert, R. L. Shoeman, R. G. Sierra, H. Soltau, J. C. H. Spence, F. Stellato, S. Stern, L. Strüder, J. Ullrich, X. Wang, G. Weidenspointner, U. Weierstall, C. B. Wunderer, and H. N. Chapman. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature photonics*, 6:35–40, 2012.

[8] J. L. P. Benesch, B. T. Ruotolo, D. A. Simmons, and C. V. Robinson. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chemical reviews*, 107(8):3544–3567, 2007.

[9] M. Bergh, G. Huldt, N. Timneanu, F. R. N. C. Maia, and J. Hajdu. Feasibility of imaging living cells at subnanometer resolutions by ultrafast X-ray diffraction. *Quarterly Reviews of Biophysics*, 41(3-4):181–204, 2008.

[10] J. D. Bernal, I. Fankuchen, and M. F. Perutz. An X-ray Study on Chymotrypsin and Haemoglobin. *Nature*, 141(3568):523–524, 1938.

[11] M. J. Bogan, W. H. Benner, S. Boutet, U. Rohner, M. Frank, A. Barty, M. M. Seibert, F. Maia, S. Marchesini, S. Bajt, B. Woods, V. Riot, S. P. Hau-Riege, M. Svenda, E. Marklund, E. Spiller, J. Hajdu, and H. N. Chapman. Single particle X-ray diffractive imaging. *Nano letters*, 8(1):310–6, 2008.

[12] M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, Cambridge, England, 7th edition, 1999.

[13] C. Bostedt, J. D. Bozek, P. H. Bucksbaum, R. N. Coffee, J. B. Hastings, Z. Huang, R. W. Lee, S. Schorb, J. N. Corlett, P. Denes, P. Emma, R. W. Falcone, R. W. Schoenlein, G. Doumy, E. P. Kanter, B. Kraessig, S. Southworth, L. Young, L. Fang, M. Hoener, N. Berrah, C. Roedig, and L. F. DiMauro. Ultra-fast and ultra-intense x-ray sciences: first results from the Linac Coherent Light Source free-electron laser. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 46:164003, 2013.

[14] J. D. Bozek. AMO instrumentation for the LCLS X-ray FEL. *The European Physical Journal Special Topics*, 169(1):129–132, 2009.

[15] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. Bajt, H. Benner, R. A. London, E. Ploenjes, M. Kuhlmann, R. Treusch, S. Duesterer, T. Tschentscher, J. R. Schneider, E. Spiller, T. Moeller, C. Bostedt, M. Hoener, D. A. Shapiro, K. O. Hodgson, D. der Spoel, F. Burmeister, M. Bergh, C. Caleman, G. Huldt, M. M. Seibert, F. R. N. C. Maia, R. W. Lee, A. Szoeke, N. Timneanu, and J. Hajdu. Femtosecond diffractive imaging with a soft-x-ray free-electron laser. *Nature Physics*, 2(12):839–843, 2006.

[16] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro. High-resolution ab initio three-dimensional x-ray diffraction microscopy. *Journal of the Optical Society of America A*, 23(5):1179–1200, 2006.

[17] H. N. Chapman, C. Caleman, and N. Timneanu. Diffraction before destruction. *Philosophical Transactions of the Royal Society B*, page 20130313, 2014.

[18] H. N. Chapman, P. Fromme, A. Barty, T. White, R. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M. J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Hömke, C. Reich, D. Pietschner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, M. Messerschmidt, J. D. Bozek, S. P. Hau-Riege, M. Frank, C. Y. Hampton, R. G. Sierra, D. Starodub, G. J. Williams, J. Hajdu, N. Timneanu, M. M. Seibert, J. Andreasson, A. Rocker, O. Jönsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, K. E Schmidt, X. Wang, I. Grotjohann, J. M. Holton, T. R. M. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J. C. H. Spence. Femtosecond X-ray protein nanocrystallography. *Nature*, 470(7332):73–7, 2011.

[19] H. N. Chapman and K. A. Nugent. Coherent lensless x-ray imaging. *Nature Photonics*, 4(12):833–839, 2010.

[20] R. B. Cole. *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation, and Applications*. Wiley, New York, USA, 1997.

[21] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301, 1965.

[22] NVIDIA Corporation. Nvidia cuda fast fourier transform library (cufft). https://developer.nvidia.com/cufft, 2016.

[23] B. Daurer, K. Okamoto, et al. Considerations for imaging small virus particles from a test experiment at the lcls. *in preparation*, 2017.

[24] D. P. DePonte, U. Weierstall, K. Schmidt, J. Warner, D. Starodub, J. C. H. Spence, and R. B. Doak. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics*, 41(19):195505, 2008.

[25] T. Ekeberg, M. Svenda, C. Abergel, F. R. N. C. Maia, V. Seltzer, J. M. Claverie, M. Hantke, O. Jonsson, C. Nettelblad, G. Van Der Schot, M. Liang, D. P. DePonte, A. Barty, M. M. Seibert, B. Iwan, I. Andersson, N. D. Loh, A. V. Martin, H. Chapman, C. Bostedt, J. D. Bozek, K. R. Ferguson, J. Krzywinski, S. W. Epp, D. Rolles, A. Rudenko, R. Hartmann, N. Kimmel, and J. Hajdu. Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser. *Physical Review Letters*, 114(9):1–6, 2015.

[26] V. Elser. Phase Retrieval by Iterated Projections. *Journal of the Optical Society of America A*, 20(1):40–55, 2003.

[27] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, a. Brachmann, P. Bucksbaum, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, a. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, a. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda. First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photonics*, 4(9):641–647, 2010.

[28] G. S. Espie and M. S. Kimber. Carboxysomes : cyanobacterial RubisCO comes in small packages. *Photosynthesis Research*, 109(1):7–20, 2011.

[29] H. M. L. Faulkner and J. M. Rodenburg. Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm. *Physical Review Letters*, 93(2):023903–1, 2004.

[30] J. R. Fienup. Reconstruction of an object from the modulus. *Optics Letters*, 3(1):27–29, 1978.

[31] J. R. Fienup. Phase Retrieval Algorithms: A Comparison. *Applied Optics*, 21(15):2758–2769, 1982.

[32] J. R. Fienup. Phase retrieval algorithms: a personal tour [Invited]. *Applied Optics*, 52(1):45–56, 2012.

[33] L. Galli, S. K. Son, T. R. M. Barends, T. A. White, A. Barty, S. Botha, S. Boutet, C. Caleman, R. B. Doak, M. H. Nanao, K. Nass, R. L. Shoeman, N. Timneanu, R. Santra, I. Schlichting, and H. N. Chapman. Towards phasing using high X-ray intensity. *IUCrJ*, 2:627–634, 2015.

[34] A. Ganan-Calvo. Generation of Steady Liquid Microthreads and Micron-Sized Monodisperse Sprays in Gas Streams. *Physical Review Letters*, 80(2):285–288, 1998.

[35] A. M. Ganan-Calvo and J. M. Montanero. Revision of capillary cone-jet physics: Electrospray and flow focusing. *Physical Review E*, 79(6):1–18, 2009.

[36] R. W. Gerchberg and W. O. Saxton. A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures. *Optik*, 35(2):237–246, 1972.

[37] J. M. Guss, E. A. Merritt, R. P. Phizackerley, B. Hedman, M. Murata, K. O. Hodgson, and H. C. Freeman. Phase determination by multiple-wavelength x-ray diffraction: crystal structure of a basic ”blue” copper protein from cucumbers. *Science*, 241(4867):806–811, 1988.

[38] S. P. Hau-Riege, R. A. London, G. Huldt, and H. N. Chapman. Pulse requirements for x-ray diffraction imaging of single biological molecules. *Physical Review E*, 71:061919, 2005.

[39] S. P. Hau-Riege, R. A. London, and A. Szoke. Dynamics of biological molecules irradiated by short x-ray pulses. *Physical Review E*, 69(5 Pt 1):051906, 2004.

[40] R. Henderson. Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction. *Proceedings of the Royal Society B: Biological Sciences*, 241(1300):6–8, 1990.

[41] W. A. Hendrickson and M. M. Teeter. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature*, 290(5802):107–113, 1981.

[42] B. L. Henke, E.M. Gullikson, and J.C. Davis. X-ray interactions: photoabsorption, scattering, transmission, and reflection at e=50-30000 ev, z=1-92. *Atomic Data and Nuclear Data Tables*, 54(2):181–342, 1993.

[43] K. Henzler-Wildman and D. Kern. Dynamic personalities of proteins. *Nature*, 450:7–9, 2007.

[44] W. C. Hinds. *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*. Wiley, New York, USA, 2 edition, 2006.

[45] R. Holyst, M. Litniewski, D. Jakubczyk, K. Kolwas, M. Kolwas, M. Kowalski, S. Migacz, S. Palesa, and M. Zientara. Evaporation of freely suspended single droplets : experimental , theoretical and. *Reports on Progress in Physics*, 76:034601, 2013.

[46] C. V. Iancu, H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen. The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. *Journal of molecular biology*, 372(3):764–73, 2007.

[47] ISO. Particle size analysis – Laser diffraction methods. Standard, International Organization for Standardization, Geneva, CH, October 2009.

[48] D. Jakubczyk, M. Kolwas, G. Derkachov, K. Kolwas, M. Zientara, and D Jakubczyk. Evaporation of Micro-Droplets : the Radius-Square-Law Revisited. *Acta Physica Polonica A*, 122(4), 2012.

[49] M. F. Jarrold. Unfolding, refolding, and hydration of proteins in the gas phase. *Accounts of Chemical Research*, 32(4):360–367, 1999.

[50] R. Juraschek, T. Du, and M. Karas. Nanoelectrospray — More Than Just a Minimized-Flow Electrospray Ionization Source. *Journal of the Americal Society of Mass Spectrometry*, 10:300–308, 1999.

[51] S. Kassemeyer, J. Steinbrener, L. Lomb, E. Hartmann, A. Aquila, A. Barty, A. V. Martin, C. Y. Hampton, S. Bajt, M. Barthelmess, T. R. M. Barends, C. Bostedt, M. Bott, D. Bozek, N. Coppola, M. Cryle, D. P. DePonte, R. B. Doak, W. Epp, B. Erk, H. Fleckenstein, L. Foucar, H. Graafsma, L. Gumprecht, A. Hartmann, R. Hartmann, G. Hauser, A. Hömke, P. Holl, O. Jönsson, N. Kimmel, M. Liang, F. R. N. C. Maia, S. Marchesini, K. Nass, C. Reich, D. Rolles, B. Rudek, A. Rudenko, J. Schulz, R. L. Shoeman, R. G. Sierra, H. Soltau, J. C. H. Spence, D. Starodub, F. Stellato, S. Stern, M. Svenda, G. Weidenspointner, U. Weierstall, A. Thomas, C. Wunderer, M. Frank, H. N. Chapman, L. Strüder, M. J. Bogan, and I. Schlichting. Femtosecond free-electron laser x-ray diffraction data sets for algorithm development. *Optics Express*, 20(4):4149–4158, 2012.

[52] L. E. Kay. NMR studies of protein structure and dynamics. *Advances in Magnetic Resonance*, 173:193–207, 2005.

[53] J. Kern, R. Alonso-Mori, R. Tran, J. Hattne, R. J. Gildea, N. Echols, C. Glöckner, J. Hellmich, H. Laksmono, R. G. Sierra, B. Lassalle-Kaiser, S. Koroidov, A. Lampe, G. Han, S. Gul, D. Difiore, D. Milathianaki, A. R. Fry, A. Miahnahri, D. W. Schafer, M. Messerschmidt, M. M. Seibert, J. E. Koglin, D. Sokaras, T.-C. Weng, J. Sellberg, M. J. Latimer, R. W. Grosse-Kunstleve, P. H. Zwart, W. E. White, P. Glatzel, P. D. Adams, M. J. Bogan, G. J. Williams, S. Boutet, J. Messinger, A. Zouni, N. K. Sauter, V. K. Yachandra, U. Bergmann, and J. Yano. Simultaneous femtosecond x-ray spectroscopy and diffraction of photosystem ii at room temperature. *Science*, 340(6131):491–5, 2013.

[54] T. Kimura, Y. Joti, A. Shibuya, C. Song, S. Kim, K. Tono, M. Yabashi, M. Tamakoshi, T. Moriya, T. Oshima, T. Ishikawa, Y. Bessho, and Y. Nishino. Imaging live cell in micro-liquid enclosure by X-ray laser diffraction. *Nature communications*, 5:3052, 2014.

[55] C. Kupitz, S. Basu, I. Grotjohann, R. Fromme, N. A. Zatsepin, K. N. Rendek, M. S. Hunter, R. L. Shoeman, T. A. White, D. Wang, D. James, J.-H. Yang, D. E. Cobb, B. Reeder, R. G. Sierra, H. Liu, A. Barty, A. L. Aquila, D. Deponte, R. A. Kirian, S. Bari, J. J. Bergkamp, K. R. Beyerlein, M. J. Bogan, C. Caleman, T.-C. Chao, C. E. Conrad, K. M. Davis, H. Fleckenstein, L. Galli, S. P. Hau-Riege, S. Kassemeyer, H. Laksmono, M. Liang, L. Lomb, S. Marchesini, A. V. Martin, M. Messerschmidt, D. Milathianaki, K. Nass, A. Ros, S. Roy-Chowdhury, K. Schmidt, M. Seibert, J. Steinbrener, F. Stellato, L. Yan, C. Yoon, T. A. Moore, A. L. Moore, Y. Pushkar, G. J. Williams, S. Boutet, R. B. Doak, U. Weierstall, M. Frank, H. N. Chapman, J. C. H. Spence, and P. Fromme. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature*, 513(7517):261–265, 2014.

[56] L. A. Lane, B. T. Ruotolo, C. V. Robinson, G. Favrin, and J. L. P. Benesch. A Monte Carlo approach for assessing the specificity of protein oligomers observed in nano-electrospray mass spectra. *International Journal of Mass Spectrometry*, 283(1-3):169–177, 2009.

[57] H. Liu. *Fundamental Phenomena and Principles in Droplet Processes*. Noyes Publications, New York, USA, 2000.

[58] P. Liu, P. J. Ziemann, D. B. Kittelson, and P. H. McMurry. Generating Particle Beams of Controlled Dimensions and Divergence: II. Experimental Evaluation of Particle Motion in Aerodynamic Lenses and Nozzle Expansions. *Aerosol Science and Technology*, 22(3):314–324, 1995.

[59] W. Liu, D. Wacker, C. Gati, G. W. Han, D. James, D. Wang, G. Nelson, U. Weierstall, V. Katritch, A. Barty, N. A Zatsepin, D. Li, M. Messerschmidt, S. Boutet, G. J. Williams, J. E. Koglin, M. M. Seibert, C. Wang, S. T. A. Shah, S. Basu, R. Fromme, C. Kupitz, K. N Rendek, I. Grotjohann, P. Fromme, R. A. Kirian, K. R. Beyerlein, T. A. White, H. N. Chapman, M. Caffrey, J. C. H. Spence, R. C. Stevens, and V. Cherezov. Serial Femtosecond Crystallography of G Protein-Coupled Receptors. *Science*, 342:1521–1525, 2013.

[60] N.-T. D. Loh and V. Elser. Reconstruction algorithm for single-particle diffraction imaging experiments. *Physical Review E*, 80(2):1–20, 2009.

[61] D. R. Luke. Relaxed Averaged Alternating Reflections for Diffraction Imaging. *Inverse Problems*, 21:37–50, 2005.

[62] M. M. Seibert, S. Boutet, M. Svenda, T. Ekeberg, F. R. N. C. Maia, M. J. Bogan, N. Tîmneanu, A. Barty, S. Hau-Riege, C. Caleman, M. Frank, H. Benner, J. Y. Lee, S. Marchesini, J. W. Shaevitz, D. A. Fletcher, S. Bajt, I. Andersson, H. N. Chapman, and J. Hajdu. Femtosecond diffractive imaging of biological cells. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 43:194015, 2010.

[63] John M J Madey. Stimulated emission of bremsstrahlung in a periodic magnetic field. *Journal of Applied Physics*, 42(5):1906–1913, 1971.

[64] F. R. N. C. Maia. The coherent x-ray imaging data bank. *Nature Methods*, 9(9):854–855, 2012.

[65] F. R. N. C. Maia, T. Ekeberg, D. Van Der Spoel, and J. Hajdu. Hawk: The image reconstruction package for coherent X-ray diffractive imaging. *Journal of Applied Crystallography*, 43(6):1535–1539, 2010.

[66] F. R. N. C. Maia and J. Hajdu. Comment : The trickle before the torrent — diffraction data from X-ray lasers. *Scientific Data*, 3:1–3, 2016.

[67] F. R. N. C. Maia, T. A. White, N. D. Loh, and J. Hajdu. editorial CCP-FEL : a collection of computer programs for free-electron laser research. *Journal of Applied Crystallography*, 49:1117–1120, 2016.

[68] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence. X-ray image reconstruction from a diffraction pattern alone. *Physical Review B*, 68(14), 2003.

[69] A. V. Martin, J. K. Corso, C. Caleman, N. Timneanu, and H. M. Quiney. Single-molecule imaging with longer X-ray laser pulses. *IUCrJ*, 2:661–674, 2015.

[70] P. H. Mcmurry. A review of atmospheric aerosol measurements. *Atmospheric Environment*, 34:1959–1999, 2000.

[71] I. McNulty, J. Kirz, C. Jacobsen, E. H. Anderson, M. R. Howells, and D. P. Kern. High-Resolution Imaging by Fourier Transform X-ray Holography. *Science*, 256(5059):1009–1012, 1992.

[72] J. Miao, D. Sayre, and H. N. Chapman. Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. *Journal of the Optical Society of America A*, 15(6):1662, 1998.

[73] A. Munke, J. Andreasson, A. Aquila, S. Awel, K. Ayyer, A. Barty, R. J. Bean, P. Berntsen, J. Bielecki, S. Boutet, M. Bucher, H. N. Chapman, B. J. Daurer, H. DeMirci, V. Elser, P. Fromme, J. Hajdu, M. F. Hantke, A. Higashiura, B. G. Hogue, A. Hosseinizadeh, Y. Kim, R. A. Kirian, H. K. N. Reddy, T.-Y. Lan, D. S. D. Larsson, H. Li, N. D. Loh, F. R. N. C. Maia, A. P. Mancuso, K. Mühlig, A. Nakagawa, D. Nam, G. Nelson, C. Nettelblad, K. Okamoto, A. Ourmazd, M. Rose, G. van der Schot, P. Schwander, M. M. Seibert, J. A. Sellberg, R. G. Sierra, C. Song, M. Svenda, N. Timneanu, I. A. Vartanyants, D. Westphal, M. O. Wiedorn, G. J. Williams, P. L. Xavier, C. H. Yoon, and J. Zook. Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source. *Scientific Data*, 3:160064, 2016.

[74] W. K. Murphy and G. W. Sears. Production of Particulate Beams. *Journal of Applied Physics*, 35(6):1986, 1964.

[75] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu. Potential for biomolecular imaging with femtosecond x-ray pulses. *Nature*, 406(6797):752–757, 2000.

[76] R. L. Owen, D. Axford, J. E. Nettleship, R. J. Owens, J. I. Robinson, A. W. Morgan, A. S. Dore, G. Lebon, C. G. Tate, E. E. Fry, J. Ren, D. I. Stuart, and G. Evans. Outrunning free radicals in room-temperature macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 68(7):810–818, 2012.

[77] R. L. Owen, E. Rudino-Pinera, and E. F. Garman. Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc. Natl Acad. Sci. USA*, 103:4912–4917, 2006.

[78] D. M. Paganin. *Coherent X-ray Optics*. Oxford University Press, Oxford, England, 2006.

[79] P. Persson, S. Lunell, A. Szöke, B. Ziaja, and J. Hajdu. Shake-up and shake-off excitations with associated electron losses in X-ray studies of proteins. *Protein Science*, 10:2480–2484, 2001.

[80] B. D. Rae, B. M. Long, M. R. Badger, and G. D. Price. Functions, Compositions, and Evolution of the Two Types of Carboxysomes: Polyhedral Microcompartments That Facilitate CO2 Fixation in Cyanobacteria and Some Proteobacteria. *Microbiology and Molecular Biology Reviews*, 77(3):357–379, 2013.

[81] M. G. Rossmann. *The Molecular Replacement Method*. Gordon & Breach, New York, USA, 1972.

[82] D. Sayre. Some implications of a theorem due to Shannon. *Acta Crystallographica*, 5(6):843–843, 1952.

[83] E. A. Schneidmiller and M.V. Yurkov. Photon beam properties at the European XFEL. Technical Report XFEL.EU TR-2011-006, European X-Ray Free-Electron Laser Facility GmbH, 2011.

[84] M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, J. Joensson, D. Odic, B. Iwan, A. Rocker, D. Westphal, M. Hantke, D. P. DePonte, A. Barty, J. Schulz, L. Gumprecht, N. Coppola, A. Aquila, M. Liang,

T. A. White, A. Martin, C. Caleman, S. Stern, C. Abergel, V. Seltzer, J.-M. Claverie, C. Bostedt, J. D. Bozek, S. Boutet, A. A. Miahnahri, M. Messerschmidt, J. Krzywinski, G. Williams, K. O. Hodgson, M. J. Bogan, C. Y. Hampton, R. G. Sierra, D. Starodub, I. Andersson, S. Bajt, M. Barthelmess, J. C. H. Spence, P. Fromme, U. Weierstall, R. Kirian, M. Hunter, R. B. Doak, S. Marchesini, S. P. Hau-Riege, M. Frank, R. L. Shoeman, L. Lomb, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, C. Schmidt, L. Foucar, N. Kimmel, P. Holl, B. Rudek, B. Erk, A. Hoemke, C. Reich, D. Pietschner, G. Weidenspointner, L. Strueder, G. Hauser, H. Gorke, J. Ullrich, I. Schlichting, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kuehnel, R. Andritschke, C.-D. Schroeter, F. Krasniqi, M. Bott, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, H. N Chapman, and J. Hajdu. Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature*, 470(7332):78–U86, 2011.

[85] C. E. Shannon. Communication in the presence of noise. In *Proceedings of the Institute of Radio Engineers (IRE)*, volume 37, pages 10–21, 1949.

[86] J. M. Shively, F. Ball, D. H. Brown, and R. E. Saunders. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. *Science*, 182(4112):584–6, 1973.

[87] J. Solem. Imaging biological specimens with high-intensity soft x rays. *Journal of the Optical Society of America B*, 3(11):1551–1565, 1986.

[88] S. K. Son, H. N. Chapman, and R. Santra. Multiwavelength anomalous diffraction at high X-ray intensity. *Physical Review Letters*, 107(21):1–5, 2011.

[89] L. Strüder, S. Epp, D. Rolles, R. Hartmann, P. Holl, G. Lutz, H. Soltau, R. Eckart, C. Reich, K. Heinzinger, C. Thamm, A. Rudenko, F. Krasniqi, K.-U. Kühnel, C. Bauer, C.-D. Schröter, R. Moshammer, S. Techert, D. Miessner, M. Porro, O. Hälker, N. Meidinger, N. Kimmel, R. Andritschke, F. Schopper, G. Weidenspointner, A. Ziegler, D. Pietschner, S. Herrmann, U. Pietsch, A. Walenta, W. Leitenberger, C. Bostedt, T. Möller, D. Rupp, M. Adolph, H. Graafsma, H. Hirsemann, K. Gärtner, R. Richter, L. Foucar, R. L. Shoeman, I. Schlichting, and J. Ullrich. Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 614(3):483–496, 2010.

[90] M. Suga, F. Akita, K. Hirata, G. Ueno, H. Murakami, Y. Nakajima, T. Shimizu, K. Yamashita, M. Yamamoto, H. Ago, and J.-R. Shen. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature*, 517(7532):99–103, 2015.

[91] G. Taylor. The phase problem. *Acta Crystallographica - Section D Biological Crystallography*, 59(11):1881–1890, 2003.

[92] G. F. R. S Taylor. Disintegration of water drops in an electric field. *Proceedings of the Royal Society A - Mathematical, Physical and Engeneering Sciences*, 280(1382):384–397, 1964.

[93] The HDF Group. Hierarchical data format, version 5, 1997-2016. http://www.hdfgroup.org/HDF5/.

[94] P. Thibault, V. Elser, C. Jacobsen, D. Shapiro, and D. Sayre. Reconstruction of a yeast cell from X-ray diffraction data. *Acta Crystallographica Section A: Foundations of Crystallography*, 62(4):248–261, 2006.

[95] N. Timneanu, C. Caleman, J. Hajdu, and D. Van Der Spoel. Auger electron cascades in water and ice. *Chemical Physics*, 299(2-3):277–283, 2004.

[96] D. van der Spoel, E. G. Marklund, D. S. D. Larsson, and C. Caleman. Proteins, lipids, and water in the gas phase. *Macromolecular bioscience*, 11(1):50–9, 2011.

[97] Marin Van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250–262, 2005.

[98] U. Weierstall, R. B. Doak, J. C. H. Spence, D. Starodub, D. Shapiro, P. Kennedy, J. Warner, G. G. Hembree, P. Fromme, and H. N. Chapman. Droplet streams for serial crystallography of proteins. *Experiments in Fluids*, 44:675–689, 2008.

[99] G. Williams, H. Quiney, B. Dhal, C. Tran, K. Nugent, A. Peele, D. Paterson, and M. de Jonge. Fresnel Coherent Diffractive Imaging. *Physical Review Letters*, 97(2):1–4, 2006.

[100] M. S. Wilm and M. Mann. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes*, 136(2-3):167–180, 1994.

[101] L. Young, E. P. Kanter, B. Krässig, Y. Li, A. M. March, S. T. Pratt, R. Santra, S. H. Southworth, N. Rohringer, L. F. DiMauro, G. Doumy, C. A. Roedig, N. Berrah, L. Fang, M. Hoener, P. H. Bucksbaum, J. P. Cryan, S. Ghimire, J. M. Glownia, D. A. Reis, J. D. Bozek, C. Bostedt, and M. Messerschmidt. Femtosecond electronic response of atoms to ultra-intense X-rays. *Nature*, 466(7302):56–61, 2010.

[102] F. Zarrln, S. L Kaufman, and J. R. Socha. Droplet Size Measurements of various nebulisers using differential electrical mobility particle sizer. *Journal of Aerosol Science*, 22:S343–S346, 1991.

[103] B. Ziaja, D. Van Der Spoel, A. Szoeke, and J. Hajdu. Auger-electron cascades in diamond and amorphous carbon. *Physical Review B*, 64(21):18, 2001.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1451

Editor: The Dean of the Faculty of Science and Technology

Distribution: publications.uu.se
urn:nbn:se:uu:diva-306609

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2016