Using Coherent Diffraction Imaging to Obtain The Structure of a Metaphase Chromosome

Laura Shemilt

April 16, 2012

Chapter 1

Introduction

1.1 X-rays and their uses

Humans have always desired to see further, and have built instruments to do so. We can now see distances both farther and smaller than we can imagine. With each observed scale new details are found, old details are lost and a new world of discovery awaits. Being able to see allows us to study phenomena and make new developments in science and technology on the observed scale. The first microscopes provided a gateway for huge advancement in medicine and biology and the first telescopes allowed the positions of the planets and stars to be tracked. In more recent years a similar explosion of activity has occurred in the field of nanotechnology which has only been possible due to the invention of techniques able to image the nanoscale.

In 1895 the discovery of x-rays by Röntgen allowed the body to imaged without using invasive techniques and since then x-ray imaging has been applied to many scientific and medical problems. X-rays have been brought into use in clinical medicine and have become an integral part of diagnosing patients. With the discovery of Bragg's Law the field of crystallography has allowed structures of macromolecules and proteins to be obtained. The Protein Data Bank grows each year with resolved structures almost entirely due to crystallography.

New powerful sources of X-rays, synchrotrons, have lead to the development of x-ray imaging techniques such as x-ray microscopes, tomography, topography and diffraction imaging as well as a whole host of techniques to tackle other problems in science. The latest machines to be developed are the X-ray Free Electron Lasers (X-FEL) which could help to solve the structures of large biological macromolecules [1].

The constant development of x-ray sources since their discovery has yielded some incredible results, and will hopefully continue to solve problems in science. This report looks at the area of X-Ray imaging with the main focus on the technique of Coherent Diffraction Imaging. The report describes the recent work of that has been done in this field especially in its applications to biology and how these can be built upon to address current problems. The

report will discuss the possibility of using this technique to image the metaphase human chromosome and analysis these images to answer some of the questions about chromosome structure.

1.2 Problems in Structural Genetics

When Gregor Mendel first did his experiments with pea plants he realised that traits were passed down generations. Twenty years after this research, with the advent of better microscopes, chromosomes were first found. It was concluded that chromosomes contained genes that were passed between generations. It was Watson and Crick who first unravelled the structure of the DNA helix and found the A-T, C-G base parings. It was realised that DNA contained the entire information on the organism coded in these base pairs. The DNA and RNA are responsible for making proteins and this process is known as the central dogma of genetics. Using only the code embedded in the DNA the genes direct the synthesis of proteins necessary for life. When a gene goes through mutation it can no longer make the desired proteins. Fragments of DNA can be used to isolate specific genes learn about their organization, expression, nucleotide sequence and evolution, this process is called genomics. The Human Genome Project was established in the 1990 in order to decode the genetic information of human beings. The entire gene sequence has not been established and this project is still ongoing. Despite the progress of unravelling the genetic sequence little is known about how the genes are structured in the chromosome. It is thought that microscopy techniques could help probe the 3D structure of the chromosome.

Chapter 2

An Introduction to Coherent Diffraction Imaging

Coherent Diffraction Imaging (CDI) is a lensless imaging technique that uses synchrotron sources and computerised algorithms to retrieve an image. Synchrotrons are used for their high brilliance, a term used in x-ray science to describe the quality of the source. CDI is a particularly powerful imaging technique because it is can yield high resolution images of matter on the nanoscale due to the avoidance of resolution limiting optics. This chapter will look at the considerations behind CDI, what is needed in this technique to resolve an image from its diffraction pattern and some of the important theoretical considerations such as uniqueness of solution and sufficient signal. At the end of this chapter the applications of CDI in biology will be discussed especially the advantages and limiting factors when using CDI in this field.

2.1 Exploring the Phase Problem

Coherent Diffraction Imaging is a lensless based imaging technique that uses algorithms to reconstruct the image of an object from its diffraction pattern. In order to do this information about the phase and amplitude of the diffraction pattern must be known. However, in measuring the diffraction pattern at the detector the phase information is lost and the image cannot be retrieved from the diffraction pattern. This is known as the phase problem and is explained in detail below.

The idea of the phase problem is well established in optics, where it is known that the an image can be obtained from the amplitude and phase information of an exiting wave. In the Fraunhofer (far-field) regime the amplitude of a wave exiting from a sample can be accurately described by the Fourier Transform of the transmission function of the object $\rho(\boldsymbol{r})$:

$$F(\mathbf{q}) = \int \rho(\mathbf{r}) \exp(i\mathbf{q}.\mathbf{r}) d^2\mathbf{r} \tag{2.1}$$

where $\mathbf{q}$ is the scattering vector and $\boldsymbol{r}$ is the real space vector. In the case of X-rays the transmission function, $\rho(\boldsymbol{r})$, is the complex electron density of the sample. Hence, the image reconstructed from this signal gives information about the electron density of the scattering object. From equation 2.1 it is evident that the transmission function can be recovered from the diffraction pattern by taking the inverse Fourier Transform. However, when a diffraction pattern is captured by a detector only the intensity is measured. The intensity $I(\boldsymbol{q})$ is the square of the Fourier Modulus:

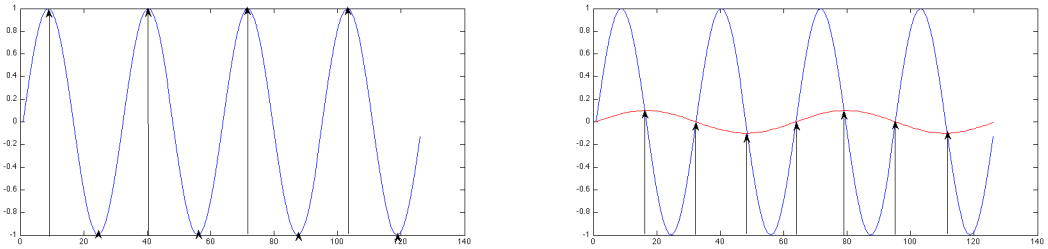
$$I(\boldsymbol{q}) = F(\boldsymbol{q})F^*(\boldsymbol{q}) = |F(\boldsymbol{q})|^2 \tag{2.2}$$

In measuring this intensity all the phase information about the sample is lost therefore $\rho(\boldsymbol{r})$ cannot be simply found by taking the inverse Fourier transform. The phase information, however, is still connected to this measured diffraction pattern and an image can be obtained from it. In an optical system a lens can be used to create an image, in this way the lens acts to preserve the phase information. In the case of x-rays, lenses and other focussing optics limit resolution and are very difficult to manufacture. Coherent Diffraction Imaging is advantageous in that it avoids the use of lenses by implementing iterative algorithms to retrieve the lost phase information, therefore allowing an image to be reconstructed from the diffraction pattern.

With all inverse problems there is the question of uniqueness of the solution, for the CDI method to work it must give one solution to the phase problem it is trying to solve. In the case of CDI the algorithms must provide a unique image from a given diffraction pattern. A unique solution is defined as a signal that can be completely and uniquely determined by the inverse of its Fourier Transform. It was shown by Bates [2] that for 2-D and 3-D problems there is almost always a unique solution to the phase problem. Non-uniqueness in problems greater than 1-D is very rare, however there are other solutions produced by the algorithms due to the properties of the Fourier transform. These are the reflection and the image propagated in space. This occurs due to the following being indistinguishable because of the lost phase information:$f(\boldsymbol{r})$, $f(\boldsymbol{r} + \boldsymbol{r}_0)$ and $f^*(-(\boldsymbol{r} + \boldsymbol{r}_0))$. For a unique solution to be found from the algorithms the data has to be suitably oversampled.

2.2 Oversampling

In information theory the Nyquist theory describes the minimum number of sampling points needed in order for a signal to be reconstructed from the sample. The Nyquist theory states that the minimum number of sample points needs to be taken at twice the frequency of the highest frequency present in a signal, $f_{sampling} \geq 2f_{max}$. Fig 2.1.(a) shows a signal sampled at the Nyquist limit it can be seen that the signal can be completely reconstructed from the



(a) (b)

Figure 2.1: (a) a signal sampled at the Nyquist density shows the minimum number of sample points needed in order to fully reconstruct the signal (b) a signal sampled below the Nyquist density cannot be reconstructed from the sample points to form the original signal but instead traces out a signal of lower frequency. This behaviour is known as aliasing.

sample points. If the signal is not sufficiently sampled the high frequencies get mixed with the lower frequencies, this phenomena is called aliasing. In fig. 2.1(b) the wave is sampled at a frequency below the Nyquist limit. When the signal is not sufficiently sampled a sine wave of a lower frequency and amplitude than the original is reconstructed from the sample points.

A crystal lattice is defined in frequency space by its reciprocal lattice points, occurring at $1/a$ where a is the lattice parameter, hence this is equivalent to sampling by the Nyquist theory. This level of sampling, however, only gives information on the amplitude but not the phase which is needed to solve the phase problem. It was noticed by Sayre in his short communication [3] that if a crystal were to be sampled at half the spacing of the unit cell $1/2a$, it would be possible to retrieve both the amplitude and the phase and hence solve the phase problem. He termed this sampling at twice the Nyquist frequency as oversampling.

So far we have considered continuous exit waves and Fourier transforms but in practical situations we have discrete datasets. A diffraction pattern measured by the detector is made up of pixels and therefore is a discrete dataset, this is the sampling of the diffraction pattern. The sampled amplitude of a diffraction pattern is given by the discrete Fourier transform:

$$|F(\boldsymbol{q})| = |\sum_{x=0}^{N} \rho(\boldsymbol{x}) \exp^{i\boldsymbol{q}.\boldsymbol{x}/N}| \tag{2.3}$$

As discussed above this sampling needs to meet the oversampling criteria. The oversampling condition proposed by Miao et al. is used as the guiding limit of oversampling for 3D diffraction pattern datasets that can be inverted using the phasing algorithms [4]. This argument is based on equation counting, Miao argues that the amplitude of a discretely measured diffraction pattern given by eq. (3) is really a set of equations to be solved. If this magnitude is considered to be real and Friedel's Law [1]is taken into account then the number of equations for a 1-D problem is

[1]Friedel's Law states that all diffraction patterns are symmetrical

$N/2$, for a 2-D problem $N^2/2$ and for a 3D problem is $N^3/2$ but the number of unknowns is N, N^2 and N^3 as the value at each pixel has a real and imaginary part. The number of equations is half the number of unknowns leaving the problem underdetermined by a factor of two. This can be solved by the oversampling method. In the phasing algorithms described in section 2.3 a guessed electron density (or support) is used to start the first iteration. If the dataset is oversampled finely enough a finite support can be used with a the area outside the support being zero. This support is a set of known valued pixels in a space of unknowns. In order for the phase problem to be solved the the ratio proposed by Miao et al. is:

$$\sigma = \frac{\text{total pixel number}}{\text{no. pixels with unknown value}} \tag{2.4}$$

From the description above the phase problem is undetermined by a factor of two, therefore $\sigma > 2$ in order for it to be solved. Extending this to two- and three-dimensions σ must equal, $2^{1/2}$, $2^{1/3}$ respectively in each dimension. Miao showed successful, unique reconstructions from 2-D simulated data for $\sigma > 2.57$. In a recent paper by Song a deconvolution method was used to oversample exactly simulated data [5]. It was shown by oversampling this data improved the quality of the image however this as not yet been applied to real data.

2.3 Algorithms

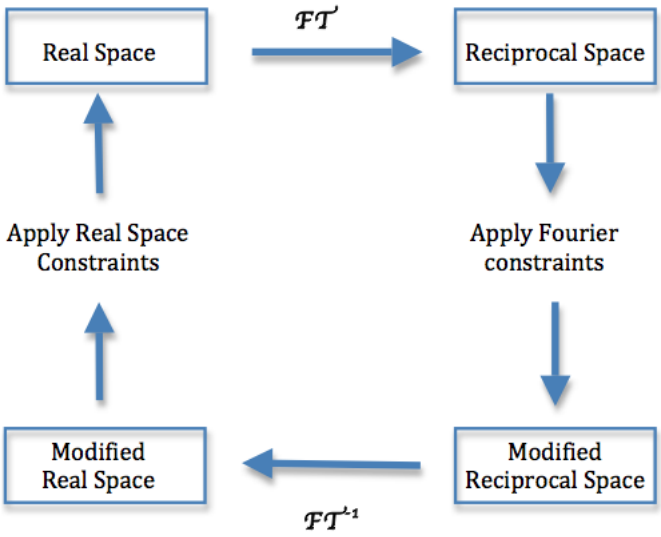


Figure 2.2: Schematic showing how a general iterative process would work. The Fourier transform of the support in real space is modified in reciprocal space, this is then taken back into real space constraints are applied. The iterative process continues until a solution has been reached.

There are many variations of the phasing algorithms but they all follow a similar procedure that is described in fig. 2.2. The algorithm moves between real space and Fourier space applying constraints between each iteration and

ultimately converging to a solution. The two most used algorithms are the Error Reduction (ER) algorithm and the Hybrid Input-Output (HIO) algorithm which are variations of the generalised algorithm shown in fig. 2.2. The following description of the algorithms is based on that given by Fienup [6].

The ER algorithm works by firstly guessing an object $g(x)$ (called a support) and then taking the Fourier transform:

$$\mathcal{F}\{g(x)\} = |G(u)|e^{i\phi} \tag{2.5}$$

where ϕ denotes the phase.

The modulus of the Fourier transform from the guessed object $|G(u)|$ is replace by the modulus of the measured object $|F(u)|$ this is the modulus constraint described in fig 2.2. The inverse Fourier transform is taken of this modified reciprocal space gives the new object $g'(x)$ that should conform to the real space domain constraints. The real space domain constraint is that the object must be non-negative. The ER algorithm can be described as follows:

$$g_{k+1}(x) = \begin{cases} g'_k & \text{if } x \notin \gamma, \\ 0 & \text{if } x \in \gamma \end{cases}$$

where γ is a set of points at which $g'(x)$ violates the real space domain constraints.

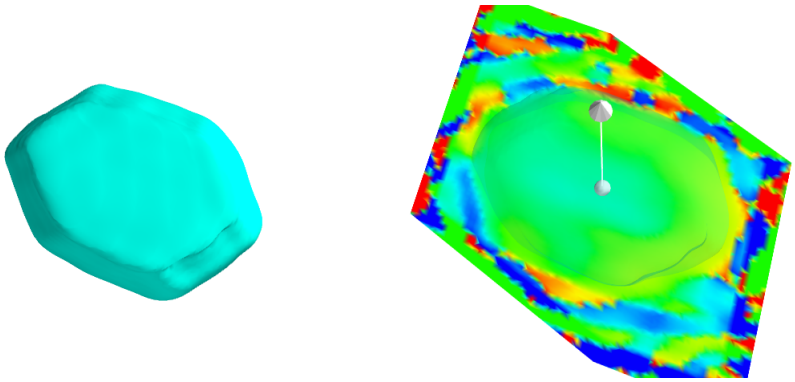
The new object $g'(x)$ is then used as the object $g(x)$ in the next iteration. This process is repeated until the algorithm has converged however ER is prone to stagnation behaviour. This stagnation occurs because the solution obtained after the k^{th} iteration becomes the starting guess for the $k + 1^{th}$ iteration. When the solution reaches a local minimum the starting guess does not vary from the previous iteration. This stagnation can be overcome by using the HIO algorithm.

The first few steps of the HIO algorithm are the same as ER the difference comes at the real space domain constraints. Instead of the previous iterate becoming the next guess as in the case of ER the output is slightly modified:

$$g_{k+1}(x) = \begin{cases} g_k & \text{if } x \notin \gamma, \\ g(k) - \beta g'(k) & \text{if } x \in \gamma \end{cases}$$

In this case, the output of the k^{th} iteration, $g'(x)$, is not used as the next guess but is used to modify the previous iteration $(k - 1^{th})$ by some small amount determined by β, where $0 < \beta < 1$. HIO resolves the stagnation issues in the ER because of the slight modification to the output. However, because of the alteration of the solution HIO never converges. It is, therefore, best to use both these algorithms in conjunction: HIO to search for the solution and ER to converge to that local minimum. The support has been briefly mentioned but is an integral part of these phase methods. The support is the object that is chosen as the starting guess $g(x)$ and is essentially the area of

space in which the image is allowed to form. It is also important in the case of the sampling ratio number of known pixels given to the algorithm is the size of the support. The support can be any shape but it is normally defined by a box which contains either a random guess or all of the same number. Different types of support algorithm can be used to make the image better such as shrinkwrap [7] or guided methods [8].



(a) (b)

Figure 2.3: (a) Reconstructed amplitude of Au nanocrystal (b) reconstructed phase the from the same Au nanocrystal using ER and HIO algorithms. The phase profile is used to map strain inside the crystal.

Fig. 2.3 shows an example of the reconstructions gained by this method. The image is of an Au nanocrystals inverted from its diffraction pattern using the HIO and ER algorithms. Fig. 2.3(a) shows the reconstructed amplitude of the 3-D image and (b) the reconstructed phase. The phase information has been used to map strain within the nanocrystals.

There are more variations of these phase retrieval algorithm with different properties, all of which can be used in CDI with different effectiveness, a comprehensive review is given by Marchensini [9].

2.4 Coherence of the Source

Coherent sources are essential in CDI experiments, without a coherent source detailed speckle diffraction patterns cannot be obtained. These speckles are not random but come from the scattering centres of the atoms. This allows a great amount of information about the object to be extracted from the diffraction pattern. The coherence of an electromagnetic field is a subtle concept and this section aims to define coherence and how it is produced from an x-ray source.

Coherence is a measure of how well two points of an electromagnetic field are correlated. For a perfectly coherent source if the electric field is defined at one point it is then known at all other points.

Fig. 2.4 shows a propagating wave at time t with two points marked on the wavefront P_1 and P_2. Coherence of an electromagnetic is described mathematically by the mutual coherence function (eq. 2.6). This is a two point correlation function of the electric field at P_1 and P_2:

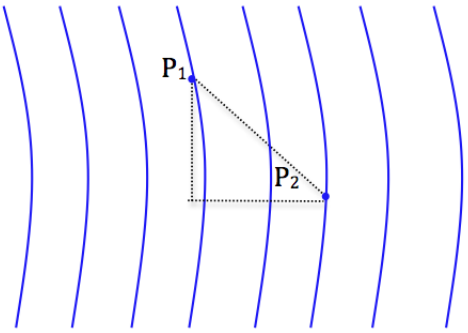


Figure 2.4: A wave front with two points P_1 and P_2 defined. The dashed vertical line shows the direction of transverse coherence and the dashed horizontal line shows the direction of longitudinal coherence.

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) = \langle E(\boldsymbol{r}_1, t) E(\boldsymbol{r}_2, t+\tau) \rangle \tag{2.6}$$

where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the positions of P_1 and P_2 respectively and τ is some measure of time from t.

This can be normalised as:

$$\gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) = \frac{\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau)}{\sqrt{\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_1, \tau)}\sqrt{\Gamma(\boldsymbol{r}_2, \boldsymbol{r}_2, \tau)}} = \frac{\langle E(\boldsymbol{r}_1, t) E(\boldsymbol{r}_2, t+\tau) \rangle}{\sqrt{\langle I(\boldsymbol{r}_1, t) I(\boldsymbol{r}_2, t) \rangle}} \tag{2.7}$$

where $I(\boldsymbol{r}_1, t)$ and $I(\boldsymbol{r}_2, t)$ are the intensities at P_1 and P_2.

For a perfectly coherent source$\gamma = 1$, which means that if the electric field is known at P_1 it is also known at P_2. For a completely incoherent source $\gamma = 0$ meaning that if the electric field is known at point P_1 it is not known at P_2. If γ lies between 0 and 1 the system is said to be partially coherent. Effects of partial coherence can be reduced with the phase retrieval algorithms by modifying the constraints in reciprocal space to include a correction. This takes the form of convolving the diffraction pattern with a Gaussian function. By applying this coherence correction the quality of the reconstruction can be increased.

Coherence does not have to be the same in all directions and therefore it helps to define longitudinal and transverse coherence, shown in fig 2.5. Longitudinal coherence can be described in the following way: if two wave fronts of wavelength λ and $\lambda - \Delta\lambda$ start off in phase the longitudinal coherence length is the distance the wavefronts have to travel before they are back in phase. It is easy to see from fig. 2.5 that the longitudinal coherence length is described by the following relationship:

$$d_L = \frac{\lambda^2}{2\Delta\lambda} \tag{2.8}$$

Transverse or spatial coherence can be defined from an extended source. Fig. 2.6 shows waves being emitted from an extended source of size D. The waves being emitted from the extremities of the source coincide at point P.

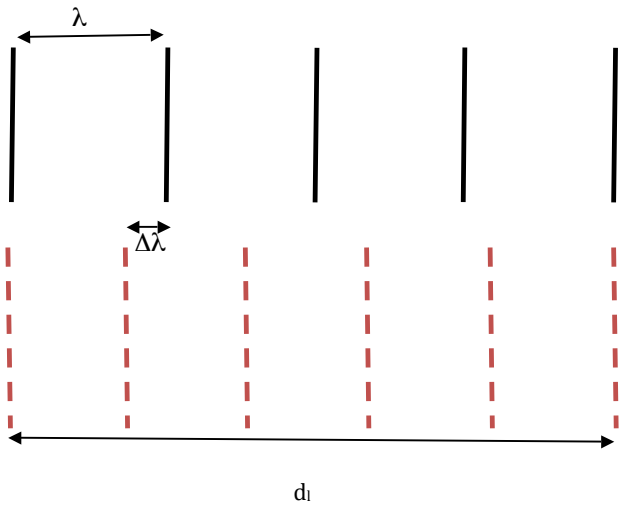


Figure 2.5: Longitudinal coherence two waves of different wavelengths start off in phase and over a distance start to become out of phase. The coherence length is the distance travelled before they come back into phase. Figure adapted from reference[10]

Travelling along one wavefront away from direction P a point is reached where the waves are out of phase this is the definition of the transverse coherence length. The wavefronts are back in phase at twice the transverse coherent length. It can be shown in fig. 2.6. that the relationship for transverse coherence length is an angular one. The transverse coherence length is given by :

$$d_T = \frac{\lambda}{2\Delta\theta} \tag{2.9}$$

If the distance between the source and the detector is large then $\Delta\theta = D/R$ therefore the transverse coherence length can be defined as:

$$d_T = \frac{\lambda R}{2D} \tag{2.10}$$

Coherence is at the core for the technique of imaging therefore the coherence properties of the source are very important. For synchrotrons sources the coherence is set by the undulators that produce the x-ray radiation from the electrons and set the size of the beam. The resulting beam shape is a Gaussian and therefore has a vertical and horizontal transverse coherence length given by:

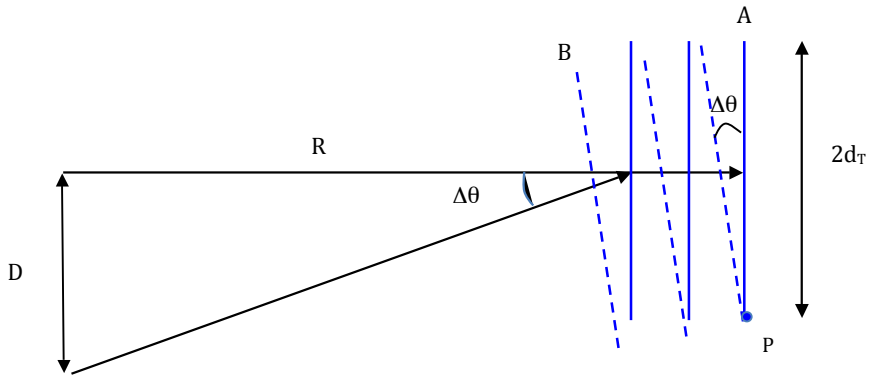


Figure 2.6: Transverse coherence. Waves emitted from different parts of the extended source source, size D coincide, at point P distance R away from the source. The transverse coherence length (d_T) is the half distance from point P to when the waves are completely out of phase. Figure adapted from reference [10]

$$d_{TH} = \frac{\lambda R}{2\pi\sigma_H} \quad d_{TV} = \frac{\lambda R}{2\pi\sigma_V} \tag{2.11}$$

where d_{TH} and d_{TV} are the vertical and horizontal transverse coherence lengths, σ_H is the horizontal size of the beam andσ_V is the vertical size of the beam. The transverse coherence can be improved by using slits at the far end to trim the size of the beam. In CDI experiments this coherence length needs to be set to the size of the sample.

In synchrotrons longitudinal coherence is inversely proportional to the bandwidth of beamline optics, therefore, coherence beamlines have the least optics possible. The biggest contributor to the obtainable longitudinal coherence is the bandpass of the monochromator which is related to its Darwin width. The optical path length difference between two different parts of the beam travelling through the monochromator determines the longitudinal coherence of the beam. If the OPLD is smaller than d_L then the beam is said to be in the coherent limit.

The OPLD is also important in looking at samples and windows in small angle transmission geometry. Fig 2.7 shows the scattering at angle θ from a beam incident on a window (S_1) and a sample (S_2) separated at a distance r. It can be seen from the geometry of the setup that the OPLD between the two rays is given by :

$$OPLD = r(1 - cos\theta) \tag{2.12}$$

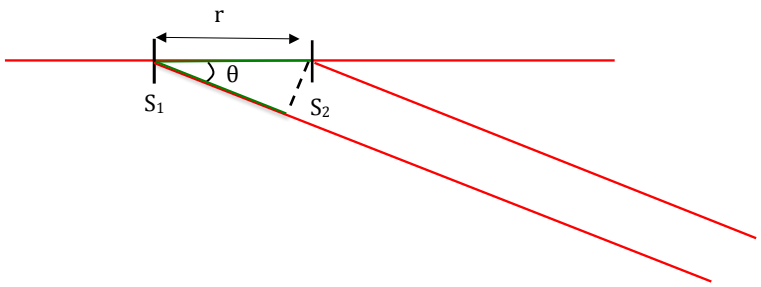


Figure 2.7: Optical path length difference between given (by the green line) from scattering at angle θ from a window S_1 and sample S_2. If the optical path length difference is greater than the longitudinal coherence length then there will be no coherent interference between the scattering from the window and sample planes.

If this OPLD is greater than the longitudinal coherence length then there is no coherent interference between the scattering from the window and the sample. This means that the signal from the window can be subtracted from the signal from the sample. This is an important consideration in order to get the cleanest signal possible.

2.5 Forward Scattering and Bragg Scattering

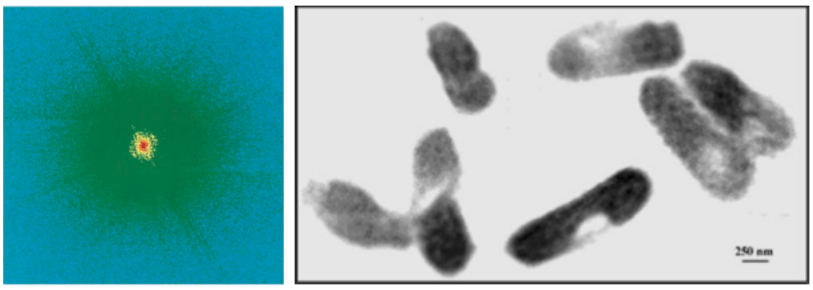
Coherent Diffraction Imaging can be done in two experimental setups: forward scattering (or transmission) geometry and Bragg geometry. Non-crystalline samples are usually measured in forward scattering mode, when the sample and the detector are in line. To gain the 3D diffraction pattern different angular projections are taken by rotating the sample through 180° in the beam.

Samples that have a Bragg peak are measured with a different alignment. The detector is offset at the reflection angle given by Bragg's Law. In this case the rocking curve around the Bragg peak is sampled in order to gain the 3D data set.

2.6 The use of CDI in Biology

Coherent Diffraction Imaging is being used to image a variety of real life examples and more importantly analyze these images to obtain useful information on a variety of structural and physical properties both in two and three dimensions. There are many examples of the use of CDI in materials science including mapping strain inside objects [11] creating images of industrial materials[12]. CDI has also been extended to biology where the main problem is the degradation of the sample due to the high levels of radiation damage. This limits the achievable resolution and the quality of data that can be taken. Despite these boundaries CDI is a technique that images the mesoscale and could be useful in biology where it is currently difficult to image these scales in thick whole specimens, especially in 3D. This section will discuss some of the work that has been done with CDI and related techniques and the problems to overcome when using synchrotron sources to image biological materials

One of the first examples of using Coherent X-ray Diffraction to image biological structures was looking at E. Coli bacteria in a dried frozen state[13]. The E. Coli bacteria were stained with MnO to improve the scattering power, this stain was also protein specific.



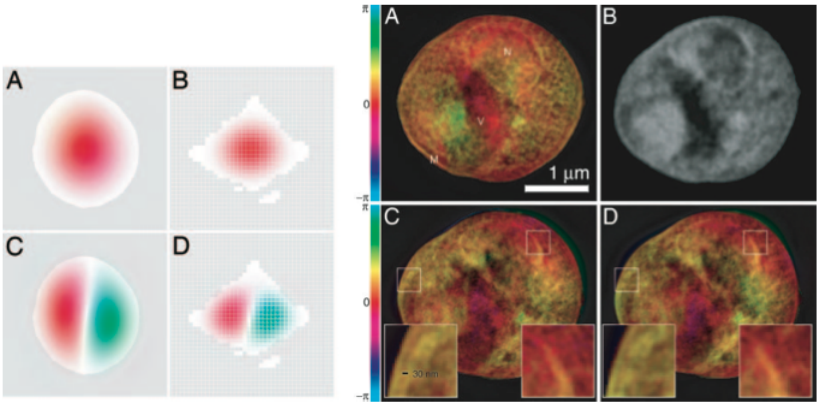
(a) (b)

Figure 2.8: (a) Diffraction pattern from MnO stained E.coli bacteria collected by coherent diffraction, (b) the reconstructed amplitude image from the diffraction pattern showing a collection of E. Coli bacteria, the dark areas show the high density regions. Figure from reference [13]

Fig. 2.8 shows the diffraction pattern and the image reconstructed from it. The dark patches in the reconstructed image show areas of high density proteins, the same areas can also be seen in fluorescence images of E. Coli. Using CDI allowed thick, whole cells to be imaged and some information about the internal structure was viewed. These images gave details of the protein structure inside the bacteria at a 30 nm resolution. However these cells were dried, frozen and stained which would inevitably change their internal structure and no 3-D images were obtained from this experiment.

Shapiro et al. managed to capture an image of a freeze dried yeast cell using CDI[14]. Using new data handling techniques to account for lost and noisy data they reconstructed images of several angular projections through a yeast cell with 30 nm resolution.

Fig. 2.9 shows the diffraction pattern and the images obtained in this experiment. Due to the weak scattering



(a) (b)

Figure 2.9: (a) The coloured regions in the centre of the diffraction pattern and the reconstruction where estimated modes were used to to reduce the effects from the missing data caused by the beamstop. The data suppressed by the beamstop and the round shape of the yeast cell meant that there was not sufficient information to form a unique reconstruction therefore modes had to be imposed to get unique convergence. (b) Reconstructions of the phase of the yeast cell from different projections. The top left image is a STXM image for comparison. The same information on the density of the regions is shown in both the STXM and the CDI images. Figures from reference [14]

of the proteins and the high energy of the beam there is a need to protect the detector from the most intense part of the beam. In this study a square with a bevelled corner was placed between the sample and the detector masking one quarter of the diffraction pattern that include the most intense part of the beam. The missing quadrant can be filled in afterwards using symmetry arguments (the pattern is assumed to be centro-symmetric). The only missing part of the data is a small black square covering the most intense part of the beam, this part of the diffraction pattern carries information about the lowest spatial frequencies. This diffraction data was very difficult to reconstruct because there was a large variance in the object's values. In order to gain a solution there had to be more constraints put on the diffraction data.

Fig. 2.9 shows the two modes of lowest constraint that were implemented on the support and the diffraction pattern in order to recover the lowest spatial frequencies. By implementing these modes a solution was obtained shown in fig. 2.9(b). A, C and D show reconstructions of the yeast cell at different angular projections. Image B was taken by Scanning Transmission X-ray Microscopy and is used for comparison. This study shows that computational techniques can be modified in order to facilitate reconstructions of difficult data. The images in fig. 2.9(b) clearly show different parts of the cell and the resolution is much better than that of the STXM image.

Very recently, steps forward in 3D imaging of yeast samples has been made by Jiang et al. who used CDI to image a whole unstained cell [15]. They used CDI to reconstruct images from 2D slices of the data at different projects which were stitched together using tomography to make a 3D image. Yeast cells are quite resistant to radiation damage so they underwent only a small amount of chemical treatment before being imaged therefore the cell structure was quite well preserved.

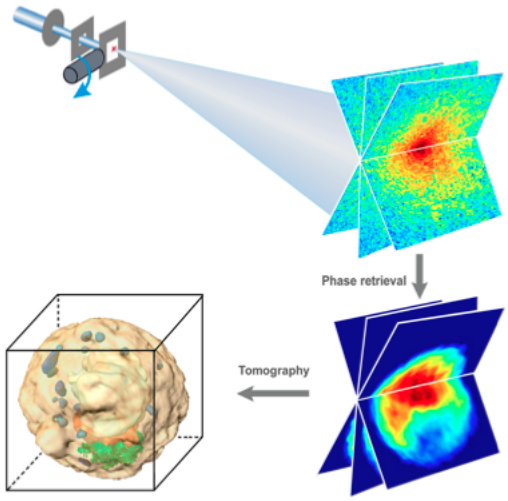


Figure 2.10: Set up of CDI -tomography experiment showing the final result of a 3D yeast cell. In the top left part of the image the experimental set up is shown, the incident beam is selected by an aperture and scattering from the aperture is blocked by the square aperture. The sample is rotated and at each rotation a diffraction pattern is taken to produce a 3D set shown in the top right. This tomographic set is then used to form an amplitude reconstruction in each plane (bottom right) which is finally made into a 3D image (bottom left). Figure from reference [15]

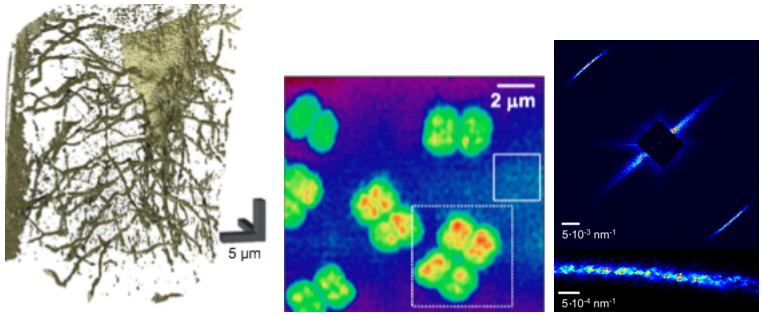
Fig. 2.10 shows in detail the 3-D model of the cell reproduced from the stack of the 2-D images. This image gives 3D spatial structure of the main parts of the cell and also a detailed surface structure with 50-60 nm resolution.

These works still have not achieved a fully 3-D reconstruction from CDI only; the work done by Jiang et al. used tomography to create their 3-D image. However, these works are a starting base that can be improved on and in doing so could have the potential to solve problems in biology that other microscopy techniques cannot. The images produced so far have better resolution than optical microscopes but do not yet provide the detail and resolution given by electron microscopes. If the resolution of CDI could be improved and more 3D images produced it could provide a way of mapping biological structures without having to section samples as in TEM.

2.7 Ptychography in Biology

Ptychography is a coherent diffraction technique that is used to image extended samples. With standard CDI the size of the sample that can be measured is limited by the achievable coherence length and the geometry of the experiment. Ptychography is a useful branch of diffraction imaging because it has a large field of view. It was first developed by Rodenburg and Faulkner in the domain of electron microscopy and has been used extensively in X-ray imaging [16]. The technique uses overlapping regions of illumination to provide constraints for the phase retrieval. The Ptychography algorithms will not be explained here but further information about current techniques can be found in references [17, 18].

There have been many examples where this technique has been used to image extended biological samples and coupled with tomography has produced 3D images.



(a) (b) (c)

Figure 2.11: Examples of Ptychography in biology (a) 3-D reconstruction of a mouse femur from data collected by the ptycho-tomography method. Figure from reference[19](b)reconstructed phase image of *D. Radiodurans* a radiation hard bacteria. This image was produced by ptychography. Figure form reference [20] (c) diffraction pattern from collagenous tissue, the centre of the diffraction pattern is suppressed by a beamstop and the 3rd order reflection can be clearly seen. In the lower part of the picture shows a zoom of the 3rd order reflection. The speckles in the diffraction pattern are apparent in this image. Figure from reference [21]

Fig.2.11 shows examples of the some the work done on biological structures in Ptychography. Fig. 2.11(a) shows the 3D structure of a mouse femur [19]. The sample was imaged using a combination of ptychography and tomography to give 3D information about the sample. This image gives very detailed information of the blood vessels inside bone. Some of the first studies using Ptychography in biology where taking images of radiation resistant bacteria , *D. Radiodurans* [20]. The image (fig.2.11(b)) shows the reconstructed phases of cluster of these bacteria with 50 nm resolution. Fig.2.11(c) shows a diffraction pattern from collagenous tissue, the outside fringes are from the 3rd order reflection of collagen. This reflection comes from the 67nm characteristic D-banding of collagen fibres. It is hoped that an image can be obtained from these regions of the diffraction pattern alone[21].

The images shown above are all of radiation resistant samples. Ptychography is a very high dosage technique but the resultant reconstructions give detailed information about an extended sample.

2.8 Summary

CDI is a useful technique for imaging with nanoscale resolution. It is currently in the stage of its development where it is being used to tackle scientific problems. There are many things still unknown, however, at the heart of this technique. There are still datasets that cannot be reconstructed by the algorithms. The oversampling limit is still in question, it is practical to collect very oversampled data in order to be sure that the criteria are met.

There have been many attempts to solve problems in biology using CDI but it is very difficult to achieve. In order to answer the questions on chromosomes structure with this technique there needs to be leaps forward with both resolution and 3D imaging. To date there has only been a few examples of achieving a 3-D image of a biological sample using CDI and it is often difficult to analyze these images to get some useful or new information.

Chapter 3

Oversampling on real data

In section 2.2 the oversampling of a diffraction pattern was discussed. Tests of the critical sampling limit proposed by Miao have only be done with simulated data. In practice this limit is adhered to but in order to ensure good data the sampling limit is far exceeded and often experiments are not attempted because the limit oversampling is being reached. I have undertaken some tests of the oversampling on a real dataset. One study was to look at the effect of nearing the critical sampling level on the reconstructed image. A correction for poorly sampled data was proposed based on modified modulus constraint algorithms and this was tested on a real data set.

3.1 Au Nanocrystals and sampling

Gold nanocrystals are good test systems to study with CDI because they are not highly complex objects, scatter well and produce clearly facetted diffraction patterns. An example of a diffraction pattern from an Au nanocrystal is shown in fig.3.1(a). This data is easy to reconstruct with the current versions of algorithms. This data was taken at the 34-ID-C beamline at the Advanced Photon Source, Illinois. The following conventions will be used throughout this chapter: the detector plane samples in 2D the horizontal direction being x and the vertical direction being y. The sampling of the rocking curve is defined as the sampling in the z direction.

It is important to think about the many ways in which data is sampled and the type of sampling that occurs. Fig. 3.1 shows the three types of sampling that can be done on the data set. Fig.3.1 (a) shows the diffraction pattern sampled at the detector, (b) the data after skipping and (c) the data after binning. The first sampling of the diffraction pattern is at the detector where at each pixel the incident signal is convolved with the point spread function of the detector. This can be thought of as convolution between the detector function (often thought of as a 2D sinc function) and the diffraction pattern. The next level of sampling comes at the data analyzing stage when the data is binned to increase the signal to noise ratio. The binning occurs by averaging together several pixels. This stage is not always done if the sample is being measured at a rate very close to the oversampling rate. Another

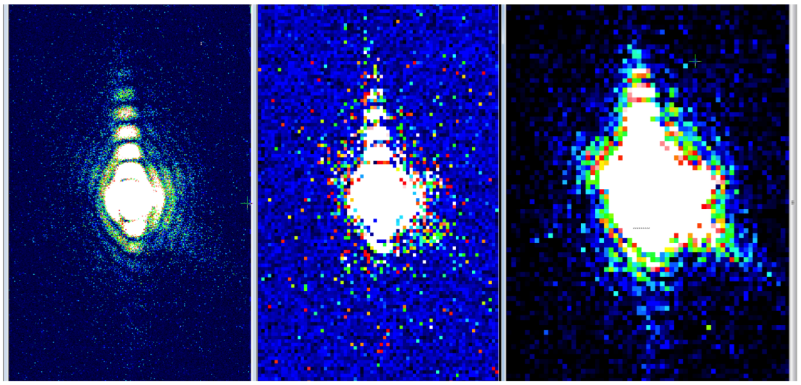


Figure 3.1: Left, Diffraction pattern sampled at the detector . Middle, diffraction pattern after skipping a skipping process. Right, diffraction pattern after binning process.

way to sample the diffraction pattern is by skipping which deletes every n^{th} pixel from the pattern. It is thought that this is the best way to sample data in the analysis stage because there is no contribution from the pixels that are deleted, unlike in the case of binning when all the information is averaged into one pixel. However, data that has undergone skipping results in poor reconstructions.

The sampling in z is different to that at the detector, as the 3D part of the diffraction data is taken by taking snapshots at different points of the rocking curve. This is the simplest definition of sampling a measurement is taken at certain intervals over a range. In this case at sampled at different angles around the Bragg peak.

The definition of the sampling rate for each type of sampling is given below. In experiments the sampling rate at the detector is decided by the geometry of the experiment. A guide to make sure that the oversampling criteria is met is calculated from as simple relationship:

$$\sigma = \frac{\lambda D}{SP(binning)} \tag{3.1}$$

where λ is the wavelength, D is the sample to detector distance, S is the sample size and P is the pixel size. The *binning* factor is included at the analysis stage. The binning averages together the pixels hence increase the size of each pixel therefore can be added into the sampling equation as a simple multiplication of the pixel size. How the data is sampled here affects the size of the support used in the reconstruction so the sampling rate here and the sampling rate defined by the support should be the same. For example the size of an Au nanocrystal is approximately 250nm, $D = 0.5$m, $\lambda = 0.139$nm and the pixel size is 23 microns at a binning of 6 in the$x$ direction and 7 in the y direction, giving an oversampling limit of $\sigma_x = 1.73$ and $\sigma_y = 2.01$. The sampling in z is given as $\sigma_z = 2.5$. When doing the reconstruction the support used is 80 pixels in x , 70 pixels in y and 64 pixels in z in an array size of 128 giving a sampling rate of $\sigma_x = 1.6$, $\sigma_y = 1.82$, $\sigma_z = 2$ this gives nearly the same number a little less due to the fact that supports have to be given to whole pixels. It is evident from this calculation that the support size needs to increase when the binning level increases.

From equation 3.1 it can be seen that experimental geometry is critical to the sampling. Often when using detectors with large pixels or when the geometry of the experiment necessitates the detector being put closer to the sample the oversampling limit is reached. It may be useful to have some way of correcting for this poorly sampled data. One way to do this is to use a type of modified modulus constraint HIO algorithm. This technique convolves the data with a correction function at the modulus constraint stage of the algorithm and updates the result. In the case of partial coherence correction function is a Gaussian of a certain width [22]. Taking the argument above that the sampling at the detector is analogous to a convolution between the diffraction pattern and a sinc function, the sampling is set by the frequency of the sinc. For a lower sampling rate a lower frequency would be used. Binning is used to sample the data at a lower rate and model large pixels. The binning can be corrected by modifying the modulus constraint in the HIO algorithm to take into account the averaging. This can be thought of as a convolution of the diffraction pattern and a box function which can be modelled by a Gaussian of the same width. When the width of the Gaussian becomes narrower than the fringes of the diffraction pattern there is a smearing out of the data, which is analogous to partial coherence. Hence, techniques used to solve for partial coherence could be used to solve correct for binning.

3.2 Simulations

Oversampling was simulated by taking a dataset and binning to different oversampling rates. An image was then reconstructed from the binning using 200 iterations of HIO and 20 iterations of the ER algorithm with a box support. A box support was used to eliminate the smoothing effects that a shrinkwrap support puts on the reconstruction. The box support was calculated so it would form tightly around the image. These box supports can contain pixels of all one value, (known as a flat start) or can contain random valued pixels, (known as a random start). Regardless of the starting support the algorithm should reach the same solution. A good way to test if the reconstructions are working is to do 15 random starts for the dataset and see if the 15 resultant solutions are the same.

The sampling rates have been calculated by equation 3.1 with the pixel size being multiplied by the level of binning. In the sampling rate in the z direction σ_z has been assigned the ratio 2.5 because that is the rate that each fringe of the pattern is sampled. The experimental data was sampled at 5 pixels per fringe but two frames were added together to create a binning in order to make the sampling rates in x, y and z more equal and so the to approach the oversampling limit proposed by Miao. Even with this binning in z the closest value to the critical sampling that can be gained is 3.09 because the support in x and y would have to be larger than the array which leaves the problem unconstrained.

The next step was to test if partial coherence corrections would improve the reconstructions badly oversampled data. This was run with 20 iterations of ER and 480 iterations of modified constraint HIO. In order to correct for the binning the Gaussian width of the coherence correction must be changed with the level of binning.

The diffraction data was binned till the reconstructions started failing. For each level of binning ten reconstructions were formed from ten random starts. The difference in the density between the 10 starts can be analyzed by looking at a map of the standard deviation of the density over the mean. These maps show where the areas of the reconstructions are different. It can also be used to compare how much the coherence corrects for the data by looking at the difference in the density between 10 corrected reconstructions and reconstructions without.

3.3 Results

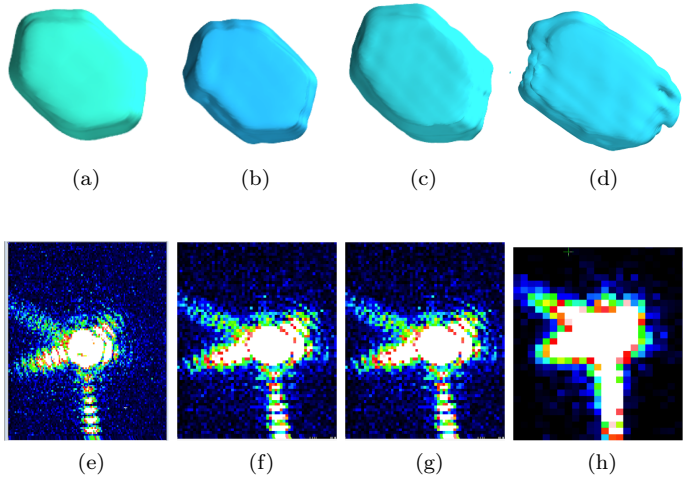
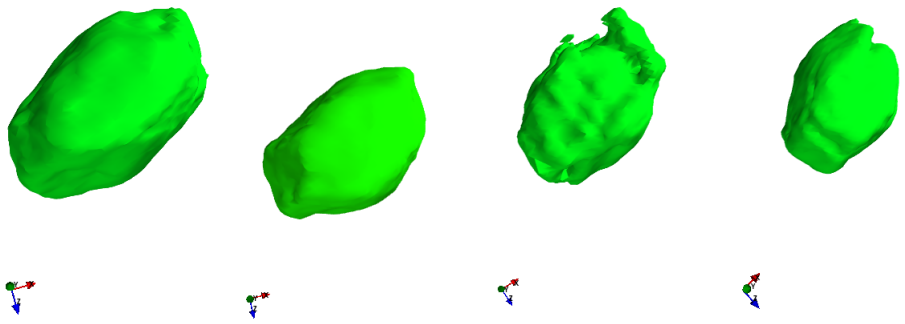


Figure 3.2: Reconstructions of Au nanocrystals and the central slice of the 3D diffraction patterns binned at different sampling rates (a)(d) $\sigma = 48.69$, (b)(e) $\sigma = 18.26$, (c)(f) $\sigma = 10.43$, (d)(g) $\sigma = 3.68$. The degradation of quality of the reconstruction can be seen with the increased binning. At $\sigma = 3.68$ the reconstruction fails resulting in the edges of the crystal not being accurately formed.

The fig. 3.2 shows the gradual degradation of the reconstruction with the binning in x and y. The effects are mainly on the edges of the crystals, which are not so sharply defined in the case of the lower binned reconstructions. The general shape of the body of the crystal remains the same in all the reconstructions. When the oversampling rate nears the oversampling limit then the reconstruction starts to fail. Fig. 3.2 (d) and (h) show the reconstruction and the not sufficiently oversampled diffraction pattern. The low sampling has resulted in the reconstruction losing density at the edges. The reconstructions (a), (b), (c), (d) are for diffraction patterns sampled at 48.7, 18.3, 10.43 and 3.68 respectively . All the were sampled at a rate larger than the oversampling limit proposed by Miao.

The partial coherence correction improves the quality of the image in all cases however this improvement is slight for well sampled data and more apparent in poorly sampled data. In all cases the faceting of the nanocrystals is sharpened by the correction though in the case of well sampled data the effect is not so easily observed, the partial coherence does not improve the image significantly (see fig.3.3(a) and (b)). Fig. 3.3 (d) shows a reconstruction from a sampling rate of $\sigma = 3.01$ that has been corrected, it can be seen that the coherence correction makes the



(a) (b) (c) (d)

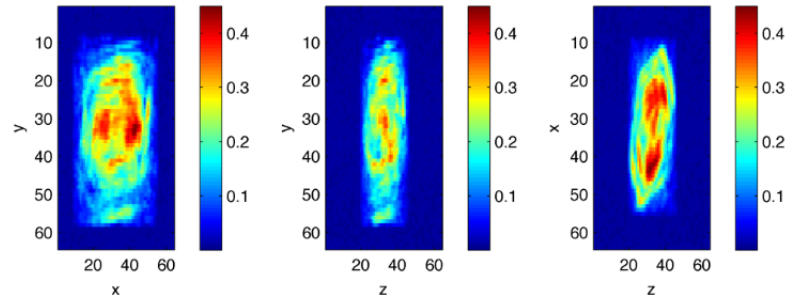
Figure 3.3: (a) Reconstruction of data sampled at $\sigma = 8.06$ (b) coherence corrected data sampled at $\sigma = 8.06$ (c) reconstruction of data sampled at $\sigma = 3.01$ (d) coherence corrected data sampled at $\sigma = 3.01$. At $\sigma = 3.01$ the coherence correction sharpens the image and adds amplitude that was missing in the no coherence correction case.

facets sharper and fills out the area of missing density in fig. 3.3(a). Here the sampling rates are still above the oversampling limit proposed by Miao but the reconstructions are starting to fail at a sampling rate of $\sigma = 3.01$.

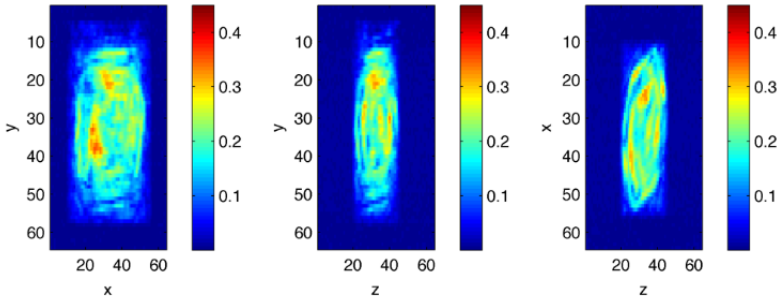
The map of density for the standard deviation over the mean shows how different the density of the reconstruction is between the 15 random starts. This shows two things: the first the uniqueness of solution and how well the coherence correction improves the density in the reconstructions. Fig. 3.4 shows cut throughs of the density in the x, y, z planes of an uncorrected and a coherence corrected reconstruction from data sampled at $\sigma = 3.01$. The areas in red show where there is the greatest difference in the density between reconstructions. In the uncorrected case there is great difference in density variations between the 15 reconstructions especially in the centre of the crystal. For the coherence corrected case there is not a great difference in the densities of the 15 reconstructions. The coherence correction improves the reconstructions and makes the solutions from the random starts more similar.

3.4 Conclusion

It was found that the data sampled above the oversampling limit of $\sigma > 2$ defined by Miao was reconstructed however the quality of the reconstruction worsened as the data became less oversampled. This particularly affected the edges of the images of the crystals and resulted in missing density from images that were sampled at a rate of $\sigma \approx 3$. It is unclear how quickly the reconstructions get worse as the oversampling limit is neared. To test this the difference in the density of the reconstructions could be compared to a well sampled reconstruction. Looking at the relationship between loss of density in the reconstruction against the sampling rate may show a trend in decreased quality of the reconstruction. To check the oversampling criteria limit as proposed by Miao is met binning would



(a)



(b)

Figure 3.4: Cuts in the x, y and z planes show ratio of the standard deviation and mean density of 15 reconstructions (a) non coherence corrected case (b) coherence corrected case. The dark red areas show where the amplitude of the 15 reconstructions differ, this usually occurs in the centre of the crystal. It can be clearly seen that there is greater agreement between reconstructions in the coherence corrected case.

have to be performed again in the z direction to reduce the sampling rate.

It is shown that the binning correction can never compensate for the loss of information due to low sampling. The correction shown for a sampling rate of $\sigma = 3.01$ does not fill in all the missing density and the sampling rate is still above Miao's sampling criteria. A poorly sampled image will always form at lower quality than a better oversampled image however the coherence correction may sharpen features that were less distinctive. The ratio of the standard deviation to mean showed that in some cases the partial coherence corrected the density mainly in the centre of the crystal. This analysis suffered from the effects of averaging between the 10 reconstructions and it may be more insightful to perform analysis on a crystal by crystal basis. It is hypothesized that a dataset sampled below the critical sampling limit and then treated with the coherence correction will not result in a reconstruction. This is because there is not sufficient information at the outset in order to solve the problem and a convolution will not add to the information present. Even though partial coherence is analogous to binning perhaps the modified modulus constraint HIO could include a correction for binning. This would be a convolution with a box function with the same width as the binning. This could be used when performing experiments with limiting geometry such as a small hutch size and small pinholes leading to the data being sampled near the limit.

3.5 Studying Chromosome Structure

The human nucleus contains 2 metres of DNA which is organized into chromatin fibres. During the cell cycle the chromatin changes shape, in the interphase the chromatin fibres are decondensed and look like strands jumbled up inside the nucleus. During mitosis the chromatin is organized in chromosomes ready for cell division. This chapter will discuss what is known about the organization of this chromosomes structure and the importance of its role. This chapter will compare different imaging techniques used in the study of chromosomes and what information can be gained from each and how if possible this information can be collated.

3.6 Levels of Structure in Mitotic Chromosomes

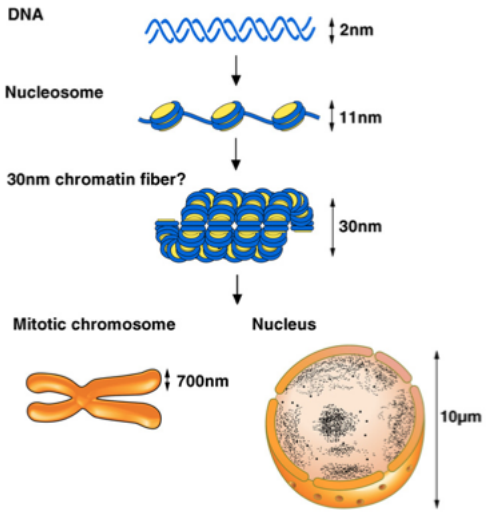


Figure 3.5: The currently agreed model of how DNA is organized into chromosomes from the level of the 2nm double helix, through the winding of the DNA around the histones to form the nucleosome of 11nm and the hypothesized solenoidal 30nm fibre. After the 30nm scale there is very little knowledge of the structure of the chromosome. Figure from reference [23]

The famous double helix model of DNA, discovered by Watson and Crick, forms the basic level of structure in the chromosome architecture. Fig. 3.5 shows how DNA is coiled into a chromosome from the DNA double helix through to the level of the chromosomes. Each nucleus inside a human cell contains a length of DNA 50 Mega base pairs long and coiled up with proteins into a structure know as chromatin. The DNA is wrapped around histone proteins, around 11nm in diameter, forming a "beads on a string" model. The unit of DNA wrapped around a histone protein is called a nucleosome. It is unclear how this chromatin coils up into a thicker fibre, a popular model is the solenoid shape of loops of nucleosomes 30nm in diameter. The chromatin is then condensed into chromatids, a pair of which form a chromosome.

The chromosome itself has a very distinct structure of two arms joined together at the centromere, the position of which defines the shape of the chromosome. There are four chromosomes shapes metacentric (centromere in

the middle), submetacentric (centromere between the middle and the end), acrocentric (centromere near the end) and telocentric (centromere at the end). The most important role of the is to protect the DNA from breakages during cell division but it is thought that the chromosome architecture also holds the answers to many questions in genomics. There is almost nothing known about how the chromatin fibres are packed into the chromosome form but in what seems a jumble of proteins and DNA lies a carefully organized structure. When looking at banding (discussed in section 3.1) the genes always appear in the same place of the chromosome and it is thought that there are links in the genetic sequence several hundred base pairs apart that must come into contact at some point [1].

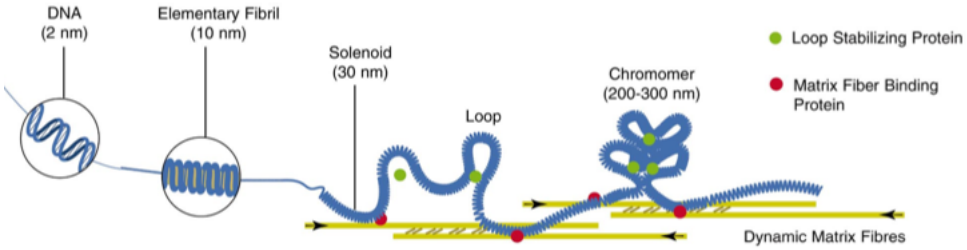


Figure 3.6: Diagram of the levels of organization inside a chromosome showing the proposed chromomere structure starting from a single strand of DNA through the nucleosome fibril and 30nm fibre to the newly proposed chromomere structure . Figures from reference [24]

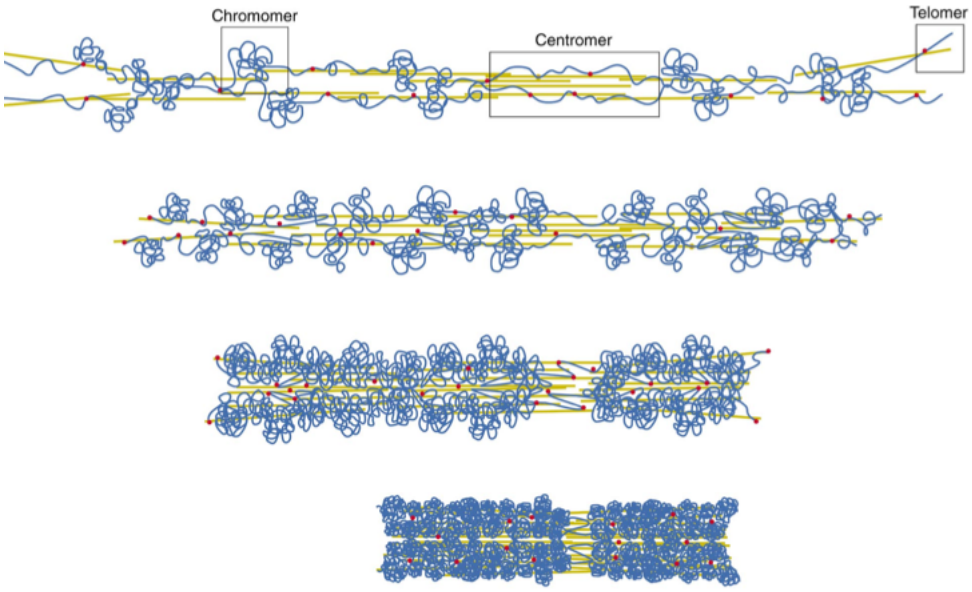


Figure 3.7: Schematic showing the formation of chromosomes in the methaphase. The DNA (shown in blue) is coiled around matrix fibres (in yellow). These matrix fibres close together during the condensation of DNA into chromosomes.

One of the most recent chromosome models has been presented by Wanner et al. who used SEM to capture high resolution images of barley chromosomes [24] shown in fig. 3.12 . Careful sample preparation allowed the chromosomes to be imaged at different stages of condensation during mitosis. These images suggest the existence of a collection of loop structures joined together by proteins, termed chromomere, of the order of 200-300 nm (shown in fig. 3.6). These form bundles of proteins that are held together with parallel matrix fibres. Fig 3.7 shows how

the chromomeres are coiled round the matrix fibres and then become gradually more condensed.

There are many approaches to unravelling the structure of the chromosome and there is certainly a need for 3D imaging to look at the chromosome architecture. There has been much work in looking at gene expression block by block but very little done on structures of the genes. A non-imaging method of studying long range interactions within the nucleus was established by Job Dekker et al. known as the 3C method [25]. This method involves looking at chromatin in the interphase and using chemical treatments to fix together strands of chromatin that approached each other. By studying this single fusion of the DNA, a three dimensional problem is reduced. This is an informatics approach of solving the structure problem.

3.7 Imaging of Chromosomes

3.7.1 Light Microscopy

The name chromosome comes from an imaging perspective, it is greek in origin from 'chromo' meaning colour and 'soma' meaning body, after the colourful dyes that bind with the different genes. The first studies of chromosomes were done with light microscopy, however under a microscope it is very difficult to distinguish between the different chromosomes. A process called G-banding was discovered in the 1970s that allowed chromosomes to be uniquely identified under a light microscope. This method uses a gene specific stain that produces dark bands that can be viewed under a light microscope. This banding is unique for each pair of chromosomes, rather like a barcode it allows them to be identified, the collection of the identified chromosomes in a nucleus is called a karyotype . This staining involves the digestion of the mitotic chromosomes with the proteolytic enzyme trypsin followed by Giemsa staining.

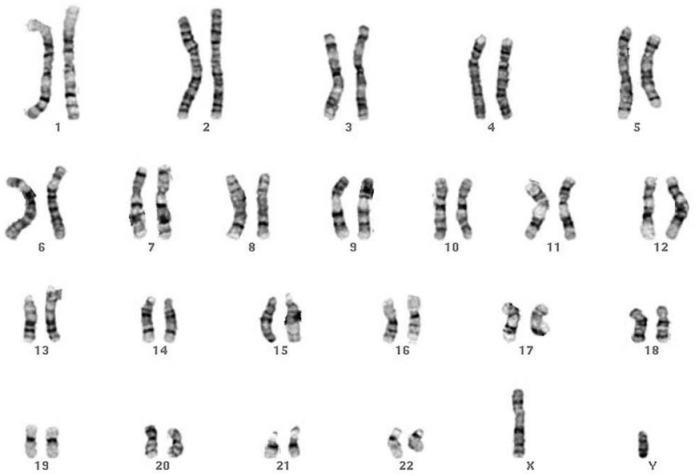


Figure 3.8: A human karyotype displayed with G-Banding produced by Giemsa staining. The band pattern is unique for each chromosome allowing them to be sorted into the 23 chromosome pairs. Figure from reference [26]

Fig. 3.8 shows an example of G-banding of the human karyotype. This type of imaging is too low resolution to tell

us anything about the structure of the chromatin fibres but it does highlight the positions of genes, sequences of structure, in the chromosomes. The fact that this can be used as identification shows that the process of packing DNA inside a chromosome cannot be completely random and that structure may play a role in genetics. The use of gene specific stains is heavily used in cytometry which have allowed methods of staining and imaging to develop. The idea of imaging gene specific stains is the basis of the technique of fluorescence microscopy.

3.7.2 Fluorescent Microscopy

Fluorescence microscopy is a technique that is used extensively in cytometry because it provides high resolution images of karyotypes. A fluorescence microscope is comprised of a high pressure Mercury or Xenon vapour source that contain wavelengths from ultraviolet to infrared. The required wavelength is selected out with a filter and the light is reflected from a dichromatic mirror through and condenser lens. This wavelength causes the dye inside the specimen to fluoresce, giving a photon with a slightly different wavelength, this is then passed through the objective lens (which is the same as the condenser lens). The emitted light passes through the mirror and then the desired wavelength is selected out with an emissions filter. This type of microscopy is extremely useful in the study of chromosomes as the fluorescent dyes can be made to target DNA, proteins or certain genes. This gives structural information about the positions of certain constituents of the chromosome.

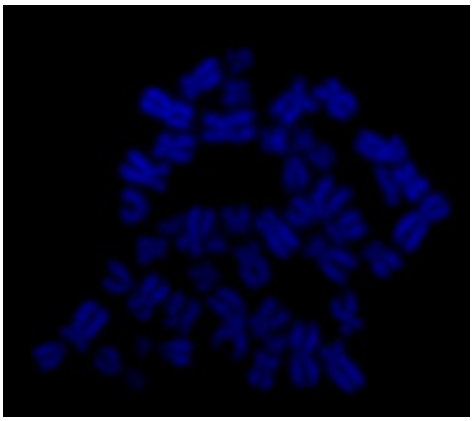


Figure 3.9: Fluorescent images of a human chromosome spread stained with a DNA sensitive stain. The DNA inside the chromosome can be clearly seen in blue, however this image contains little information on the DNA poor regions such as the centromere and telomeres.

Fig. 3.9 shows an image of a chromosome spread stained with a DNA specific stain DAPI, this was taken by a Zeiss Axio Imager 2.0 fluorescence microscope using the cyan filter. Here the DNA is clearly labelled but the resolution is not sufficient to get any information of the coiling of the DNA however it can be seen that the DNA is wound into this condensed chromosome shape and is less dense at the centromere region.

Multiple colour fluorescence in situ microscopy (M-FISH) is used heavily in cytometry. This method stains the chromosomes with different dyes that fluoresce under different wavelengths. The chromosome spread is viewed under successive filters that correspond with each dye. Different bands are expressed by different dyes and each pair

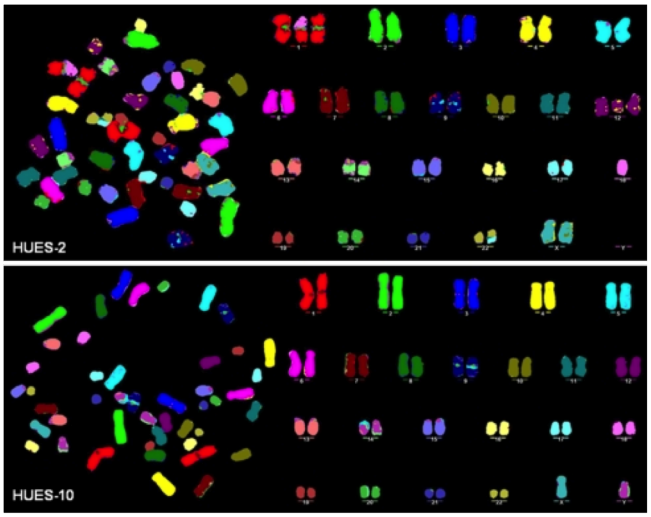


Figure 3.10: M-FISH karyotype of stem cells. Figure from reference [27]

of chromosomes has a unique code that can be seen when correlating the images taken. This method, therefore, allows the chromosomes to be identified. It also allows for a single gene to imaged in a spread of chromosomes. Fig. 3.10 is an M-FISH image of all 23 pairs of human chromosomes.

This technique of fluorescence microscopy is very useful for seeing where a gene occurs inside a chromosome but cannot tell us about the structure of that gene because the resolution of fluorescence microscopy (200 nm) if much greater than the size of the chromatin fibres (11nm). However, this technique provides a good basis to compare the X-ray images to because it identifies both the shape of the chromosome and areas sensitive to a particular stain. It is difficult to know what resolution will be achievable in the X-ray experiments and how dyes and stains may appear in the images.

3.7.3 Electron Microscopy

Electron Microscopy is a high resolution technique that has many applications. The use of Transmission Electron Microscopy is limited for the study of chromosome structure because thin samples need to be used. The metaphase chromosome is too thick to be imaged by TEM. It would be very difficult to section a metaphase chromosome without damaging the intricate network of chromatin fibres. TEM provides detailed high resolution images for samples with thicknesses below 100 nm. It was by TEM that the “beads on a string model" of histone proteins and DNA were found [28]. Scanning Electron Microscopy (SEM) has been used to image chromosomes and has revealed some interesting information about their structure, one of the benefits of SEM is that it has a large depth of field. SEM, however is limited to producing 2-D images of surface structure. This will not be able to yield information on the full 3D structure of an object.

An SEM works by radiating a sample with electrons and measuring the signal from emitted electrons that interact with the atoms in the surface of the sample. An SEM typically has two detectors one to measure secondary electrons

and the other to detect backscattering electrons, these two types of electrons give different information about the sample. X-rays emitted by the specimen from the incident electrons can also be measured to give chemical specific information about the sample. When the incident electrons collide with the sample the low energy valence electrons from the surface are emitted, these are called secondary electrons (SE). Due to the low energy of the secondary electrons only those very near the surface are emitted , the ones produced deeper being absorbed by the sample, therefore this technique is very surface sensitive. Backscattered electrons (BSE) have higher energy than secondary electrons and are emitted from deep within the sample and therefore contain information from deeper than the sample surface. The emission of BSE is caused by the scattering of the incident electrons inside the sample.

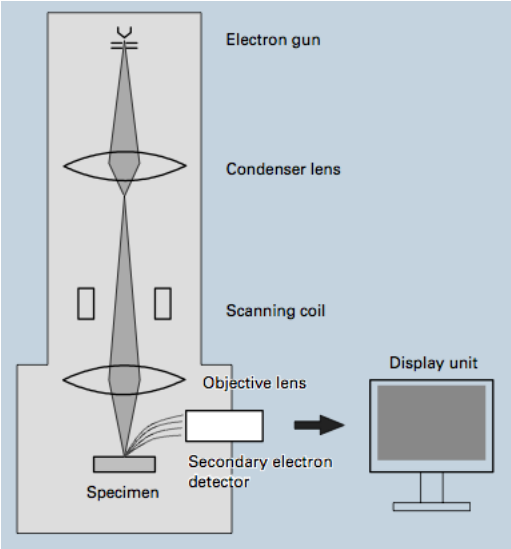


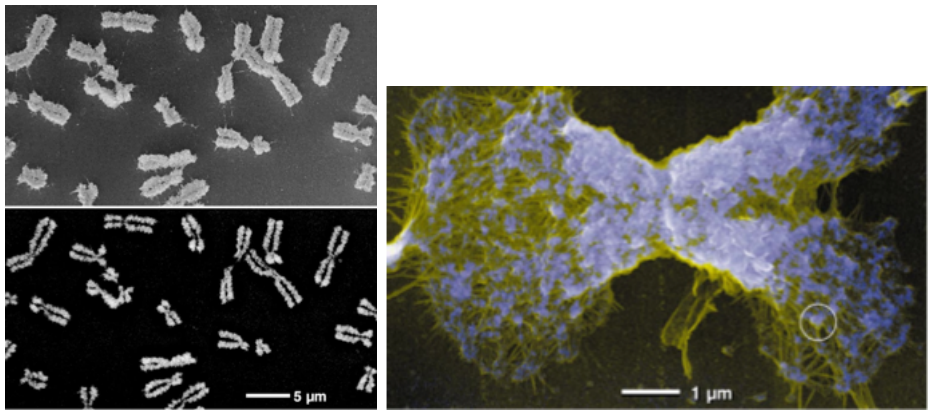
Figure 3.11: Diagram of the SEM. The electron beam is focused onto the sample by the magnetic condenser and objective lenses, electrons scattered from the sample are captured by the detector and the images are displayed on a computer monitor. Figure from reference [29]

An SEM uses a thermionic emission gun to produce electrons of tunable energy, these electrons are then focused into a probe that is scanned over the sample. In fig. 3.11 the optical set up of an SEM is described. The electrons pass through the magnetic condenser lens and aperture and then a magnetic objective lens to form a fine probe used in the scanning the sample. The scattered electrons are captured by a detector and an image is then projected onto the screen.

SEM only works with conductive samples so non-conductive samples have to be coated and viewed under low vacuum to reduce charging effects. The samples are usually sputter coated with a thin deposit of precious metal (such as gold or platinum) or a conductor like carbon. This coating is distributed in a very controlled manner so the surface contours, to which SEM is sensitive are conserved. Another way of avoiding charging effects is to look at a hydrated sample, however this has to be done in low vacuum or with an environmental SEM.

If the charging effects can be overcome, SEM is a useful imaging tool in biology. It provides very high resolution images (order of 5 nm) and has very low radiation dosage. It has a large depth of field which coupled with tilt

series and images taken in stereo can provide a three dimensional element to the image, but does not give full 3-D information. The backscattering electrons can probe deeper into the sample and therefore can obtain images of deeper structure within the sample.



(a) (b)

Figure 3.12: (a) Images of a human chromosome spread taken with both SE (top) and BSE (bottom). (b) Image of a single chromosome with BSE and SE measurements overlaid. Blue corresponds to the BSE images and yellow to the SE images. Figures from reference [24].

A careful study by Wanner et al. [24] has imaged many different parts of barley chromosome and whole chromosomes in great detail and have used this information to try and build a new model of the chromosome see section on chromosome structure. Fig. 3.12(a) shows images of the same chromosome spread with both secondary electron (SE)(top) and backscattered electron (BSE) (bottom). The chromosomes were stained with platinum blue stain, a DNA specific dye that interacts with the backscattered electrons because it contains platinum a heavy metal. In the BSE images it can be seen that there is a higher density of DNA in the chromatids then at the centromere which is also seen in the fluorescence image of the DAPI stained chromosomes (fig.3.9). The SE image shows the surface structure of the chromosomes, showing the wrapped DNA and proteins in the chromatids and the parallel protein fibres in the centromere.

Fig. 3.12 shows the SE (yellow region) and BSE (blue region) images super imposed. Here the chromosome is in a decondensed state therefore less compact here the chromomeres seem to separate from the protein matrix fibres and can be imaged. The DNA in the chromomeres is stained with platinum stain, which the BSE are sensitive to so the DNA dense regions are shown in blue in the image. Bobbles of DNA can be seen to be held around the protein fibres. These images are some of the most detailed of chromosome structure taken with any form of imaging. The high resolution of SEM and the surface sensitivity shows well the contours of the fibres wrapped up inside the DNA. However, it is very unclear from these images if there is any order to the structure, the looped chromomeres can be seen but it is difficult to see an organization of these.

These image also only give 2D information so it is very difficult to see if there is a higher order of organization. An emerging technique called 3-View is an SEM based technique that produces 3D images made from a stack of

2D images [30]. The method uses a microtone to thinly section a sample embedded in a resin. This block is then imaged by SEM, another surface layer is removed, and the block is imaged again. A stack of high resolution 2D images are built up in this way and then are stitched together using software by AMIRA to create a 3D image. This technique has been used to obtain images of several biological systems, including brain tissue [31]. We are in the process of trying to prepare chromosomes to be imaged in this way. There are several problems associated with this method for imaging chromosomes arising from how to successfully embed chromosomes in the resin. The technique of 3-View, although it has been applied to biological samples works best with hard samples because these do not move around in the resin when sectioned. If the sample moves to much whilst being cut the 3D image cannot be accurately created from the 2D slices. If this problem can be overcome the technique of 3-View could be an excellent high resolution way of imaging the 3D structures of chromosomes.

3.7.4 Scanning Transmission X-ray Microscopes

X-ray microscopes use zone plate lenses to create images, the resolution of these images is dependent of the lens used. A Scanning Transmission X-ray Microscope (STXM) uses a zone plate to focus the x-rays onto a sample, this sample is then scanned through the beam and the image is recorded on a detector. This method of microscopy is quite suitable for biological samples because it is a low dosage technique with a large field of view, it also can be used for measuring samples that are too thick for electron microscopy. Williams et al. used the STXM at Brookhaven National Laboratory to image chromosomes from a Vicia Faba plant prepared in a number of ways[32].

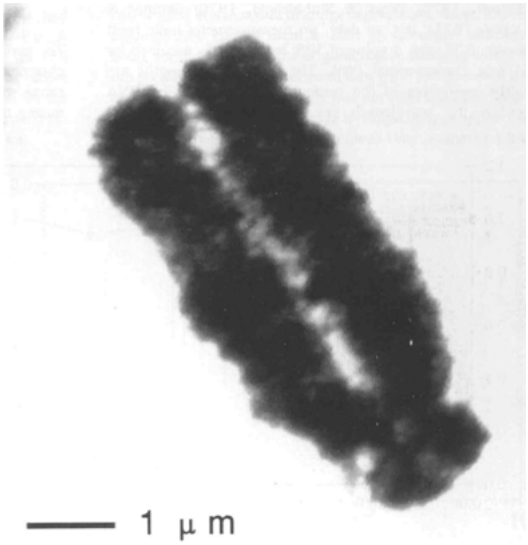


Figure 3.13: STXM micrograph of a mitotic Vicia Faba chromosome. Figure from reference [32]

An example of a STXM image of a Vicia faba chromosome shown in fig. 3.13. The image produced shows the shape of the chromosome but it is very difficult to discern any other features. Compared to the other techniques discussed nothing new can be learnt about structure from the STXM images. However, one of the most useful results of this

work is how sample preparation can effect radiation damage and will be discussed further in section 4.1.

3.7.5 Imaging Chromosomes using CDI

Nishino et al. have imaged a chromosome using CDI at the high coherence beamline, BL29XUL, in Spring-8, Japan [33]. They achieved 3D images of chromosomes at a resolution of 120 nm and were able to distinguish some of the internal structure. The chromosomes were chemically dried, fixed and mounted on a membrane by gentle centrifuging. These chromosomes were unstained but provided enough scattering power to create clear speckle diffraction patterns (see fig 3.10). Like the experiments discussed in section 2.1 the signal from the direct beam was blocked by beam stop. Three quarters of the diffraction pattern were measured by the detector, the fourth quadrant was blocked out by the square that formed the beam stop. The 4th quadrant was then inferred from symmetry arguments.

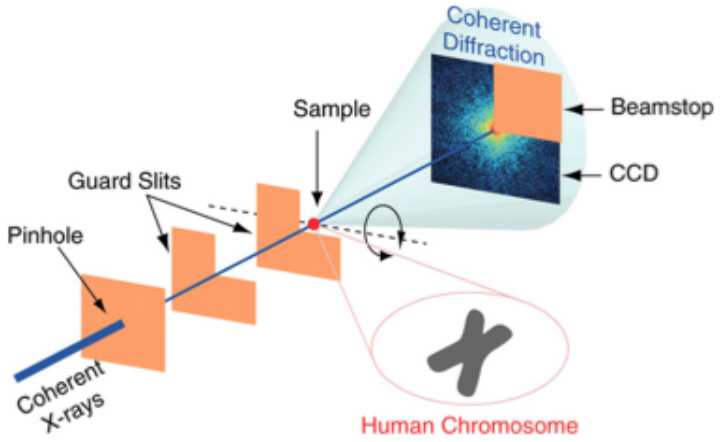
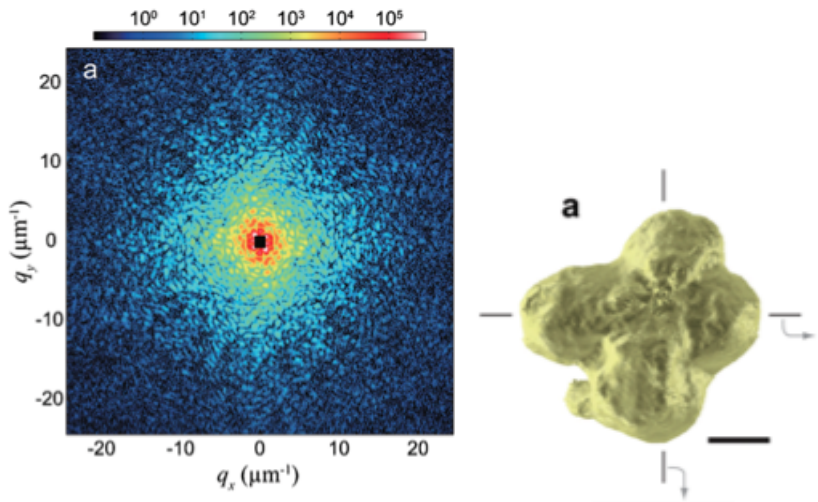


Figure 3.14: Diagram of experimental set up for imaging the chromosome. The experiment is done in transmission geometry where the incident coherent beam is cleaned up using the guard slits and the beam size is set by the aperture. The bevelled square is used in place of a beamstop this blocks out the most intense part of the beam as well as a quadrant of the diffraction pattern. The lost data can be inferred using symmetry arguments. Figure from reference[33]

A diagram of the experimental set up is shown in fig. 3.14. The area blocked by the beamstop was 23 pixels and in order to reconstruct the image the support was altered after every 500 iterations to be the previous reconstruction. The reconstructed images were take from 15 different random starting electron densities and the 10 most similar images were averaged together to achieve the reconstruction.

Two and three dimensional reconstructions of the chromosome give information about the external and internal structures. The isosurface of the 3-D reconstruction is shown in fig. 3.15 the 'X' structure of a mitotic chromosome is clearly displayed and there are some contours around the centromere region. However, in the 3D image does not give much detailed information about the internal structure. There is know information about the stacking of the fibre in the chromosomes.



(a) (b)

Figure 3.15: (a) Diffraction pattern with 4th quadrant filled in using symmetry arguments. However there is still a dark region is formed due to the need to block out the most intense part of the beam. (b) 3D reconstruction showing isosurface of a human metaphase chromosome. Figures from reference [33]

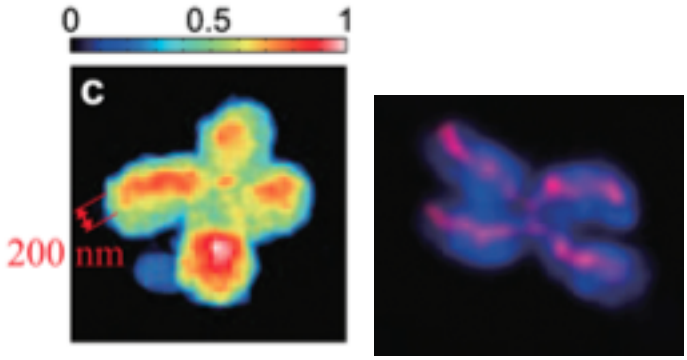
The 2D image shows some of the internal structures of the DNA (fig. 3.16), when comparing this reconstruction to a fluorescence image of a chromosome stained with DAPI (a DNA specific stain) the same information is present. Both show a wiggling spine of DNA following the arm of the chromosome, this shows that CDI is capable of giving images of deeper structures and could be a complimentary technique to staining methods as genes and DNA structures could be imaged. For this to be possible the resolution of CDI must be improved.

This work shows that it is possible to obtain images of chromosomes using CDI methods. The images shown here are not to sufficient resolution in order to be able to distinguish the 30nm chromatin fibre. These images show some higher order structure but there is not a great amount of information to be obtained from this.

3.8 Conclusions

The structure of the chromosomes is unknown past the scale of the nucleosomes. There has been much evidence for and against the existence of the 30nm chromatin fibre \cite{Maeshima10} and the way the nucleosomes are arranged in this fibre is still debated. There have been new models proposed to how this fibre is then looped into the chromosome but there is still know knowledge about how these loops are compiled. From fluorescence techniques the position of the genes in the chromosome is revealed, and more over, these can be used as unique identifiers for chromosomes. The positions of the genes never change which suggests there is some organization in the formation of DNA inside the chromosome yet there is no information about how this is structured.

The advantages and limitations of some of the many techniques that have been applied in order to answer the question on chromosome structures have been discussed above. How far can imaging techniques go to solve these



(a) (b)

Figure 3.16: (a) Two dimensional CDI reconstruction of a human metaphase chromosome (b) Fluorescence image of a chromosome stained with DAPI. The bright regions of amplitude in the CDI image correspond to the backbone of DNA shown by the DAPI stain in the fluorescence images. Figures from reference [33]

problems? The first problem in most imaging techniques is the need for 3D structure. CDI is a technique that can image in 3D but as shown in section 4.7 the resulting reconstructions do not give a lot of information on the internal structure of the chromosome.

The second problem is of resolution, it is the limiting factor in the level of detail each technique is able to measure. Electron microscopy has the best resolution of any technique however it cannot be used to image thick samples. Sectioning of samples and taking 2-D images can give some indication of 3D structure but the destruction of the sample during sectioning can ruin the accuracy of the 3D image. CDI is useful because it can image thick samples without sectioning and the limit to the resolution with the set up at I-13 at the Diamond Light source will be around 11nm- the level of the nucleosomes.

However, in order to achieve these high-resolution 3D images of chromosomes there are many considerations and technical problems to overcome, these are discussed at length in chapter 5.

Chapter 4

Experimental Considerations

In order to obtain successful reconstructions of metaphase chromosomes there are several preliminary investigations that must be done in order to test the capabilities of the system. Also, in applying this technique to biology much thought must be given to sample preparation. This chapter discusses the future experiments planned and how these will help in the process of imaging the chromosome and difficulties that will have to be faced and addressed.

The first part of this chapter focusses on the unavoidable radiation damage and how it effects both resolution and the sample degradation. Sample preparation can be used to combat this to a certain extent and is also integral in the success of getting a good resulting diffraction pattern. A section is dedicated to talking about sample preparation methods currently employed in cytometry, electron microscopy and X-ray studies and how these can be applied to CDI. The latter part of this chapter will focus on an experiment planned for later this year at the I-13 beamline at the Diamond Light Source in order to test the facility and look at some of the issues that need to be addressed for the chromosome experiments.

4.1 Radiation Damage

Radiation damage is a great concern with imaging in the life science. This damage ultimately limits resolution and in order for CDI to be a useful branch of microscopy the resolution must be comparable or better to existing techniques. The goal of the field of x-ray microscopy is to provide images of whole frozen, hydrated biological samples on the nanometer scale, however it may be better to try and image dried samples especially if they are more radiation sensitive and give greater signal. Howells et al. looked at the possibility of imaging biological samples in this limit using a combination of their own measurements and other studies [34]. The argue that for imaging it is important to think about the radiation dosage in two ways: the required dose for imaging and the maximum tolerable dose for the sample. The maximum tolerable dose is the amount of radiation a sample can endure before degradation makes it impossible to image. The required imaging dose is based on the Rose criterion, and it is the

amount of dose necessary in order to obtain sufficient statistics to gain an image.

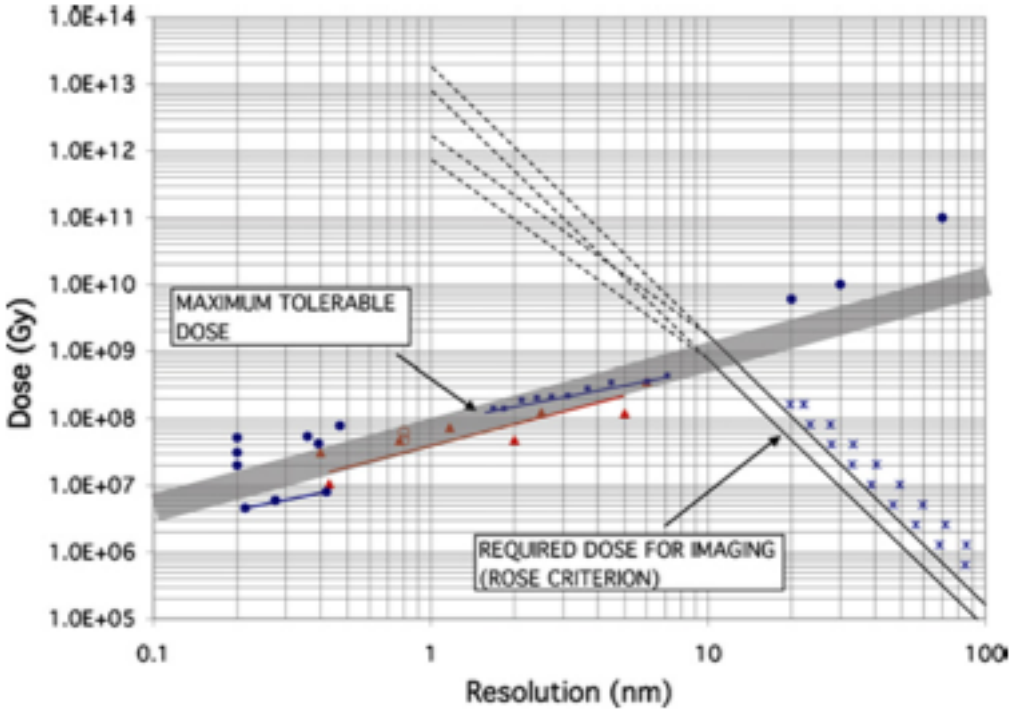


Figure 4.1: The maximum tolerable dose inferred from several studies filled circles: X-ray crystallography; filled triangles: electron crystallography; open circles: single-particle reconstruction; open triangles: electron tomography; diamonds: soft X-ray microscopy; filled squares: ribosome. The required dose calculated for a protein of density 1.35 g/cm^3 in a background of water for 1keV x-rays. Figure from reference [34]

Fig. 4.1 shows the maximum tolerable dose vs. the required dose for imaging. The required imaging dose curve was calculated for a protein, empirical formula $H_{50}C_{30}N_9O_{10}S_1$, whereas the maximum does curve was obtained by careful studies of different systems, including protein crystals, soft matter and ribosomes. The point of maximum achievable resolution lies where the two curves cross and is found to be 10nm.

There are several conditions that can be used to lessen radiation damage such as sample preparation and attention to experimental set up. In 1993, a study by Williams et al. imaged a Vicia faba chromosomes using STXM, previously discussed in section 3.6. In this study careful attention was paid to how radiation damage affected the resulting image. Using a measure of mass loss during exposure it was shown that samples that have been chemically fixed are more resistant to radiation damage than frozen hydrated samples [32]. It was also found that a single exposure of long duration gives a lower dosage than multiple images of short exposure.

Fig. 4.2 shows a chromosome that has been repeatedly exposed to different dosages of radiation, the gradual degradation of the sample can be seen.

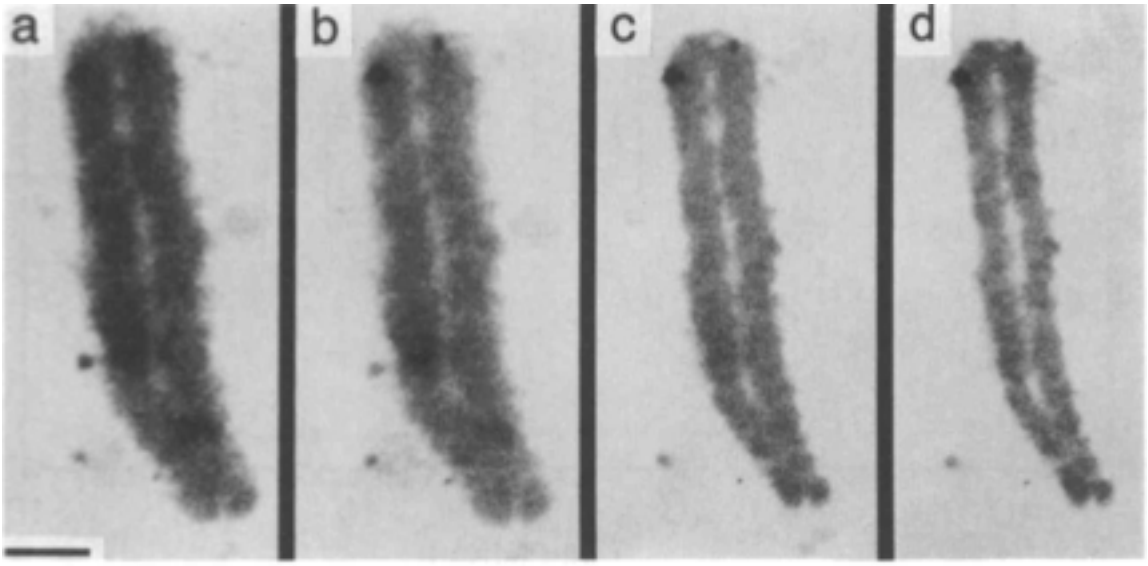


Figure 4.2: STXM images of *Vicia faba* chromosomes after 0, 35, 139 and 240 Mrad of radiation damage. These chromosomes were dried and fixed in a solution containing 0.1\% gluteraldehyde. The images series shows the gradual loss of mass with increase radiation dosage. Figure from reference [32]

4.2 Sample Preparation

Sample preparation is crucial in microscopy because it directly effects the resolution and accuracy of the image. In most cases biological samples need to be treated in some way in order to be imaged and this process can have negative effects on the structure of sample being viewed. There is a great push to image biological specimens in their natural hydrated state but this is impossible for some imaging methods. In the fight to get increasing resolution the samples must undergo some chemical treatment such as staining for fluorescence or sputter coating for SEM work.

For X-rays the preparations are currently even more intense as fixed, dried samples can withstand the incident radiation better than wet samples. However, these chemical processes can alter the 3D structure of the sample that is being so carefully imaged and this will not give accurate information of the sample in its natural state. In the case of the chromosome this chemical preparation may be useful. In all areas of biology, imaging of chromosomes has been done in either some fixed state or with some staining and this has given great insight into the genetic code. Using this work as a base point for the CDI imaging may provide a mark for resolution and the images could be correlated to expose new features. Many stains contain heavy metal which would increase the scattering power of the chromosomes they were applied to.

There are many questions in the sample preparation and protocols have to be designed and trialled. There are several important issues to address spreading the chromosomes, drying, fixing and staining. In the case of CDI it would be useful if the chromosomes were prepared in a highly condensed state giving a dense compact area of material. This would improve the signal from the chromosome at the detector. The chromosomes need to spread on

a membrane in order to be mounted in the beam and ideally they need to be isolated and not overlapping. There are three main parts to the sample preparation for chromosomes these are drying, fixing and staining.

For preparation in fluorescence microscopy, and a preparation common in cytometry the chromosomes are spread using a drop method where they are placed in solution of methanol or ethanol acetic acid this solution kills the cell and fixes the chromosomes. The methanol denatures the proteins which can result in the collapse of the histone scaffolding however the acetic acid helps to swell the cell therefore the two used in conjunction can preserve the structure of the chromosomes. Around 15 microlitres of solution is then dropped from around a 30cm height onto a glass slide. The impact with the glass causes the nucleus to shatter and the chromosomes to spread out. The chromosomes spread in clumps near the nucleus. It is very hard to replicate the size and shape of chromosomes that come out of this procedure, each batch produces a unique variety and the factors affecting this spreading method are not entirely known. One factor that is known is humidity the more humid the conditions the more widely the chromosomes are spread. It is possible in this way to control the density of the chromosome spreads in the sample [35]. Another method is the "drop-cryo" method that prepares the slides on dry ice. The cold temperatures allow the chromosomes to keep their bulky shape by slowing the evaporation.

Once the chromosomes have been spread, they are then air dried when the volatile alcohol evaporates away, which causes the chromosomes to collapse forming flat low dense regions. Air drying is not effective for preserving the surface structure that SEM is sensitive to a way to preserve the structure is to use critical point drying. In a phase diagram there are boundaries between two phases these are points of critical temperature and pressure where both phases can exist. By choosing a temperature and pressure on this boundary between the liquid and gas phase the two can exist with the same density. Critical point drying works by replacing the water in a biological sample with an inert substance (usually CO2) which can change into a gas without a change in the density or surface tension which ultimately leads to maintaining the structure upon drying. Hexamethlydisilazane (HMDS) drying is chemical alternative to critical point drying which reduces the surface tension and cross-links proteins giving the sample extra rigidity as it dries. This has produced results similar to that of critical point drying [36].

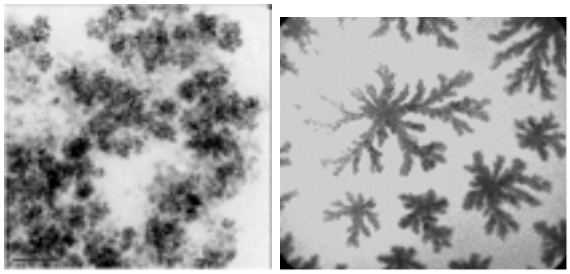
For SEM preparation the chromosomes need to be chemically fixed. Fixing cross-links the protein inside the chromosome forming a rigid mesh thus preserving the structure. The cross-linking agent used is gluteralderhyde. After it is applied the samples are then dehydrated with increasingly concentrated solutions of ethanol. In order to be viewed in an SEM the samples must be conductive they are therefore sputter coated with a thin layer of a non-reactive metal such as platinum or gold, carbon can also be used. Sputter coating is a controlled process and very accurate layer thickness can be applied.

Dyes are also heavily used in the imaging of chromosomes. There are many different dyes some protein specific, some DNA specific and some gene specific. It is important to consider chemistry of these stains, how they target and bind to certain parts of the chromosome. For work with X-rays dyes that contain heavy metals and stain the DNA will need to be used in order increase the scattering power. Work is being done to look at the chemistry of

these dyes and how thoroughly they coat the sample.

4.3 Experimental Set Up

The experiments will be undertaken on I-13 the coherence beamline at the Diamond Light Source, Harwell, Oxfordshire. This beamline can produce large coherence lengths on the scale needed to image the chromosome. However, to get in to the Fraunhofer regime the sample to detector distance must be very large. The beamline at I-13 has a 20m hutch and a detector that is mounted on a robot ; which has both a large reach and can be moved backwards and forwards. The beamline itself is 250m long and can produce highly coherent X-rays with a large coherence length. The flexibility of this beamline and its high coherence should allow for high resolution images of chromosomes.



(a) (b)

Figure 4.3: (a) TEM image of FePt nanoparticle suspensions air dried on a formvar grid. These clusters are made up of particles a few nanometers in size kept from coagulating by a surfactant (b) TEM images of Ferritin deposits. When air dried Ferritin form dendritic structures a few microns in size. These formations are similar to those of FePt because they contain a similar sized sub unit.

An experiment imaging a model chromosome is going to be attempted at this beamline. It will use model systems of a chromosome to test several aspects of the setup including: the coherence length, resolution and whether the sufficient oversampling is being achieved. The test samples will be ferritin dried on a membrane and clusters of FePt nanoparticles. These are good test systems for several reasons, they form extended structures of small particles. The dendritic structures shown in fig of the ferritin extend to 10 microns in length which is the same size as a stretched out chromosome. They contain heavy metals and so should have good scattering power but they are thin, not very dense samples so very little 3-D information can be yielded. Ferritin is a biological sample so it will give a good indication about how biological material might degrade in the beam. The FePt will not degrade and therefore will give a good test of maximum achievable resolution. It is hoped to gain sufficient resolution to see the individual particles in the reconstruction, these nanoparticles can be manufactured to be as large as 17nm. These nanoparticles are coated with a surfactant that stops the particles coalescing. When the particles are grouped together they form colloidal structures which may exhibit some local order like chromatin fibres in a chromosome. TEM images of these colloids are shown in fig. 4.3.

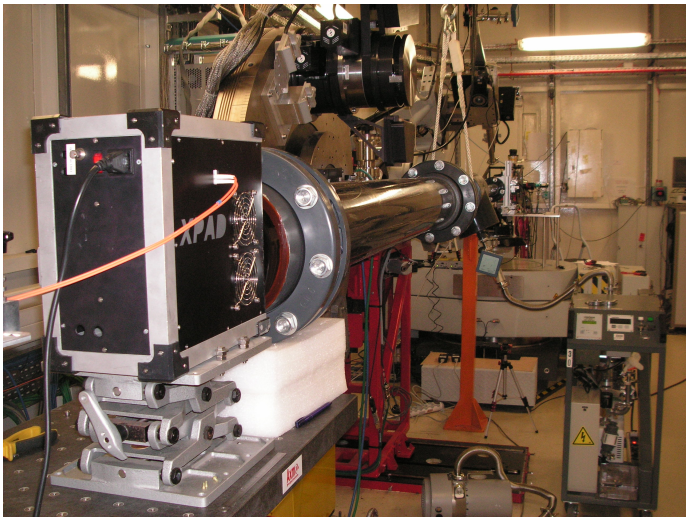


Figure 4.4: Example of a vacuum tube used in a forward scattering experiment in SOLEIL,St. Aubin, France. The vacuum tube reduces air scatter that would be seen in the diffraction pattern as noise.

In order to image the large spreads of these samples a coherence length of 50 x 50 microns is needed. As previously mentioned a long flight path is needed for this coherence length therefore there will be a lot of air scatter in the diffraction pattern. To reduce this a vacuum tube must be prepared, an example of a similar tube at the synchrotron SOLEIL is shown in fig. 4.4. Samples will be mounted on either kapton or SiN membranes, these materials are quite transparent to X-rays so will not contribute to the background noise.

The detector used in this experiment is a Medipix detector developed by the ESRF. One advantage of this detector is that is has a higher dynamic range than that of a CCD. The dynamic range of a detector is measured by the ratio of the highest intensity signal measured before the onset of saturation and the RMS of the intrinsic noise of the detector. Due to the high saturation level of the detector may mean that the most intense part of the diffraction pattern can be measured without using a beamstop. This will allow low spatial frequency information to be collected and a higher resolution achieved.

4.3.1 Application to the Chromosome

Noise will a big issue in these experiments due to the weak scattering of the chromosomes. The chromosomes are only a few microns thick, and despite the good contrast between protein and air there will be a large air path compared with size of the chromosomes. The air path, approximately 5mm, plus the windows will provide more signal than the 2 microns of chromosome hence there will be larger scatter from the background then from the chromosome. It may be possible to do a straight subtraction of the contribution from empty cell to eliminate the background, however this may not be possible due to coherence effects. As discussed in section 2.4, if the optical path length difference between the two scattering windows is larger than the longitudinal coherence length then for this set up the coherent limit is exceeded. In this scenario, it therefore, will be possible to do the straight

background subtraction.

The whole experiment will be done in vacuum. This will reduce the air scatter but will provide technical complications with the design of a vacuum chamber making it difficult to align the chromosomes in the beam. More problematic it will introduce the need for windows that will provide some scattering themselves.

One of the major problems with these type of experiments is alignment, trying to find the sample with the beam. This will be even more trying in the case of the chromosome because the current methods for spreading the chromosomes often lead them to being bunched or overlapped. The samples contain both nuclei and chromosomes spreads also strands of proteins which will be indistinguishable from their diffraction patterns. In order to gain a good image an isolated chromosomes is needed. Methods to achieve a membrane spread with just chromosomes are being tried.

Another way to easily find the sample is by using an optical microscope to align the sample. On the 34-IDC beamline at the Advanced Photon Source, Illinois , such a set up is successfully used. Marking slides with lithography patterns of heavy metals that can easily be found with an x-ray beam can also be used.

4.4 Algorithms and the Reconstructions

After discussing the experimental aspect it is important to think about the data processing. The algorithms are quite robust and can solve the phase problem under most circumstances. A data set that will reconstruct well is one that is well sampled, has little or easily removable noise and is not too affected by partial coherence. In order to get a high resolution dataset the highest spatial frequencies must be recorded as well as the information dense centre. This can hopefully be achieved with using a detector of suitable dynamic range to avoid the use of a beamstop that would completely suppress the direct beam in order to see the outer speckles. In the event of using the beamstop certain data analyzing processes can estimate the blocked data to achieve better results. Simulations with real data were run to see how a beamstop would affect the reconstruction and how the data could be accounted for. It was found that the reconstruction was extremely sensitive to the suppression of the central peak and only a small beamstop would result in a good reconstruction. There are other computational ways of getting over this problem such as estimating the central peak from the data around or creating a mask for that area. Experimental techniques can help to create very small beamstops so that there is not so many problems at the reconstruction stage, of course more data is always better than less.

As mentioned in the previous sections, the suppression of noise will be critical to this experiment. Noisy data effects greatly the reconstruction and the capability of the algorithm to give an accurate image of high resolution. There are ways to reduce the noise both in the sample preparation and the experiment but there are also ways to combat noise effects in the data handling stage. The first is to eliminate bright pixels that have obviously occurred from random sources such as cosmic rays this can be done using simple image editing software. With every diffraction

pattern measured a dark-field image is taken in order to provide a background subtraction, this is an effective way of getting rid of background noise.

Another way noise can be eliminated is in a process called "binning" this has the same definition as in information theory. When a signal is "binned" several points are summed together in order to condensed the signal a diagram illustrating. Binning is used heavily in CDI predominantly to increase the signal to noise ratio. Care must be taken when binning a signal that the oversampling criteria is still met. Many careful calculations were done looking at the effect of binning on real data and how the resulting reconstruction is effected. Theoretically , as long as the diffraction pattern is still sufficiently oversampled and the correct support is used the reconstruction should be successful each time. Simulations were done on real data of Au nanocrystals it was found that as data was sampled less and less there was a definite degradation in the image quality. In the case of the diffraction pattern of the chromosomes there may not be enough oversampling at the experimental stage to allow for binning.

The algorithms will process easily real objects but often do not find a solution when the objects have strong phase. The chromosome is a real object so the algorithms should find a solution from an adequately sampled dataset using a combination of the ER and the HIO algorithms. One way of trying to find a solution is to run many iterations of HIO. This is extremely common in complex datasets. When faced with a difficult dataset using the correct choice of support can help guide the reconstruction to the correct answer. The most simple start support is a box but it can be changed to be the reconstruction of the initial object or the shrinkwrap support can be used. The shrinkwrap support changes every few iterations and provides a tight support around the image, this results in very high definition images [7]

An issue may be the size of the datasets which could need a large amount of processing power. If the sampling is such that the diffraction pattern cannot be binned CPU processing may be needed.

Chapter 5

Future Issues

Chromosomes have been studied form many decades using a variety of techniques in order to find information on the genome. There is a need to probe the structure of the chromosome in order to answer many questions in genomics and microscopy techniques are a good way forward to probe this. The advantage of CDI over so many other techniques is that it can provide high resolution 3-D mapping of the structure of the chromosomes. There is potential with this technique to find evidence of the 30nm ordering of the chromatin fibre and the hypothesized 200-300 nm chromomere structure.

There is a question to how sensitive CDI is to the staining of the chromosomes. Staining will probably not greatly increase the scattering power of the chromosomes if the dyes contain heavy metals. By using nanoparticles of heavy metals attached to gene specific stains then the position of genes in a chromosomes may be able to be imaged with CDI. Used in conjunction with fluorescent techniques identifying the genes in CDI methods may be possible.

In order for CDI to achieve these goals the resolution must be greatly improved. Resolution is the biggest limiting factor, and it is thought that the best possible resolution that can be achieved from the set up is 10nm. This will be on the level of nucleosome so images taken at this level should be able to yield information on higher order structures. A great factor in gaining good resolution is the radiation damage to the sample. Currently work is being undertaken to find the best possible preparation conditions for the sample to reduce the radiation damage whilst preserving as well as possible the internal structure of the chromosome.

There are several computational techniques that can be used to access information about the image. Currently in CDI very little has been done in analyzing the images obtained, the techniques is starting to move in the direction of solving physical and biological problems. The images produced by CDI contain a wealth of information on the density and structure of the object which so far has not been accessed. One way to do quantitative analysis on this would be to analyze spatial correlation functions. This technique is already used in X-ray and neutron scattering because the real space 2-point correlation, otherwise known as the pair distribution function, can be found directly

from the Fourier Transform. The pair distribution function gives information about particle separation and is used heavily in the study of liquids. Other information about the arrangement of particles in a structure, such as whether they have helicity or angular correlations must be found from 3 or more dimensional correlations. This information can only be found from the real space image and not the diffraction pattern. The ability to image in 3-D will give access to the information that these higher order correlations can highlight and therefore is extremely important in the search for higher order structure.

Bibliography

[1] Baker M. Genomes in three dimensions. *Nature*, 470:289–294, 2011.

[2] Bates R.H.T. Fourier Phase Problems Are Uniquely Solvable In More. Than One Dimension .1. Underlying Theory. *Optik*, 61:247–262, 1982.

[3] Sayre D. Some implications of a theorem due to Shannon. *Acta Cryst.*, 5:843, 1952.

[4] Miao J.; Sayre D.; Chapman H.N. Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. *J. Opt. Soc. Am.*, 15(6):1662–1669, 1998.

[5] Song C.; Ramunno-Johnson D; Nishino Y.; Kohmura Y.; Ishikawa T.; Chen C-C.; Lee T-K.; Miao J. Phase Retrieval from exactly oversampled diffraction intensity through deconvolution. *Phys. Rev. B.*, 75:012102, 2007.

[6] Fienup J.R. Phase retrieval algorithms: a comparison. *Appl. Opt.*, 21(15):2758–2769, 1982.

[7] Marchensini S.; He. H.; Champman H.N.; Hau-Riege S.P.; Noy A.; Howells M.R. ; Weierstall U.; Spence J. X-ray image reconstruction from a diffraction pattern alone. *Phys. Rev. B.*, 68:140101–4, 2003.

[8] Newton M. C.; Leake S.J.; Harder R.; Robinson I.K. Three dimensional imaging of strain in a single zno nanorod. *Nature Materials*, 9:120–124, 2010.

[9] Marchensini S. A unified evaluation of iterative projection algorithms for phase retrieval. *Rev. Sci. Instr.*, 78:011301–011309., 2007.

[10] Als-Nielsen J. & McMorrow D. *Elements of Modern X-Ray Physics*. Wiley, 2010.

[11] Pfeiffer M.A. ; Williams G.J.; Vartanyants I.; Robinson I.K. Three-dimensional mapping of a deformation field inside a nanocrystal. *Nature*, 442:63–66, 2006.

[12] Chen B.; Zhang F.; Berenguer F.; Bean R.J.; Kewish C.D.; Vila Comamala J.; Chu Y. ; Rodenburg J. M.; Robinson I.K. Coherent x-ray diffraction imaging of paint pigment particles by scanning a phase plate modulator. *New J. Phys.*, 13:103022, 2011.

[13] Miao J.; Hodgson K.O.; Ishikawa T.; Larabell C.A.; LeGros M.A.; Nishino Y. Imaging whole Escherichia coli bacteria by using single-particle x-ray diffraction. *Proc. Natl. Acad. Sci.*, 100(1):110–112, 2003.

[14] Shapiro D.; Thibault P.; Beetz T.; Elser; V.; Howells M.; Jacobsen C.; Kirz J.; Lima E.; Miao H.; Neiman A.; Sayre D. Biological imaging by soft x-ray diffraction microscopy. *Proc. Natl. Acad. Sci.*, 102(43):15343Ð15346, 2005.

[15] Jiang H.; Song C.; Chen C.C.; Xu R.; Raines K.S.; Fahimian B.P.; Lu C-H. ; Lee T-K.; Nakashima A.; Urano J.; Ishikawa T.; Tamanoi F.; Miao J. Quantitative 3D imaging of whole, unstained cells by using X-ray diffraction microscopy. *Proc. Natl. Acad. Sci.*, 107(25):11234–11239, 2010.

[16] Rodenburg J.M.; Faulkner H.M.L. A phase retrieval algorithm for shifting illumination. *Appl. Phys. Lett.*, 85(20):4795–4797, 2005.

[17] Thibault P.; Menzel A.; Bunk O; David C.; Pfeiffer F. High-resolution scanning x-ray diffraction microscopy. *Science*, 321:379Ð382, 2008.

[18] Maiden A. M. & Rodenburg J. M. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 2009.

[19] Dierolf M.; Menzel A.; Thibault P.; Schneider P.; Kewish C.; Wepf R.; Bunk O.; Pfeiffer F. Ptychographic X-ray computed tomography at the nanoscale . *Nature*, 467:436–439, 2010.

[20] Giewekemeyer K.; Thibault P.; Kalbflesisch S.; Beerlink A.; Kewish C.; Dierolf M.; Pfeiffer F.; Salditt T. Quantitative biological imaging by ptychographic x-ray diffraction microscopy. *Proc. Natl. Acad. Sci.*, 2009.

[21] Berenguer F.; Wenger M.; Bean R.; Bozec L.; Horton M.; Robinson I. Coherent X-Ray diffraction from collagenous soft tissues. *Proc. Natl. Acad. Sci.*, 106(36), 2009.

[22] Whitehead L.W.; Williams G. J.; Quiney H. M.; Vine D. J.; Dilanian R. A.; Flewet S.; Nugent K. A.;. Diffractive imaging using partially coherent x-rays. *Phys. Rev. Lett.*, 103:243902, 2009.

[23] Maeshima K.; Hihara S.; Elstov M. Chromatin structure: does the 30nm fibre exist in vivo? *COCEBI*, 22:1–7, 2010.

[24] Wanner G.; Formaneck H. A New Chromosome Model. *J. Struct. Biol.*, 132:147–161, 2000.

[25] van Steensel B.; Dekker J. Genomics tools from unraveling chromosome architecture . *Nature Biotechnology*, 28(10):1089–1095, 2010.

[26] http://www.pathology.washington.edu/galleries/Cytogallery/main.php?file=human%20karyotypes, 2003.

[27] Moralli D.; Yusuf M.; Mandegar M. A.; Khoja S.; Monaco Z. L.; Volpi E.V. An improved technique for chromosomal analysis of human es and ips cells. *Stem. Cell. Rev.*, 7(2):471–4, 2011.

[28] Olins A.L.; Olins D.E. Spheroid Chromatin Units. *Science*, 4122:330–332, 1974.

[29] JEOL. Scanning electron microscope a to z: Basic knowledge for using the sem, 2009.

[30] Denk W.; Horstman H. Serial Block-Face Scanning Electron Microscopy to Reconstruct Three-Dimensional Tissue Nanostructure. *PLoS Biol.*, 2(11), 2004.

[31] Arenkiel B. R.; Elhers M. D. Molecular genetics and imaging technologies for circuit-based neuroanantomy. *Nature*, 461:900–907, 2009.

[32] Williams S.; Zhang X.; Jacobsen C.; Kirz J.; Lindaas S.; Van't Hof J.; Lamm S. Measurements of wet metaphase chromosomes in the scanning transmission X-ray microscope. *J. Microscopy*, 170(2):155–165, 1993.

[33] Nishino Y. ; Takahashi Y.; Imamoto N.; Ishikawa T.; Maeshima K. Three-Dimensional visualization of a human chromosome using coherent X-ray diffraction. *Phys. Rev. Lett.*, 102:018101–4, 2009.

[34] Howells M. R.; Beetz T.; Chapman H.N.; Cui C.; Holton J.M.; Jacobsen C.; Kirz J.; Lima E.; Marchesini S.; Miao H.; Sayre D.; Shapiro D.; Spence J.C.H; Starodub D. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *J. Elec. Spec. Rel. Phen.*, 170:4–12, 2009.

[35] Deng W.; Tsoa S. W.; Lucas J.N.; Leung C.S. ; Cheung A. L. M. A New Method for Improving Metaphase Chromosome Spreading. *Cyto. Part A*, 51A:46–51, 2003.

[36] Braet F.; de Zanger R; Wisse E. Drying cells for SEM, AFM and TEM by hexamethyldisilazane: a study on hepatic endothelial cells. *J. Microscopy*, 186:84–87, 1997.