S*upervisor*: Prof. Ian Robinson, UCL

Correlation Functions in Determining the Structure of Chromosomes

Kamila Suankulova

Submitted as part of the BSc Physics Final Year Project

UCL

ABSTRACT

Chirality is ubiquitous in nature and is important also in chromosomes. The structure of chromosomes at the 30 nm scale has being investigated as a possible link between the chirality of DNA and that of the 200 nm fibre. As 3D images are required to determine the chirality fully, which are not yet available, the investigation here focused on determining the existence of a regular structure at this scale using 2D correlation functions. We detail why we imaged the chromosomes with SEM, how the samples were prepared and obtained. The image analysis was done using direct, Fourier transform and radial distribution correlation analysis developed in MATLAB. By testing on dots, these functions were found to provide us with information about the size and spacing of various units of structure. This testing also, however, revealed problems with the radial distribution function (such as problems with normalisation) that would need to be addressed in future. By inspecting the correlation function output, it was concluded that there exists some structural units with a width and a spacing of about 30-40 nm. A full quantitative analysis of the width and spacing is yet to be done. The functions have been developed with the later adaptation into 3D in mind.

Table of Contents

1. Introduction

From the discovery of the double helical structure of DNA to the tragedy of thalidomide, the chirality of basic structures is increasingly being recognized as being of fundamental importance.

After the discovery of the structure of DNA in 1953, it was found to have right-handed chirality. This DNA wraps and folds itself in a certain way to produce the chromosomes living in the nuclei of each of our cells. We know that, in a chromosome, the structure at the 200 nm scale appears to exhibit some sort of chiral coiling[1] as the DNA wraps around the protein fibre. The intermediate stage however is yet to be determined. While there is a consensus about what happens at about 11nm (DNA wraps around proteins forming a 'beads on a string' structure called nucleosomes), there is still a question about how these strings then condense [1]; that is, the structure of the chromosome at the 30 nm scale remains unknown. In this line of research, the central question is whether the chirality of DNA is passed over to the higher order structure in chromosomes or if it is lost on the way. Several theories exist about the 30 nm nucleosome structure; some researchers claim to show that nucleosomes coil into solenoidal structures (which would be chiral) whilst others claim that no structure exists at all. This discrepancy may be due to the environment of the chromosomes when they were imaged or the imaging tool itself. In any case, to fully determine the chirality of a structure, one would need a 3D image of it, and this has yet to be achieved.

In this project we too will be working on 2D images, of both human and barely chromosomes, but we use a different imaging technique: scanning electron microscopy (SEM) to obtain them. With SEM, the preparation of the samples s very important and will be discussed. SEM may be considered an appropriate tool for this, not only due to its high surface sensitivity (with resolution reaching 1nm), but importantly also for the compositional information one can capture (up to $\sim 1 \mu m$ deep). The latter promises three dimensional information in further research. [2] We shall see that previous experiments were performed using atomic force, optical and transmission electron microscopy.

Using these 2D SEM images, we then develop functions to search for and analyse structure at the 30nm scale: these are known 2D correlation functions. This is done bearing in mind that the correlation analysis techniques on 2D images could be later applied to 3D images. A 2D correlation

[1] This has been shown in the case of fruit flies [72].

[2] See section 3.4 for an explanation of this technique.

[3] There was a significant anomaly in the bottom half of the image that was avoided.

function embodies the relationship between of two points in space that take similar values. In this project we are particularly interested in the correlation of a function with itself, which is known as an ‘autocorrelation function’. Such analysis of the SEM images of chromosomes would allow us to find the presence of preferred distance between units in space and whether such units are of a 30 nm diameter.

This autocorrelation analysis can be performed in several ways. We shall see how autocorrelation could be determined by direct calculation (by repeatedly shifting the image and multiplying by the original image), by using convolution theorem, and by employing a 1D representation of autocorrelation – the radial distribution functions (RDF).

Note that if we were to find complete disorder at that scale, this may imply the lack of chirality. If it is found to have an ordered structure on other hand the chirality of this structure would not be ruled out and would need to be investigated further with 3D images.

This analysis is performed in MATLAB with the continuous functions being replaced by an array of pixels representing the SEM output. Various regions of the images will be explored by employing the mentioned three techniques of autocorrelation analysis. The first two (direct correlation and that using the convolution theorem) are shown to give us very nearly identical results with the latter being much faster than the former. In moving to discrete data the problem of aliasing arises which we will attempt to solve by padding the area around the image. The peaks of these autocorrelation functions essentially represent the presence of a unit of structure at that location. The spacing between peaks then is then the spacing between the units of structure and a regular spacing would imply a regular structure. Furthermore the width of the peaks gives us information about the dimensions of these units of structure.

To test the effectiveness of the autocorrelation functions created in MATLAB we apply them to variations of computer generated structures, like dots and hard spheres. By varying the degree of correlation in dots, and the distribution of hard spheres, we should find the plotted correlation peaks to be in agreement with the known input.

From this analysis there appears to be some regular structure present, with a preferred distance of about 30-40 nm. Hence we suggest further modification of the 2D correlation functions explored in this project to higher order correlation functions in order to investigate 3D images for the presence of chirality. If such chirality would be found, it would address the main question set in our project;

that is, it may suggest that the chirality is maintained throughout the increasing order of structure in chromosomes. Nevertheless in this project we only address the question of the existence of a preferred distance at 30 nm, and conclude that there indeed does appear to be a regular structure at that scale.

1.1 Chirality

The chirality, a ubiquitous phenomenon in nature, was originally named after greek word ‘cheir’ for handedness. It dates back to the discovery of Louis Pasteur in 1848 when he separated left-handed from right-handed sodium ammonium tartrate (tartaric acid) crystals. [3] In chemistry a molecule which is non-superimposable on its mirror image, is described as chiral. [4]. Such chiral molecules are described as optically active: they rotate the plane of polarization of light. It means that when polarised light passes through any single chiral molecule it rotates its plane of polarisation as a result of its interactions with the molecule. These chiral molecules (in fact the majority of organic molecules), unlike achiral ones, exist as two ‘enantiomers’, either right handed (dextrorotation) for clockwise plane rotation of polarised light, or left-handed (levorotation) for anticlockwise rotation [4]. Note however that a *substance* can be made of various chiral molecules, each having a certain rotation of this plane, but that there could be an equal number of both enantiomers (called a racemic mixture [4]). In such a case the substance will be optically inactive with zero net rotation, meaning no chirality [5]. Hence the optical inactivity of a substance is determined by both the distribution of molecules, and of individual molecules it’s made up of [4]. The overall chirality of a substance is also important in, for example, how the substance tastes or smells. An interesting example of this is the case of limonene. It is found in two orientational forms: one resulting in a lemon, and the other in an orange taste/odour. [6] The same applies to the carvone molecule, with chirality being responsible for the difference in smell between caraway and spearmint [7].

The foundational bricks of life, amino acids that make up proteins are left-handed whilst the sugars in DNA are right-handed [8]. Most of the elements in human body are chiral; this means that in certain drugs for example, one enantiomer may perform a positive function whilst the other a negative function or no function at all [5]. This lesson was tragically made clear in the 1950s, when thalidomide was prescribed to pregnant women for the treatment of morning sickness. As with most artificially synthesised drugs, thalidomide was produced as a racemic mixture. It was only later discovered that while one of the enantiomers had few side effects, the other resulted in severe birth

defects [9]. Since then the chirality has been recognised as essential in drug design.

In this project we attempt to develop the basis for which the question of whether chirality at the level of DNA is passed from to the higher order organisation of chromosomes, may be answered.

In order to establish chirality, some sort of consistent structure should exist in a substance. We shall now briefly explore the background of chromosome structure at different levels and trace the possibility of chirality in chromosome as it is formed. Particularly we will be interested in a middle missing piece: the 30 nm scale. As we shall see however, the full structure information needed to determine chirality could not be recovered due to limitations in imaging the chromosomes. So far only 2D images exist and hence that implies that the direction of possible coiling cannot be determined (as for that 3 points in space are required).

2. Chromosomes

2.1 Background

In this section we shall proceed to a discussion of what chromosomes are and what is known and unknown about their structure.

Chromosomes were first described by Strasburger while observing condensation in plant cells during mitosis, when they become particularly recognisable [10]. Later the investigation of chromosomes continued with the Boveri-Sutton theory [11] [12] identifying chromosomes as carriers of genetic information with specific sites allocated to it. The discovery of Watson and Crick [13] revealed the fundamental genetic unit of the chromosome - DNA (which famously has a double helical structure).

So how does the DNA ‘bundle up’ into a chromosome? In 1976 it was postulated that the 2nm DNA winds around histone proteins forming a nucleosome fibre of about 10-11 nm. [14] [15] [16].

The nucleosomes then ‘somehow’ folds to form a higher order structure of about 200 nm called a chromomere. Finally the chromomeres fold, attaching to underlying protein matrix fibre, to form two sister arms (chromatids) joined by a central constriction called the centromere, which makes a chromosome (Fig.1). The decondensation of chromosomes was investigated, supporting the mentioned layout [17]. The regions of condensed chromomeres were found to be carriers of genetic information.

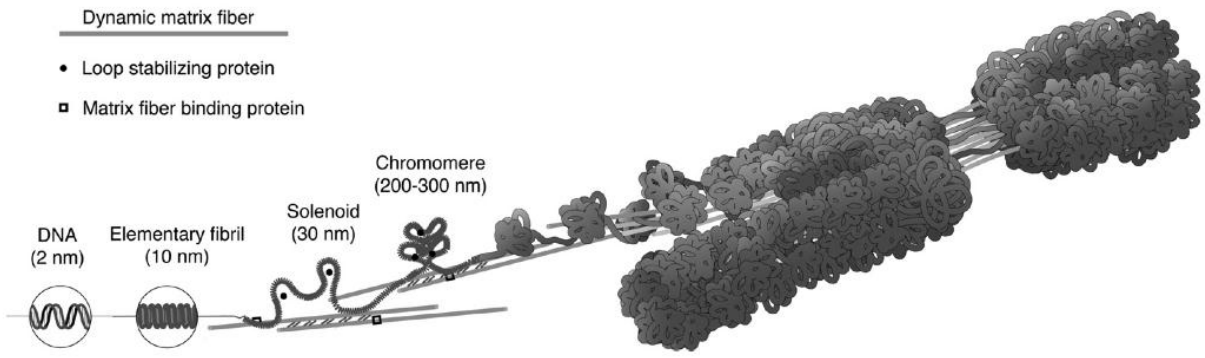


Figure 1. Representation of chromosome model. [18]

2.2. The 30 nm fibre

Our region of interest is the possible presence of an intermediate step between the 10-11 nm nucleosome fibre and the 200 nm chromomere fibre. Due to the compact structure, the 'beads on the string' nucleosome fibre was claimed to fold into a fibre with a diameter of 30 nm. [19].

For the past years this postulated 30 nm fibre has been investigating using X-ray diffraction [20], neutron scattering [21] and transmission electron microscopy [22], giving rise various theories of structure.

According to Rhodes' group [23] all of the proposed theories could be summarised into two main types: solenoid and zigzag structure.

The solenoid model consists of adjacent nucleosomes held by linker DNA, forming a one-start helix, about 33 nm in diameter (Fig. 2.a, c) [24]. The zigzag model, on the other hand, is made of nucleosomes arranged in a two-start helix, about 28 nm in diameter (Fig. 2.b, d) [25] [26]. The length of the linking DNA was said to define one type from the other. It was later claimed by the other researchers that both groups might exist simultaneously in 30 nm fibre [27], underlying the again the contentious nature of the 30 nm fibre.

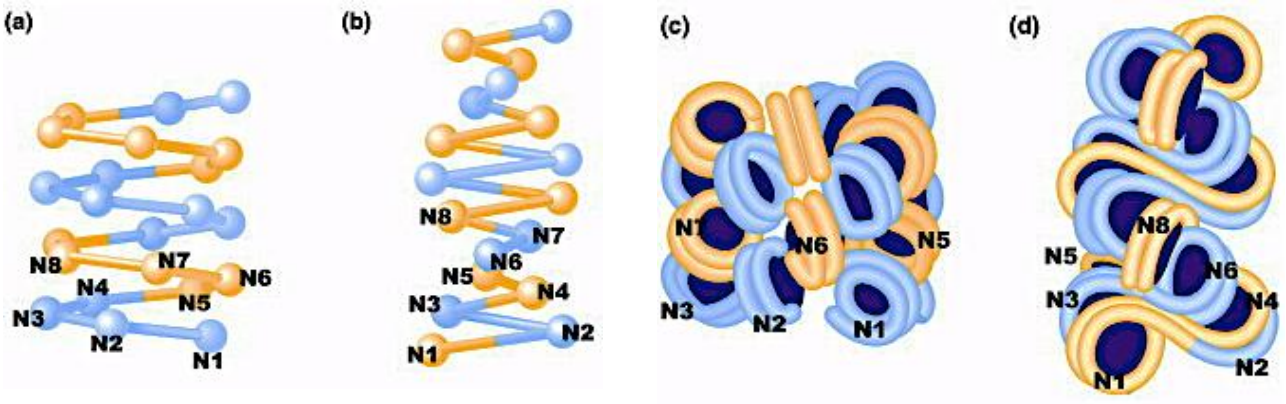


Figure 2. Display of solenoid (a),(c) one-start helix and zigzag (b),(d) two-start helix models of 30 nm fibre. [28]

The proposed models were tested *in vitro* by a group in Cambridge [29], [22] . The 10 nm chromatin fibre was reconstructed by using dialysis on e.coli DNA and on extracted chicken protein. Transmission electron microscopy (TEM) was used with the cryo technique[2] used to preserver samples close to their native state. The results suggest that the 30 nm fibre exists as a solenoidal structure [23].

In 2010 however, this was rejected by another group that found, *in vivo*, no formation of a 30nm chromatin fibre, but rather that it existed in highly disordered state similar to a polymer melt [28]. It was postulated that a similar state exists in the majority of active interphase nuclei. The proposed dynamic polymer chains were consistent with the process of condensation during mitosis and the DNA transcription and replication processes during interphase in chromosomes. [30]. Maeshima's group cited errors in interpretation and equipment on the part of Rhodes' group as the reason for obtaining discrepant results [28]. Although both groups were employing the cryo TEM method, the preparation of samples had some deviations. It is important then to understand the chromatin function, and its behaviour in live cells. Whether the preparation technique or if the variation in samples affects the 30 nm fibre observed remains an issue to be determined.

In the following sections we describe the imaging technique, method of analysis and obtained results in our investigation of this structure. We start with imaging, why SEM is appropriate and what sample preparation method was employed.

[2] See section 3.4 for an explanation of this technique.

3. Imaging

3.1 Techniques

In this section we shall proceed to discussing the imaging of the chromosomes. We begin by arguing why SEM is the most appropriate method to be used in chromosome structure analysis. What we are trying to achieve is enough resolution into the chromomere structure to be able to conduct the 30 nm structure analysis. It seems however, that something more than just resolution is needed. The surface sensitivity could be achieved for example in an atomic force microscope (AFM) [31]. AFM studies in human [32] and barley [33] chromosomes suggested a granular 30 nm structure. More recently, research by a group in Japan [34] observed meiotic chromosomes in silkworm and claimed to identify a spherical and prominent chromomere structure along the vertical axis of chromosomes. Although it has the benefit of observing samples directly in the native conditions (like in a liquid), the AFM images do not supply the necessary level of structural or compositional information of the sample, which may affect our interpretation of the 30 nm granules observed on the surface.

Such information could be resolved from the transmission electron microscope (TEM) but such would require a finite thickness of up to 100 nm in a specimen, which is not the case with chromosomes as they too large and vary from one another even in the same nuclei [17]. Hence the only technique left is scanning electron microscopy (SEM). More than just having a resolution capability down to 1 nm (with careful aligning of the electron beam), it also could be used to provide compositional information about the specimen, with electrons capable of penetrating the sample up to $\sim 1\mu$m deep with higher beam voltage applied in back scattering electron mode [35]. Such a property is not only useful in the research carried out in this project, but it also may promise the construction of 3D images, which is being currently attempted [2], [36].

Hence in this project SEM is employed for chromosome sample observation. However in order to ease the task of locating the chromosomes spreads in the sample, they are first examined under fluorescent light microscope [37]. The light microscope technique in such a case was used to estimate the 200 nm fibre [38]. Whilst SEM harms the sample with its high voltage electron beam, the light microscope allows us to locate the chromosomes spreads much quicker and without harming the sample (in addition to being faster). In the next section we shall discuss the basic operation of SEM and what makes it a special technique for imaging chromosomes.

3.1 Scanning Electron Microscope (SEM)

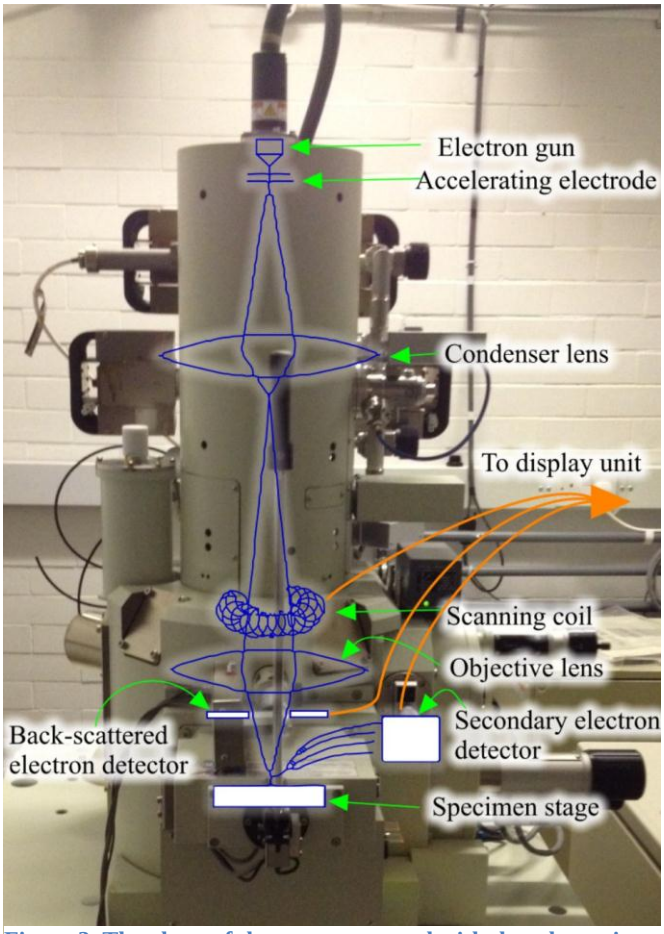


Figure 3. The photo of the apparatus used with the schematic diagram overlaid.

The basic operating principle of SEM lies in shining a high energy electron beam on the sample and detecting the scattered electrons [35 pp. 75-97]. It consists of the following features: an electron optical system, specimen stage, electron detectors, display unit and operational system.

The optical system creates an electron probe with an electron gun, condenser lens and objective lens, scanning coil. [35 pp. 21-25]

There exist several types of electron gun such as: the field-emission (FE), Schottky-emission (SE) and thermionic emission (TE) guns. [35 pp. 29-30] In this project we employed a FE gun, which extracts electrons from a wire filament by causing field electron emission. For that a voltage is applied to a thin tungsten filament with a single tungsten crystal attached to the tip, shaped to have a curvature radius of about 100 nm to make an emitter (Fig.3) [39 pp. 5-7]. A positive voltage of a few kV is then applied to the extracting electrode (located underneath the gun in the shape of disk with an opening in the middle). The electrons then tunnel through the emitter, affected by the difference in voltage. After that the electrons pass through an accelerating electrode (Fig.3), which varies the energy of the electron probe by varying the applied voltage [40 pp. 13-21].

The advantage of an FE gun in scanning electron imaging is that the extracted beam in such a case appears as thin as 5-10 nm in diameter due to the limited size of the electron source, which makes it suitable for high-resolution imaging. [40 pp. 13-21] Also, the energy spread of the electron beam is much smaller than the one produced by a TE gun (because the electrons directly tunnel through

without needing to heat the filament). This is especially important when samples are observed at a low-accelerating voltage as this energy spread determines the resolution. Hence FE appears to be superior in electron-source size, brightness (here meaning a quantity that represents the current density and parallelity of the electron beam), lifetime and energy spread (energy width) of the electron beam [35 pp. 29-40]. FE happens to perform best in morphological observations at high magnification. So after the beam is focused and accelerated to the needed energy by the electrodes, it then proceeds to the lenses.

In SEM, the lenses (Fig.3) used are magnetic and are composed of a coil of electric wire in a sealed container with a narrow gap, where a concentrated magnetic field is induced [41 pp. 39-49]. The two lenses in SEM, the condenser lens and the objective lens, are responsible for producing the fine electron beam. The strength (focal distance) of the lenses can simply be varied by changing the current flowing through the coil. The focal distance of the lenses however is also linked to the energy of the electrons going through them, because of the force exerted on the electrons due to the magnetic fields of the lenses [35 pp. 40-45]. If we keep the magnetic field on the lenses the same for example, but we change the energy of the electrons, the path of the electrons would change and therefore so would the focal distance of the lens. This is because the trajectory of higher energy electrons is less affected by the same magnetic field [40 pp. 21-35]. We can of course compensate for a change in electron energy to get the image back in focus, but this is usually limited [41 pp. 30-49]. Instead we can change the 'working distance' (WD) used (i.e. we can move the sample to the focus point) [39 pp. 9-13] to get the sample in focus once more.

By varying the excitation of the condenser lens we can adjust the thickness of the electron beam. The electron probe either narrows with stronger excitation or broadens with weaker excitation. However at stronger excitations, fewer electrons pass (i.e. a lower current). There is therefore a trade-off between beam thickness and current [35 p. 42]. With other types of SEM, there is less of a trade-off (i.e. one can adjust the current more without adversely affecting the beam thickness) [35 pp. 54-56]. However these other types generally do not reach the required resolution; FE SEM reaches this resolution because the FE gun produces a large probe current, which allows imaging the sample when the diameter goes to theoretical limit.

The scanning coils (Fig.3), located below just before the objective lens, perform the scanning function, by controlling the frame and the line times, as well as the scanning size. Some parts of the beam then proceed to the objective lens, where the beam is finally focused to the final diameter

(which then determines the resolution) [39 p. 13].

Finally the final focused beam reaches the specimen stage (Fig.3). It is basically a stable plane with motion which can be operated by the computer in following directions: horizontally (x,y) for selecting the area of interest, vertically for varying the image resolution and depth of focus; tilting and rotating for imaging different planes of a sample. [35 pp. 61-65]

After the electron beam hits the sample some of the electrons are scattered within the sample. The extent to which the electrons scatter is largely dependent on the energy of the beam and the mass and density of the constituent elements of the sample. The amount of scattering can be increased by increasing the electron energy or by increasing atomic mass and density [35 pp. 61-65]. Obviously a higher current (and hence more electrons per unit time) will also lead to more scattered electrons arriving at the detector.

When the electrons scatter off the specimen, a range of signals can be detected such as: x-rays, cathodoluminiscence, back-scattered electrons (BSE), Auger electrons (weak energy BSE) and secondary electrons (SE) [40 pp. 171-203]. This again highlights the versatility of SEM. We are interested, however, in topographical and morphological analysis, which can be obtained from the SE and BSE signals. [39 p. 27]

The scattered electrons are generally distinguished due to the energy spectra. So that SE energies do not exceed 50 eV, whilst the BSE energy ranges from the incident energy to 50 eV at least. [40 pp. 135-165]

The SE are produced from the surface of the sample when the incident beam scatters off the valence electrons. Due to the low energy of SE, those extracted at a deeper level are absorbed before reaching the surface [40 pp. 152-165]. The number of SE detected is larger if the beam hits the surface at a larger angle (rather than perpendicularly); the brightness therefore depends on the incident angle of the beam [40 pp. 152-165]. This makes SE surface sensitive and makes the electrons carriers of topographical information.

However due to their weak energy, the SE are affected strongly by any electrical charging of the sample. This results in anomalous regions in the final image. This is because the ratio between the incoming and scattered electrons is no longer unity [39 pp. 30-34]. One could reduce this effect by either reducing the electron energy or by lowering the vacuum. These would both however severely reduce the possible resolution [35 pp. 90-95]. By altering the conductivity of the sample however,

the effect can be reduced without a reduction in resolution. In the next subsection we discuss the methods we used to do this, particularly for biological samples.

Also, unevenly prominent lines give rise to an edge effect, appearing very bright on the images [35 pp. 88-97].

After the SE are scattered by the sample they get registered by the SE detector (Fig.3). This detector is made of a fluorescent scintillator (which attracts the SE by applying a high voltage to it). Before the scintillator, an Everhart and Thornley (E-T) collector electrode is positioned with a few hundred volts applied in order to control a number of SE [35 pp. 99-125]. The scintillator then emits a photon whenever an electron impacts it. This light is then directed to photo-multiplier tube (PMT) which effectively amplifies the signal, turning it back into an electric current (which is then further amplified) [35 pp. 99-125].

The second type of electrons which are scattered are BSE or reflected electrons. Due to high energy of the incident electrons, BSE collect the information from deep down into the sample, providing compositional analysis [35 pp. 75-88]. As mentioned above, the scattering of electrons inside the sample is sensitive to atomic constitution. That implies BSE yield a stronger signal with heavier atoms. There are two modes of BSE signal: TOPO (topographical) and COMPO (compositional). We are mainly interested in COMPO as this is the one that provides the compositional information [35 pp. 75-88].

The BSE are registered by positioning the detectors directly above the sample (Fig.3). The same amplification procedure as with the SE then takes place [39 pp. 13-16].

Note that, for monocrystalline samples, both the intensity of the detected BSE, and that of the SE, depends on the orientation of those crystals. This creates an ‘electron channelling contrast’ on the image (could be done by tilting sample as well) [39 pp. 28-31] [42]. However this is not the case with chromosomes.

The electric signal from the detectors is then transmitted to the display unit (Fig.3). The image formation is synchronised with electron probe scan in order to map the detected signal in space. A fast scan is used for observation, and a slow one for photos in digital format. [40 pp. 207-250]

In FE SEM the sample is kept inside a high vacuum (about 10^{-4} Pa), particularly the region of the beam path, as the electron beam is very sensitive to air and dust particles [39 pp. 4-6]. The vacuum

is achieved by the use of pumps (diffusion, turbo-molecular); particularly for FE, an oil free sputter ion pump is used to create the high vacuum. In order to maintain the vacuum an air-lock chamber is used when the specimen gets exchanged. [35 pp. 18-40]

One additional issue encountered with SEM is that, particularly in case of biological samples, the electron beam causes severe deformation of the sample. As we shall however, like the issue of sample charging, this problem can be minimised by careful sample preparation [42].

3.2 Sample Preparation

We shall now briefly discuss the methods used in preparing the chromosome samples for SEM imaging. Following Wanner's group [17], this basically consists of four stages: drop cryo, fixing, staining and drying. By using the same technique, every stage of the mitosis of rye was imaged in SEM [43]. In this project both human and barley chromosomes are prepared using this technique. For more details of the procedure see Appendix A.

Drop cryo

Having a solution of human or barley nuclei, we first need to extract the chromosomes. This is very simply achieved by dropping the solution from about 30 cm height. The simplicity of drop method allows observing mostly each of the chromosomes under a good contrast. [44] As the nuclei hit the sample slide, the cell walls break, releasing the chromosomes. After air-drying the sample is moved onto dry ice to preserve the chromosomes in a close to native state (this is called the cryo stage). Due to the cryo stage the mid drying step used elsewhere [45] could be omitted.

Fixing

The extracted chromosomes only consist of repeatedly folded DNA wrapped around a protein; this means that each structural unit is only held in place by the protein matrix and the neighbouring units, making the structure relatively unstable [17].In the next stage therefore, the chromosomes are fixed (done with a protein rich glutaraldehyde solution). Note that the samples are kept wet throughout this stage and onwards for preservation purposes [46].

Staining

As discussed above, and particularly for FE SEM, the image quality is quite sensitive to the sample

interaction with electron beam [42]. Therefore to prevent the sample from being damaged by the beam, to increase the contrast of the image, and to reduce the problems caused by sample charging, we stain the sample [18]. The stain ought to bind strongly with the DNA, and have a high atomic mass.

In the early studies of chromosome samples, osmium [47] was used as a stain. It was however successfully replaced by an organic platinum compound named 'platinum blue' $[CH_3CN]_2Pt$ [44], which was found to react with DNA-like acids. It has been successfully employed by Wanner's group to study barley chromosome samples [44]. Moreover, the platinum compound was observed to bind to regions of high chromatin concentration according to the BSE signal, which indicates those regions of the chromosome arms that are rich in DNA compared to satellite or centromere regions. However if a silver or mercury stain is applied, which binds to proteins, the BSE signal instead reveals the protein fibre surrounding the centromere [42] (which was found to be responsible for keeping tension, binding the chromosome together).

Drying

Having extracted the chromosomes, fixed and stained them, we finally need to dehydrate them. This is because SEM needs to operate in a vacuum and so all the samples have to be dried. This drying is done using ethanol [42].

We can finally proceed to actually imaging the prepared chromosome samples.

3.3 Images

To actually image the chromosomes using the Fe SEM, we first located them using a fluorescent light microscope (Fig.4.a,b.). This is done for speed, and to reduce the effects of charge build-up and damaging of the sample as discussed earlier. The rough positions of the chromosomes are identified relative to markings in the slide made with a diamond tip (Fig. 4.a).

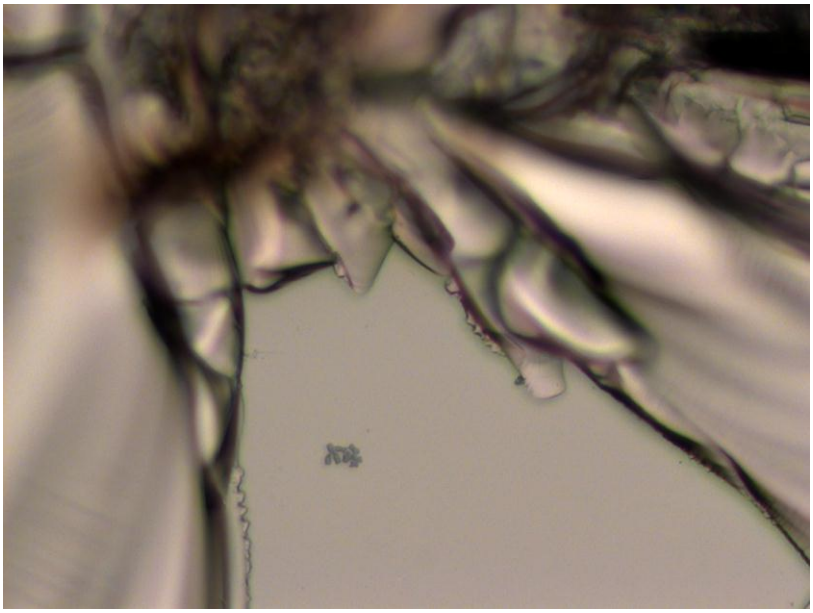


Figure 4.a. A SEM image of the chromosome spreads next to the diamond knife marker on a glass slide taken with a fluorescent optical microscope

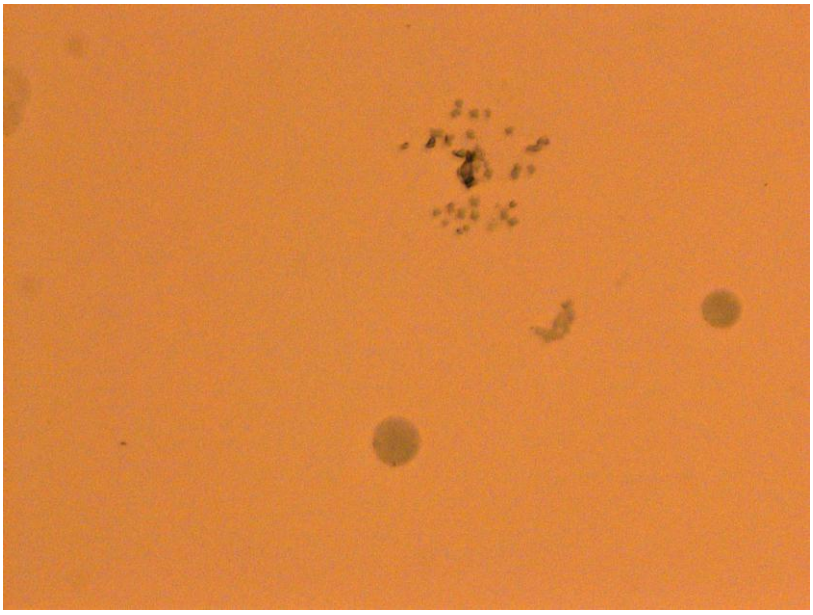


Figure 4.b. A SEM image of the chromosome spreads next to the nuclei taken with a fluorescent optical microscope

The imaging of samples is then done using a JSM 7401F-7401 with the following settings: 2.9-3mm working distance, gun voltage of 2.0 V for SE and a higher voltage of 20 kV for BSE (discussed before BSE requires a higher energy to penetrate the sample to a sufficient depth). [42]

The chromosomes were first located at the 1 μm scale (Fig.5.a) and then magnified to the 100 nm scale to achieve the high-resolution barley (Fig.5.b) and human (Fig.5.c) chromosomes images. Note that the scale of the images is not the same as resolution (the latter defined as the minimum separation that can be distinguished).

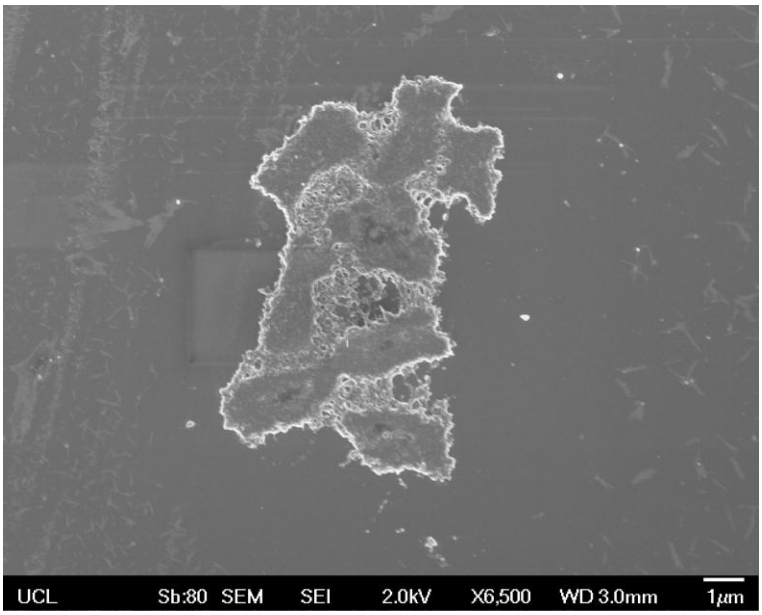


Figure 5.a. Display of the human chromosome at the 1µm scale as they are found in SEM before selecting a region of one of them and zooming in.

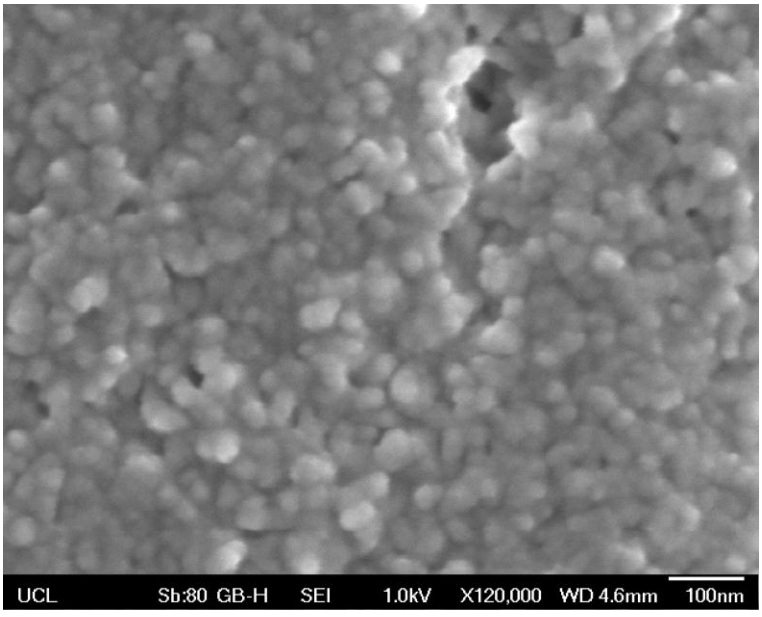


Figure 5.b. Display of the barley chromosome at the 100 nm scale. Courtesy of Laura Shemilt, UCL

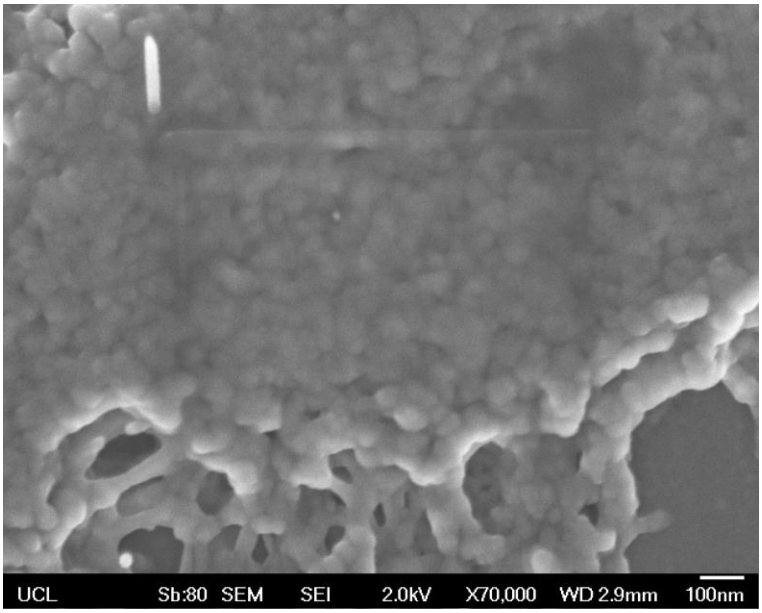


Figure 5.c. The image of the human chromosome at the 100 nm scale.

So now, having obtained images of the chromosomes, we shall discuss how the images of chromosomes are analysed using correlation functions to determine the presence of any structure.

4. Correlation functions

4.1. Autocorrelation

So here we deal with correlation functions and their application in one and two dimensions.

Two points in space are said to be correlated if they take similar values. A correlation function then is describes these relations. In order to understand the formalism behind correlation functions we begin by considering convolutions and the convolution theorem.

A convolution is defined as the integral (1) over all space of the product of one function with another function shifted by some constant, Δr. (Fig.6) [48]

$$D(\Delta r) = f(r_0) \otimes g(r_0) = \int_{-\infty}^{\infty} f(r_0) g(\Delta r - r_0) dr \qquad (1)$$

Note that the product of functions could be in the reverse order as a convolution is commutative. In other words a convolution could be imagined as a function continuously being shifted by some amount Δr, and multiplied by the other function at the value r_0 (Fig.6) [49]. Gradually

superimposing these two functions, we essentially find the weighted sum of shifted copies of one function. [50]

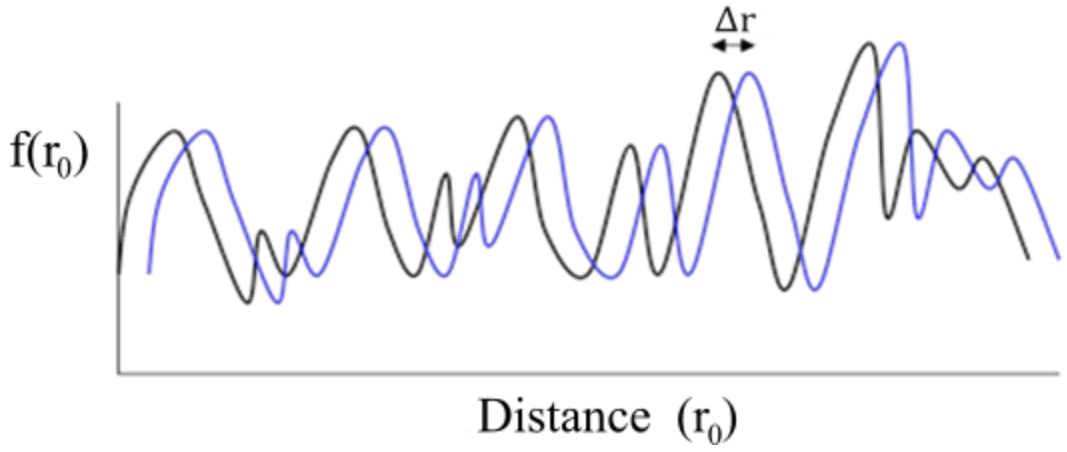


Figure 6. Display of function plotted next to the shifted function

Correlation functions are a special type of convolutions, obtained by changing the phase of one of the functions, meaning reversing of the argument from $\Delta r - r_0$ to $r_0 - \Delta r$ (2). [51] However in this project we are particularly interested in correlation of a function with itself, which is known as an autocorrelation. It allows us to determine the presence of regular structures as there would be a large correlation whenever the peaks or troughs overlap as the function is being shifted across itself. [51]

$$C(\Delta r) = f(r_0) \otimes f(r_0) = \int_{-\infty}^{\infty} f(r_0) f(r_0 - \Delta r) dr \qquad (2)$$

4.2 Fourier Transform Autocorrelation

However there exists a quicker way of obtaining an autocorrelation: by employing the convolution theorem. The theorem (3) states that: the Fourier transform of a convolution is the product of the Fourier transforms of each function [52]:

$$T\,[f(r_0) \otimes g(r_0)] = T[f(r_0)]T[g(r_0)] \qquad (3)$$

Where T indicates the Fourier transform, which is defined below (4) [52]:

$$T\,[f(r)] = \int_{-\infty}^{\infty} f(r)e^{-i2\pi s\cdot r}dr \qquad (4)$$

However when applied to autocorrelations the convolution theory changes. Due to change of phase mentioned above, the complex conjugate of one function needs to be taken [53]. The convolution theorem for correlations then becomes: the Fourier transform of the correlation function is the product of the complex conjugate of the Fourier transform of one function with the Fourier transform of the other one [52]. Given that we are particularly interested in *auto*correlations, the expression for convolution theory becomes (5):

$$T\,[f(r_0)\otimes f(r_0)] = T[f(r_0)]^{*}T[f(r_0)] = |Tf(r_0)|^2 \qquad (5)$$

Hence, by taking the inverse Fourier transform of the convolution, we will arrive to the original expression for autocorrelation (6) [52]:

$$T^{-1}|Tf(r_0)|^2 = T^{-1}T\,[f(r_0)\otimes f(r_0)] = f(r_0)\otimes f(r_0) \qquad (6)$$

This property of convolution theorem is employed in x-ray diffraction experiments [54]. In such experiments the x-ray radiation, when passing through the material, is detected as a power/intensity spectrum, which is linked with the distribution of atoms in the material. The autocorrelation is then found by taking an inverse Fourier transform of this power spectrum [54].

When the autocorrelation calculation is applied to the SEM images used in this project, the functions are represented by the array of pixels that constitutes the digital image.

4.2 Radial Distribution Function

The autocorrelation functions described above are obviously two dimensional and can thus be hard to read or analyse. It is therefore useful sometimes to represent information about correlations in space in only one dimension. The way we do this is using a radial distribution function (RDF) [54].

An RDF is simply defined as the probability of finding a particle in an areal element dr at a distance r from another particle located at $r = 0$, which is taken as the origin. (Fig.7.a,b) [54]

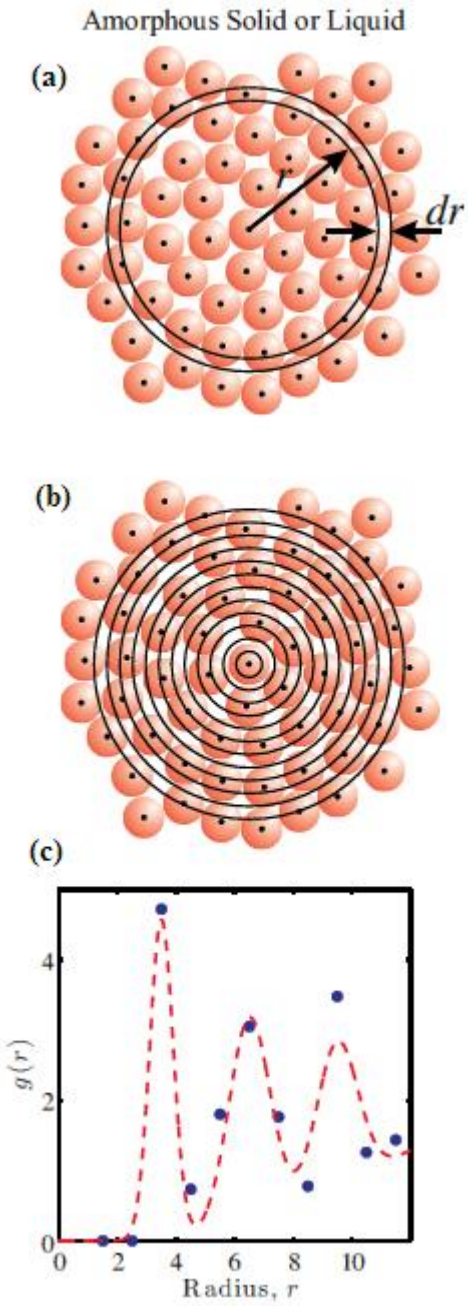


Figure 7. (a),(b)The construction of Radial Distribution Function as probability of finding two particles at the separation of radius r in non-crystalline material .(c) The resultant RDF lineout. Adapted from [54].

RDF is formally defined as a density-density autocorrelation function (7) [55]:

$$C(\boldsymbol{r}) = \langle \rho(\boldsymbol{r_i})\rho(\boldsymbol{r_i} + \boldsymbol{r}) \rangle_i \tag{7}$$

To derive this expression we should first introduce the Dirac delta function, δ, which, for our

purposes, takes non-zero value only at a single point in space, where it has a weight of unity (8) [55]:

$$\delta(r - r_j) = \begin{cases} weight\ of\ 1, & r = r_j \\ 0, & otherwise \end{cases} \tag{8}$$

These delta functions, when calculating the correlation between two points in space, function as masks or filters. This is because the integral of a delta function times another function to equal to the value of that function at the position at which the delta function is unity, [55]. Following that, the two density functions are defined in terms of delta functions as (9), (10):

$$\rho(\boldsymbol{r_i}) = \sum_{j}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r_j}) \tag{9}$$

where $\rho(\boldsymbol{r_i}) = 1$, always. [55]

$$\rho(\boldsymbol{r_i} + \boldsymbol{r}) = \sum_{j}^{N} \delta(\boldsymbol{r_i} + \boldsymbol{r} - \boldsymbol{r_j}) \tag{10}$$

where the $\rho(\boldsymbol{r_i} + \boldsymbol{r})$ only takes a non zero value whenever it is overlapped with the $\rho(\boldsymbol{r_i})$ function [55].

In order to relate this to the probability of finding a particle at some distance, r, relative to the average uniform distribution.in the same volume/area, we normalise the by the average density in the system, ρ_0 and by the average areal density, ρ_a [50] Also, to obtain the function, which is directionally independent for an isotropic system, the distribution function is averaged over the all rotational angles [54].

Hence we can define a real space representation of the correlation function of the positions of units of structure as an RDF: $g(r)$ (11) [50]:

$$g(r) = \frac{C(\boldsymbol{r})\rho_a}{\rho_0} = \frac{1}{4\pi N r^2 \rho_0} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} \delta(r - r_{ij}) \tag{11}$$

In Fig. 7 we see how an RDF is defined for the hard spheres model arranged in space (as is found directly from molecular dynamic (MD) simulation data) [54]. The peaks in the RDF correspond to

the location of the neighbouring spheres relative to the central one, averaged over the areal element. In the case of non-crystalline material, the peaks tend to cease with increasing distance, r, over a range of few spheres [54]. In this project, we look at the correlations in chromosomes and so should keep in mind that we are dealing with non-crystalline matter and hence should expect the correlation to cease with distance as displayed in Fig.7c.

4.3 X-Ray Diffraction Experiments

The density autocorrelation function is used in x-ray diffraction experiments as mentioned previously. For this, the RDF describing the structure of a liquid or crystalline material is obtained directly from the structure factor with assumption that the scattering is predominantly elastic (since the radiation energies are higher than the atomic excitation energies) [56]. Hence it is sensitive to the distribution of atomic positions in the material [54]. The intensity or power spectrum is measured [57] as I (12):

$$I = A^2\left[\sum_{ij} e^{i\boldsymbol{Q}\cdot\boldsymbol{r}_{ij}}\right] \tag{12}$$

A static structure factor S(Q) (13) is then derived from the power spectrum measured in scattering experiments [58], where g(r) is the density autocorrelation function or RDF.

$$S(Q) = 1 + 4\pi\rho \int_0^\infty r^2\,(g(r) - 1)\,\frac{\sin(Qr)}{Qr}\,dr \tag{13}$$

Hence g(r) (14) can be obtained by taking the inverse Fourier Transformation of the structure factor, normalised by the average number density of particle in the sample, ρ_0 [59].

$$g(r) = 1 + \frac{1}{2\pi^2 r\rho_0}\int_0^\infty Q\,(S(Q) - 1)\,sin(Qr)dQ \qquad (14)$$

After having looked at one of the successful applications of correlation functions as method of measuring distribution of the positions atoms in a substance, we should next proceed to analysing the structure of chromosome samples through correlation functions (in this case on SEM images).

5. Results and Analysis

Having now defined the formalism behind correlation functions, we now implement them in MATLAB to analyse our SEM images. The functions and scripts used to obtain these images can be found in Appendix B.

The analysis of correlation is performed by the direct autocorrelation calculation, by the Fourier transform method (using convolution theorem), and finally by transforming autocorrelation in 1D to achieve the radial distribution function.

Note that all the axes have been rescaled from pixels into nanometres (128 px : 100 nm).

5.1 Autocorrelation

Beginning then with the direct autocorrelation, we analysed the image of barley chromosomes (Fig.5.b). For ease of computation the image is first split into smaller (in this case into 100x125nm) regions and the direct autocorrelation is calculated by using eqn. (2). The direct autocorrelation function of the second arbitrary region (of the 8x8 possible regions) was plotted in Fig.8. To make this 2D function more clear, we also plot an x-lineout (a row of pixels) from the middle of the function. This allows us to see how the amount of correlation varies along this line.

The autocorrelation of this 100x125 region shows a distinct bright peak of correlation in the middle. This is always the case with autocorrelation functions as the central point is where the image has neither been shifted in the x nor the y directions. The x-lineout shows relatively weak peaks of correlation on the sides of the central peak which might represent a preferred distance to the nearest other structural unit in the image. As we stated previously, because chromosomes are non-crystalline, this lack of long range order in the peaks is consistent with our expectations for this

type of sample. The peaks in the x-lineout however do appear to have a width of about 30-40 nm which would correspond to the size of the structures visible in the image region. However the secondary peaks, with a width of about 20 nm, may then correspond to the presence of a second layer of structure which might be only partially visible.

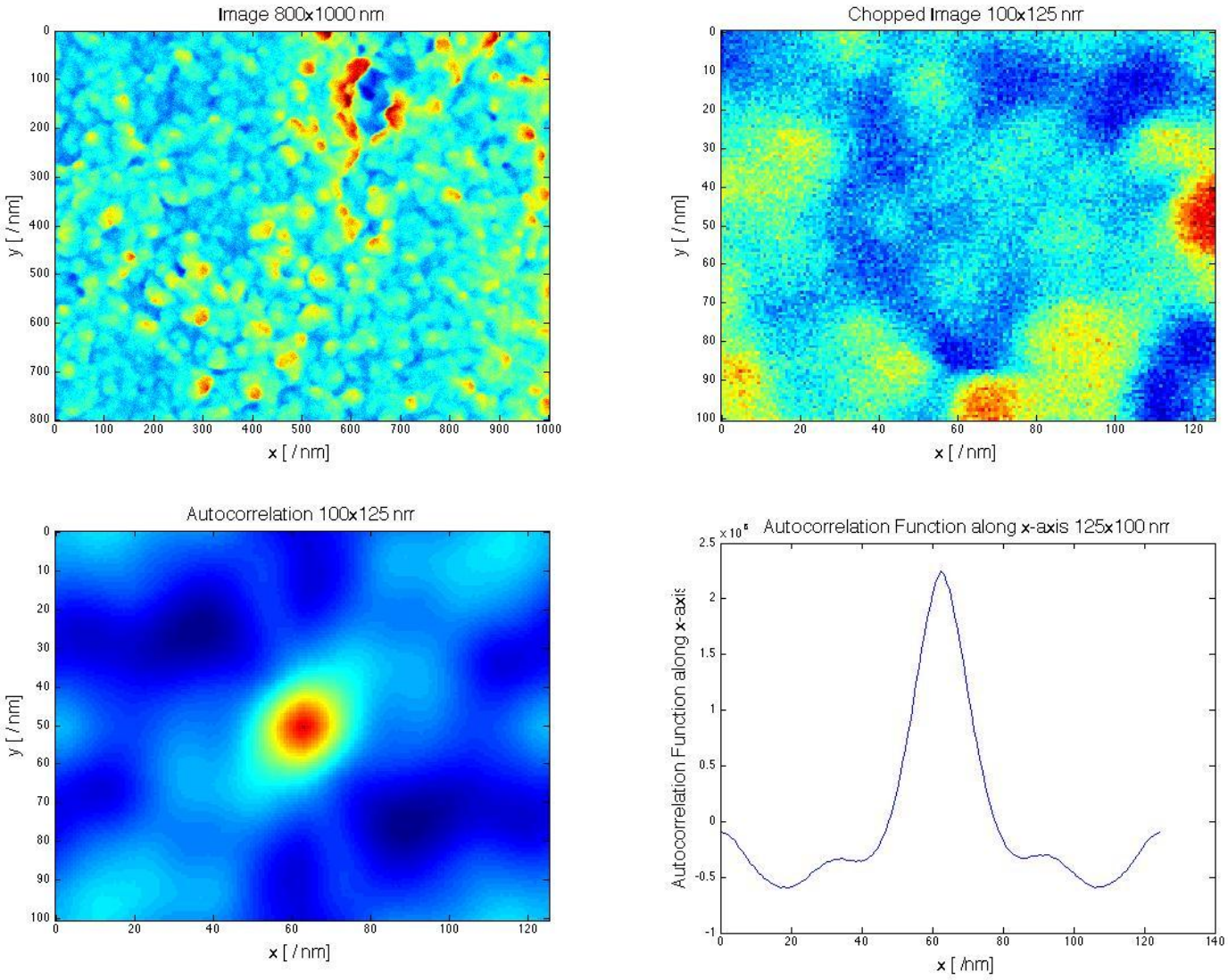


Figure 8. Display of barley image (top left), the second region chosen from the barley image after 8x8 division (top right), direct autocorrelation performed on the chosen region (bottom left), with the lineout along x-axis of autocorrelation (bottom right).

Next the autocorrelation of the same region was calculated, but this time by using the Fourier transform method (using MATLAB's in-built FFT script) corresponding to eq.(6). (Fig.9) As expected, these two methods produce an almost identical autocorrelation function.

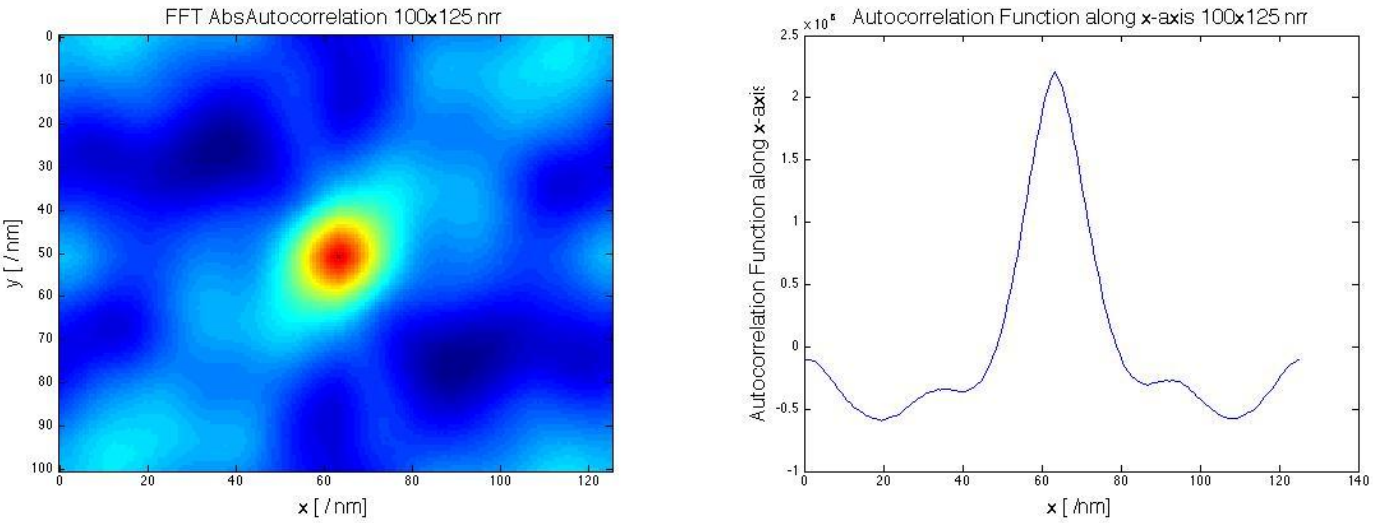


Figure 9. Display of FFT autocorrelation (left) performed on the chosen region of the barley image with the lineout along x-axis of FFT autocorrelation (right). C.f. Fig. 8.

The same procedure was then repeated for the human chromosome image (Fig.5.c). The region of 67x118 nm (Fig.10) was chosen from the relatively homogeneous[3] part of the human chromosome image. Again we see that two methods (Fig.11) appear to produce identical autocorrelations. When we compare the x-lineouts however we notice some slight differences, in particular at the central peak. Given the insignificance of these differences, and that the FFT method is so much fast than the direct correlation method we shall use the FFT method from now on. The slight difference in autocorrelations may be due to the problem of aliasing (see next subsection) in FFT; however more likely is that the similarity in the results indicates that the aliasing problem also affects the direct correlation method. This is also suggested by the rise in the x-lineout near the 3rd peak in both figures 8 and 9 – instead we would have expected this to fall due to the non-crystalline nature of the chromosome. This problem occurs in the case of the FFT method because the in-built function is

[3] There was a significant anomaly in the bottom half of the image that was avoided.

discrete [60]. We shall now describe how we go about minimising this problem.

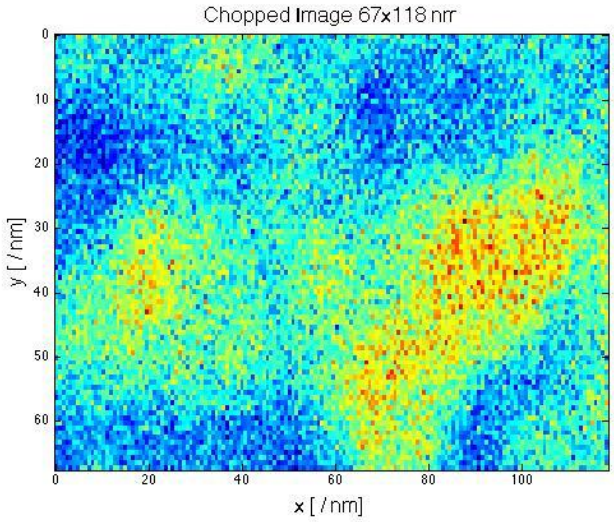
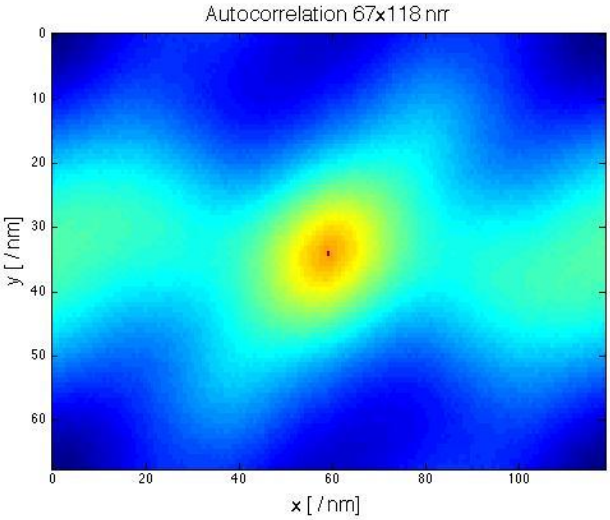
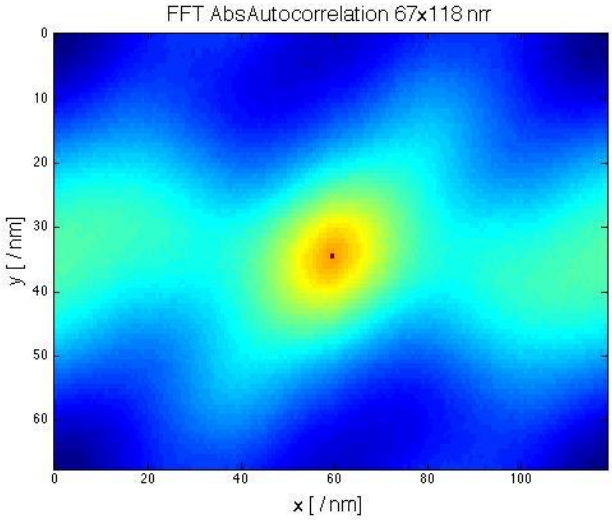


Figure 10. Display of the region of the human chromosome SEM image (Fig.5.c).





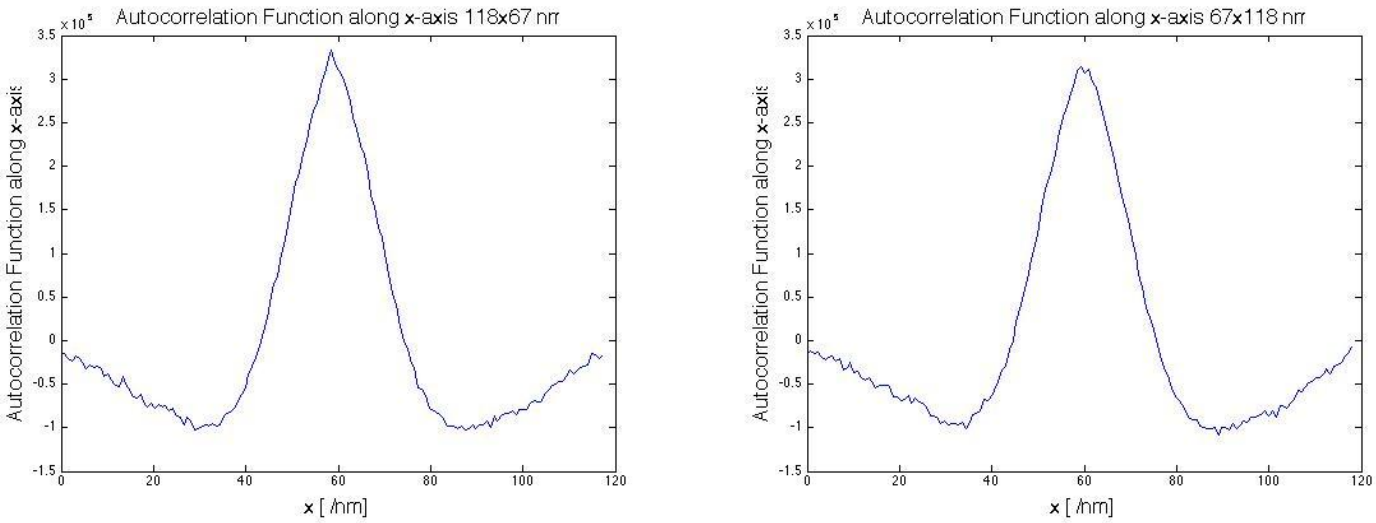


Figure 11. Display of the direct autocorrelation and lineout (left top and bottom) performed on the chosen region of the human chromosome; FFT autocorrelation and lineout (right top and bottom) performed on the same region of the human chromosome.

5.2 Aliasing

The encountered problem of aliasing is essentially the folding back of the image into itself at its boundaries. The usual way of addressing this problem is by increasing the space between the region of interest and the boundary. This can be accomplished either by 'zooming out', or by looking at a smaller sample (that would fit better into the region being analysed). [61] The space can also be increased by padding the image with zeros [60]; this is the method we attempt in Fig. 12.

To investigate this we again take same region of the barley image and pad it with zeros (Fig.12). When the autocorrelation is performed however, the result appears identical to the convolution of two uniform squares (which gives a triangle shape, with a peak whenever the squares fully overlap). We can explain this because of the very large contrast between the zero region and the image region – the function is seeing the image as an essentially uniform square relative to the zeros.

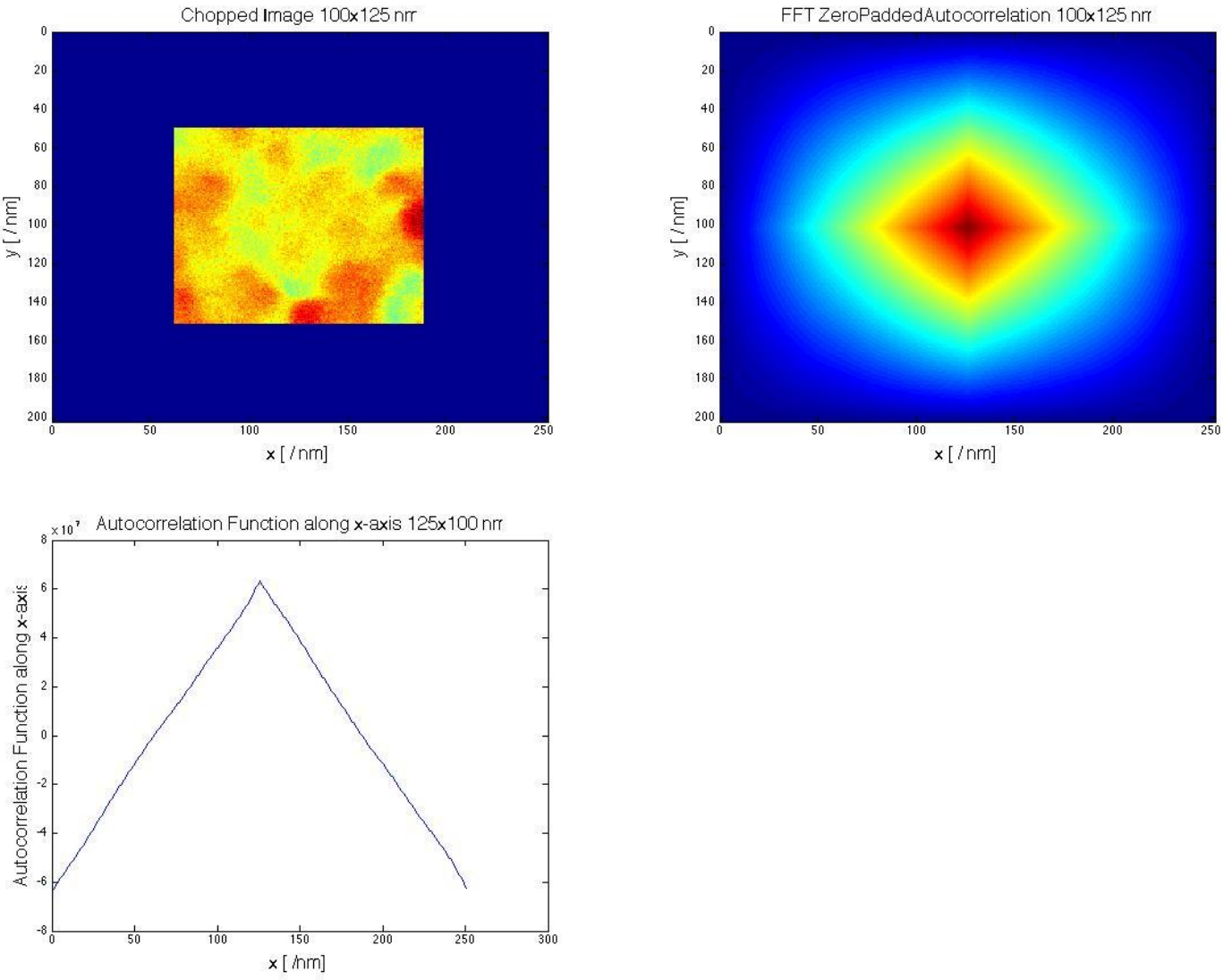


Figure 12. Display of the zero padded second region of barley image (top left) with the FFT autocorrelation (top right) and the lineout along x-axis (bottom left)

Hence, in order to reduce this contrast, we instead pad the image with the mean pixel value. The fast_padded_regions.m and for_fast_padded.m scripts were created for this (see Appendix B).

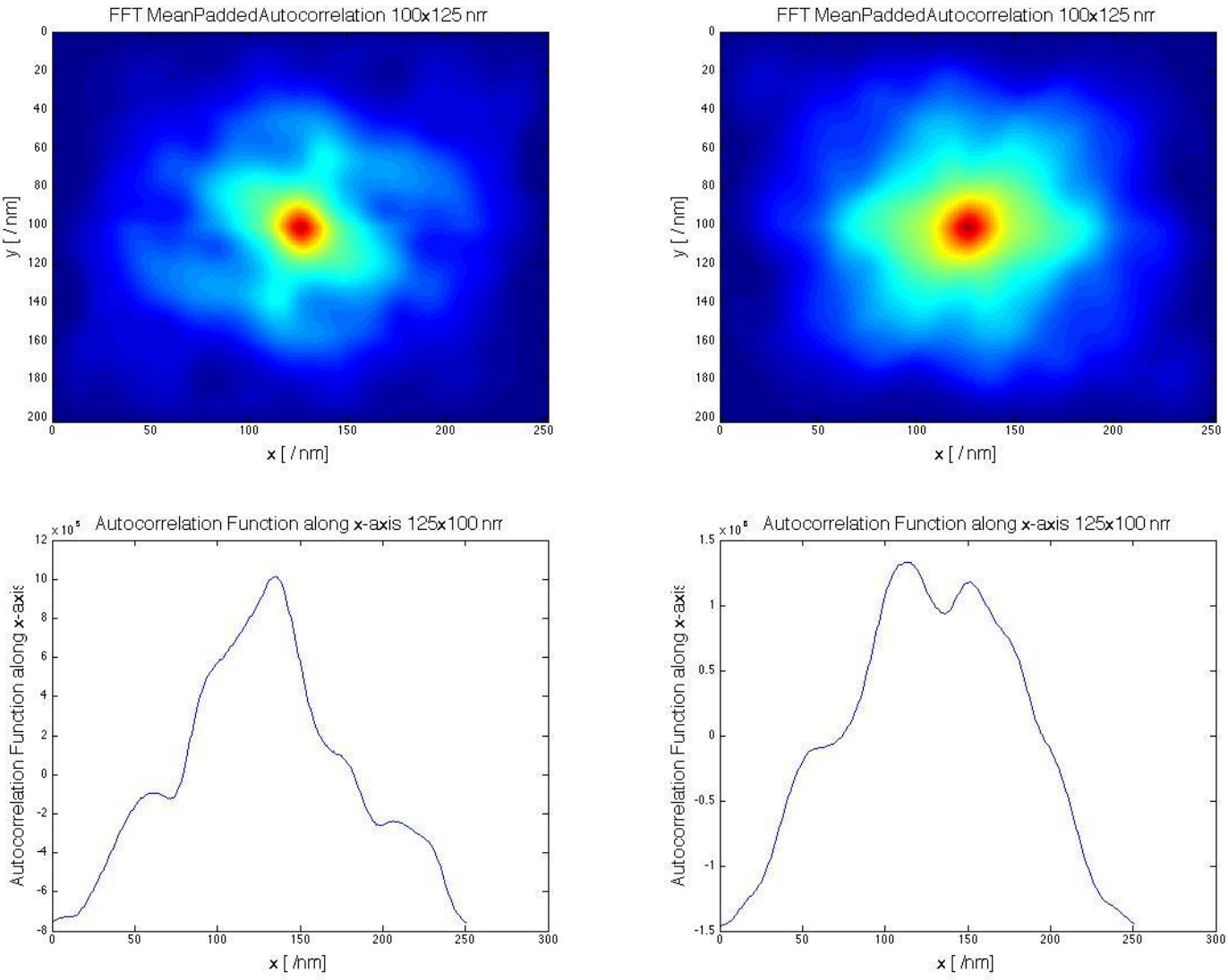


Figure 13. The FFT mean padded autocorrelations of the first (top left), second (top right) regions of the barley image, and its x-axis lineouts are shown.

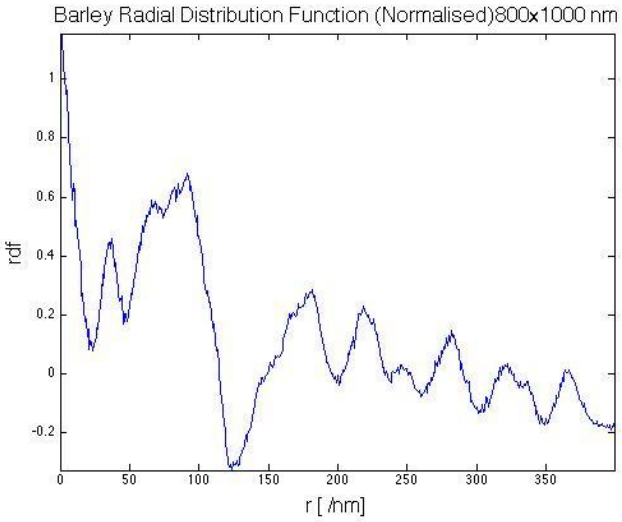
The first two regions of the barley image (again splitting into 8x8 regions) were calculated and displayed (Fig.13). As observed from Fig.13, the irregular shape of the central peak may correspond to different shaped structural units prevailing in each region. However this may also be due to the cutting off of structural units given the small size of the region analysed. Below we analyse much larger regions, producing a more consistent pattern.

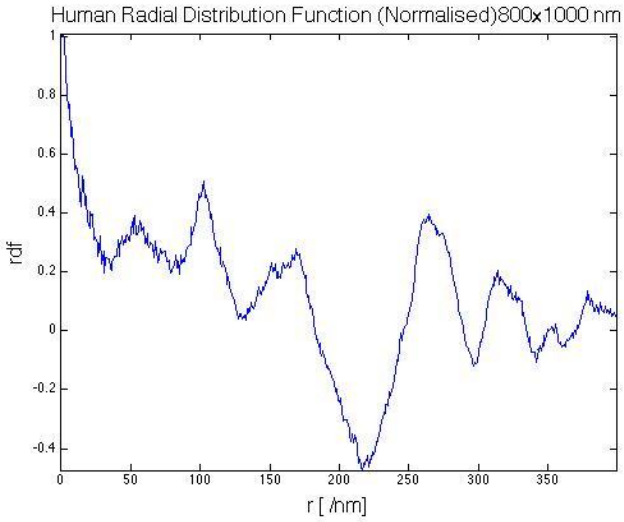
Given the irregular shape of autocorrelation peaks then, it is clear that the x-lineouts are not representative plots of these functions. We therefore proceed to calculate the RDF (which may be thought of as the average of all the unique lineouts of the autocorrelation).

5.3 Radial Distribution Function

The RDF is now employed as defined in eq. (11). The rdf_comparesz_whole.m and for_rdf_comparesz_whole.m scripts are the used (see Appendix B).

The RDF was performed by starting at the centre of the image and increasing the radius up to a maximum corresponding to half of the shortest side (400 nm in this case). The RDF was then normalised arbitrarily so as to always start at unity (i.e. rdf = 1at the centre of the image). The normalisation factor was also chosen so as to have the oscillation baseline at rdf = 0. However, as we shall see, this normalisation may be specific to certain images/regions and needs further investigation. Also, the RDF was averaged over all angles, but only starting at the centre of the image (which is obviously not always a unit of structure). In future the statistical average over all possible starting points ought to be taken.





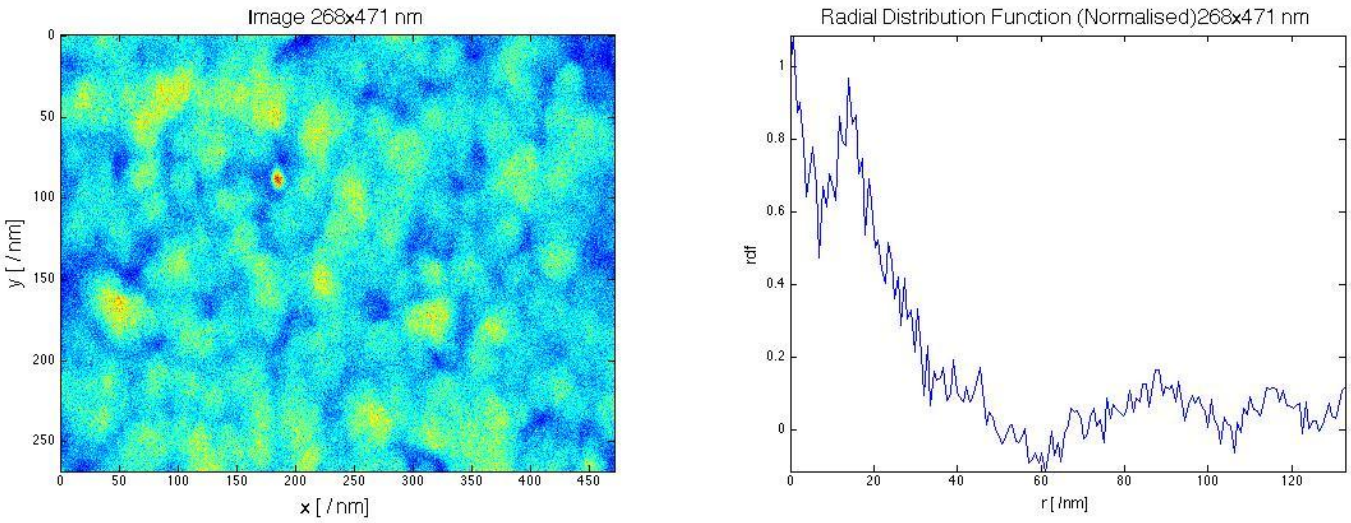


Figure 14. Display of the RDF of the whole barley (top left) and whole human (top right) images, and a smaller region (avoiding the anomalous region) of the human (bottom right) chromosome with its corresponding RDF (bottom right).

The RDF of the whole images of barley (Fig.5.b) and human (Fig.5.c) chromosomes is calculated and plotted in Fig.14. Comparing the two, we observe that there is a deep trough in the RDF of the barely image at about 125 nm; this could be due to the first region of minima following the central region of maxima at between x,y=400nm to x,y=500 nm (Fig.8 Image 800x1000 nm). The same might be the case for the human chromosome at about the 230 nm distance in RDF. The region where the continuous pattern breaks down by the edge of the chromosome arm corresponds to the trough in the RDF. The RDF was also performed on a smaller (268x471nm) region of the human chromosome, cropping out the edge of the chromosome.

By inspection, if we compare human and barley RDFs (Fig14), the barley shows two peaks appearing before the radius reaches 50 nm which could be the size of half of the first structural unit together with the second unit. However the human RDF displays only a single broad peak until 50 nm; this however could be biased due presence of the chromosome edge captured in the image. If the zoomed region of human chromosome is instead observed (Fig.14) it also displays two peaks of structural units before the radius reaches 50 nm. This may indicate a regular 3D structure, with two units of structure stacked on top of each other.

However the negative minima suggest again the need for improving the normalisation factor of the RDF. In future it also may be worth investigating different ring thicknesses (dr) as this was found to affect the result when tested on generated dots.

Finally, we shall test the functions developed on computer generated dots to demonstrate their

successes and failures.

5.4 Dots

In this section the mean padded image with RDF on autocorrelation is performed by compare_whole_image.m script, see Appendix B for this script.

In order to compare the functions and their operation we run the functions on images of organised 'hard spheres' in 2D made in a graphics suite (Xara Designer Pro X). Fig.15 shows the result of running the mean-padded FFT autocorrelation on the dots and the RDF directly on the image.

The autocorrelation appears to be consistent with the distribution and size of the dots. As expected there is a distinct central spot with peaks fading away as the array moves into the mean padding. Hence we could conclude that the autocorrelation performs analysis as expected for this case.

However the RDF shows some errors present, with peaks appearing very noisy, which could be an indication that the normalisation was miscalculated. The arbitrary normalisation factor needs, in other words, to be generalised. It also highlights the need for taking the origin itself to be a unit of structure (which is not the case here). Finally, there were array subscript mismatch errors appearing depending on the dimensions of the image being analysed. This was solved by arbitrarily subtracting or adding 1as required, but this aspect of the code may be improved in future.

Given the requirement of this RDF to start on a unit of structure, we also in Fig. 15., consider performing the RDF on the autocorrelation function itself – which by definition has a peak at the centre. The result observed was well defined, regularly spaced peaks up to about 400nm (the edge of the shortest side) at which point the peaks rapidly fade out. This strongly suggests that in future, performing the RDF on the autocorrelation itself will be more accurate.

We have seen that the RDF can give us information about the spacing of these units of structure; we now investigate if it can give us information about the size of these structures. For that we shall run the function on an image of organised small hard spheres (similar to Fig.15, only smaller). (Fig.16)

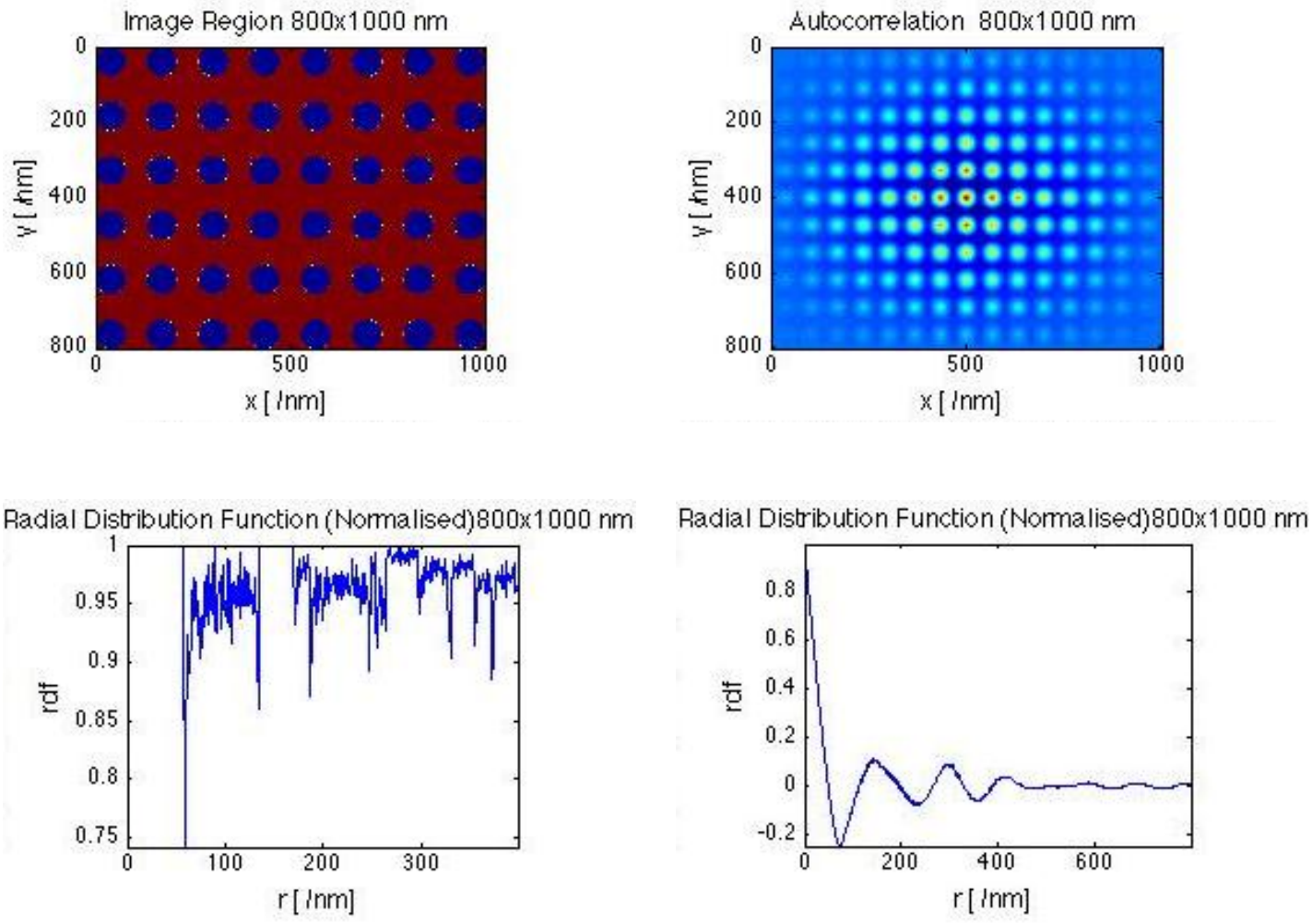


Figure 15. The display of hard spheres (top left) with the autocorrelation (top right) and RDF of the image (bottom left) and the RDF of autocorrelation (bottom right).

As expected, the peaks in Fig.16 appear similar to those in Fig.15 (large spheres), with well-defined spacing. The widths of the peaks did decrease as expected. What we could conclude from that is that both the autocorrelation and the RDF give us information about the size and spacing of the units of structure. However we should emphasize that such observations are qualitative and explicit statistical calculations ought to be performed to show this conclusively.

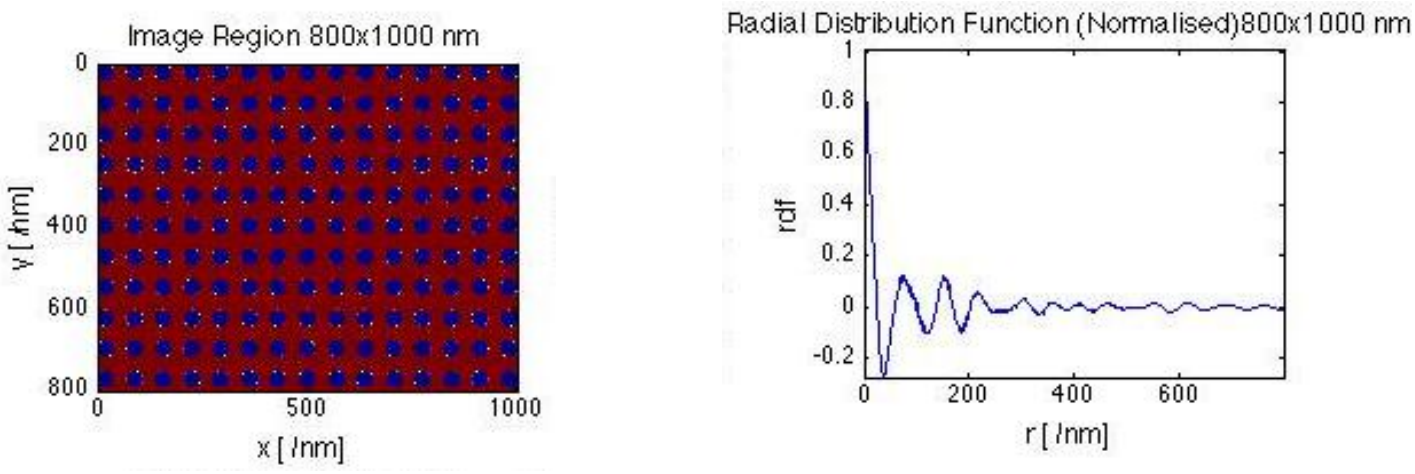
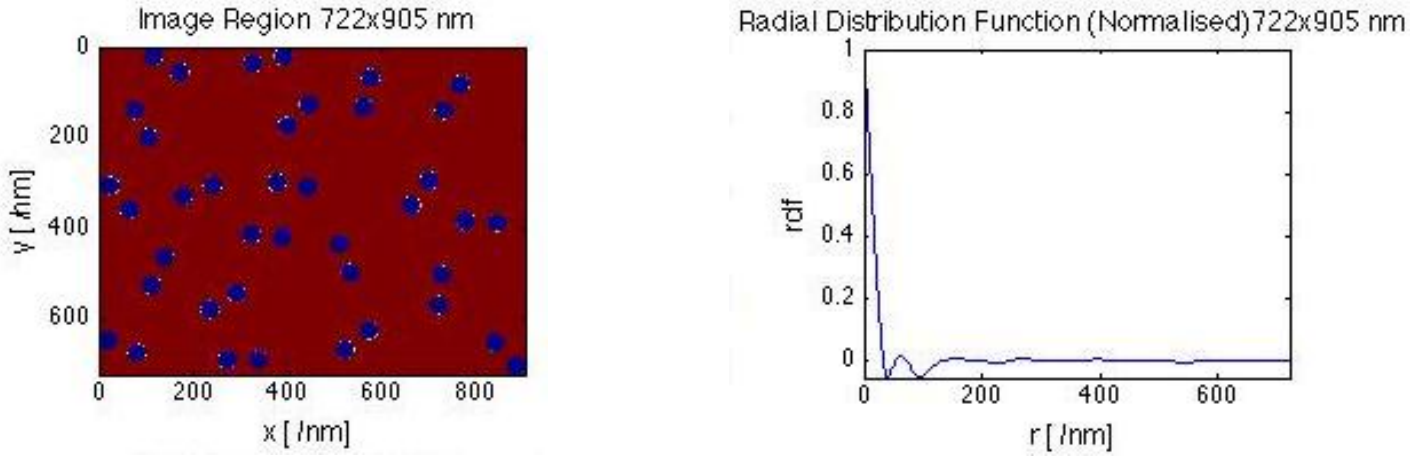


Figure 16. Display of small organised spheres (left) with its RDF on autocorrelation.

We not consider the function's ability to pick out a preferred distance along random orientations by considering pairs of spheres (orientated randomly) and less organised spheres (not in pairs) (Fig.17). Again, we run the mean padded FFT on the images to obtain the autocorrelation and then run the RDF on the autocorrelation. The RDF of the pairs reveals a distinct peak corresponding to the pair separation and then flattens out as expected. In contrast, the RDF of the less correlated spheres shows no prominent peaks at all (again, as one would expect). Hence, by inspection, we could conclude that the autocorrelation and RDF of the autocorrelation give an accurate representation of the repetitive distribution throughout the image.



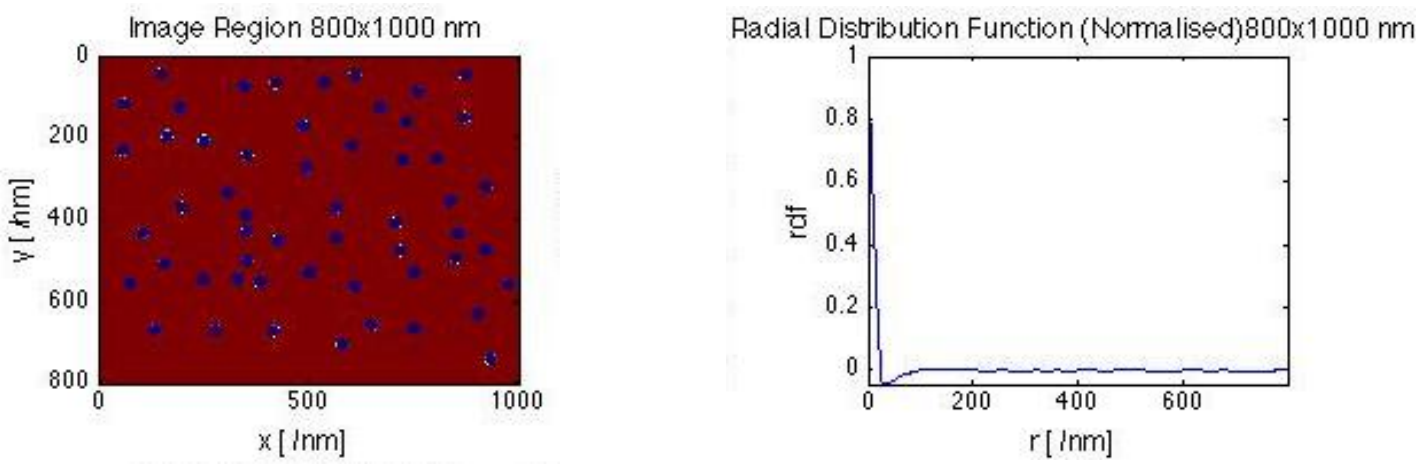


Figure 17. Display of the pairs dots, randomly orientated (top left) with the RDF on autocorrelation of it (top right), and less correlated dots (bottom left) with the RDF on autocorrelation of it (bottom right)

Having thus refined our functions, we now recalculate the RDF of (Fig.14), taking rather the RDF on autocorrelation of the barley and human chromosome images (not the images themselves). This is shown in Fig.18., where the barley auto RDF shows two peaks appearing before the 100 nm radius, which could be showing behaviour of the pair correlated spheres (as in Fig.17). This would mean that there is consistently a structural unit at a repeating distance from each unit of structure, although this peak is relatively minor. For the case of the human RDF, the lack of another distinct peak may be due either to there being a lack of regular structure, or the effect of the anomalous region identified earlier. To further investigate this, the anomalous region was cropped out (i.e. using the same region as in Fig.14), and the analysis was repeated. Now the RDF shows two distinct peaks present in the range before 100 nm indicating a regular structure.

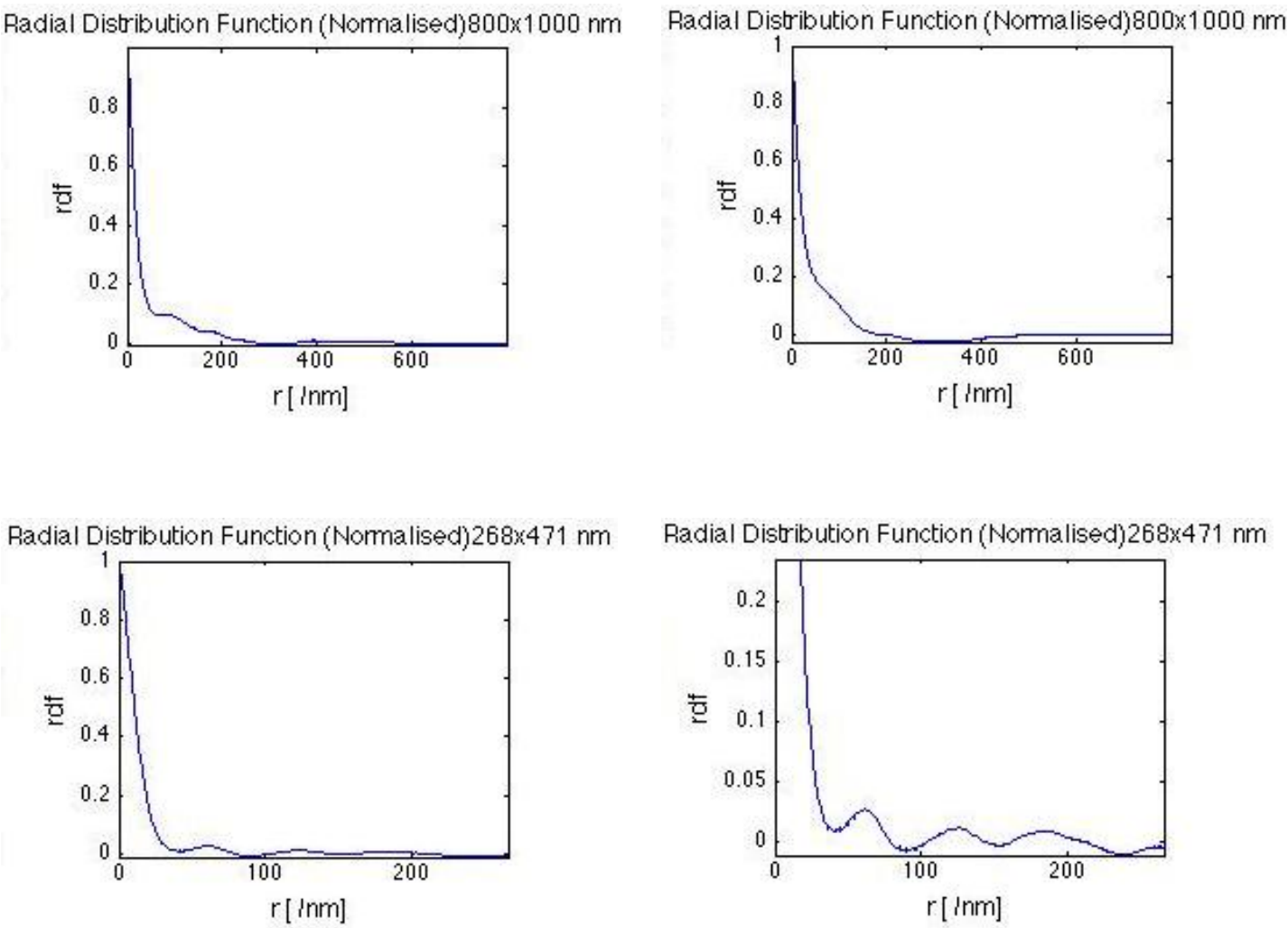


Figure 18. Display of the RDF on autocorrelation of barley (top left) human (top right), human region (bottom left), human region manually zoomed in for clarity(bottom right)

These peaks, by inspection, may be split into a first peak at the origin, with the trough between 30-60 nm, and the next peak appearing between 60-90 nm. This would agree well with the theory that there is a regular structure at the 30 nm scale.

The auto RDF performed on barley and human chromosome images therefore indicate two peaks observed in the radial distance of 100 nm from the centre; these could be taken to be two peaks with a width of 30-40 nm and with a spacing of 30-40 nm.

However, such reasoning is again only qualitative and approximate; in order to support the hypothesis many more regions of chromosomes should be chosen and analysed with the statistical

average drawn and the spacing between peaks and width of the peaks evaluated computationally.

6. Conclusion

In this project we have explored the importance of the chirality in chromosomes and in nature more generally. Following that, the aim was to determine whether the chirality in DNA is passed over to a higher order structure or whether it is lost on the way. For that we considered the importance of the 30 nm structure as the potential missing piece between the chirality of DNA and the possible chirality of the 200 nm structure. We have seen that, although 2D imaging techniques exist, the chirality could not yet be determined as 3D information on the structure of the sample has not yet been obtained. The objective of this project then was to investigate the presence of any regular 30 nm structure; if present, this structure's chirality could then be determined in further investigations, once 3D images are available. In order to analyse this structure, we chose to use SEM to image the chromosome samples and to then analyse the structure with correlation functions in MATLAB.

We looked, in some detail, into why FE SEM in particular was appropriate for imaging chromosomes in this project. In SEM imaging we discussed the essential role of sample preparation and identified the possible errors from the thin coating, nucleoplasm residue and the air drying step which could possibly be sources of aberrations in the images. Two images, of human and of barley chromosomes, were obtained. We then discussed the analysis by correlation functions and the formalism behind it.

The three methods of analysis for this project were outlined: direct autocorrelation, Fourier transform autocorrelation and the radial distribution function. These were then implemented in MATLAB. The direct autocorrelation of small regions of the image showed the peaks of width about 30-40 nm present with separation of 30-40 nm. This observation was checked in the further analysis. The direct autocorrelation and Fourier transform methods were shown to give an almost identical result, but the FFT method proved to be to be much faster. The aberrations in the FFT autocorrelation identified the presence of the aliasing effect. An attempt at fixing this by padding with zeros was made, but failed due to too great a contrast between the pixel values. Padding with mean pixel values instead however seemed to resolve this issue. In the lineouts of the mean padded autocorrelation, we saw some suggestion of the overlap between units of structure (due to the 3D nature of the chromosomes).

The RDF function was used directly on the images obtained, but showed severe errors in

normalisation and statistical averaging when tested on dots. However, it was found that the RDF function worked well when applied to the autocorrelation instead of the image directly (which might have solved one of the errors – the statistical averaging problem identified). The concept was checked on ordered dots, pair correlated and less ordered dots of various sizes. These tests showed that the autocorrelation and the RDF on autocorrelation appear consistent with the size and the degree of order of the dots. Finally, the RDF on the mean padded autocorrelation was performed on the barley and human chromosome SEM images. Both RDF functions displayed the presence of two prominent peaks in the range of 100 nm from the central point (this only became clear in the human image after cropping out anomalous region). It was suggested that this could be due to the presence of two structural units of a width of 30-40 nm, separated by about 30-40 nm.

This suggestion however is only based on qualitative analysis of the RDFs and rough estimates. In future research, a full quantitative analysis should be performed, averaging the size and spacing of all peaks. Also although the RDF appeared to work well on the autocorrelation functions, it still needs to be developed further.

However the main conclusion of the research in this project is that the correlation functions performed in MATLAB show some regular structure present at the 30 nm scale. The exact size and spacing of this structure is a subject for further research. Based on the existence of this structure, the chirality is definitely a research topic worth pursuing when 3D images become available.

Acknowledgements

The author would like to thank her supervisor, Professor Ian Robinson, for illuminating discussions and feedback; and also to thank Dr Jesse Clark for assistance with MATLAB coding and guidance throughout the project, and finally, Laura Shemilt for supervision in preparing the samples and obtaining the human chromosome images as well as constructive feedback throughout the project.

7. Bibliography

1. *Scanning electron microscopy of chromosomes.* **Wanner, G. and Schroeder-Reiter, E.** 2008, Methods in Cell Biology, Vol. 88, pp. 452-454.

2. *Focused ion beam (FIB) combined with high resolution scanning electron microscopy: A promising tool for 3D analysis of chromosome architecture.* **Schroeder-Reiter, E., Pérez-Willard, F., Zeile, U., Wanner, G.** 2009, Journal of Structural Biology, Vol. 165, pp. 97-106.

3. *Louis Pasteur's discovery of molecular chirality and spontaneous resolution in 1848, together with a complete review of his crystallographic and chemical work* . **Flack, H. D.** 65, 23 June 2009, Acta Crystallographica , Vol. A, pp. 371–389 .

4. **Morrison, R.T. and Boyd, R.N.** *Organic Chemistry.* 6th. New Jersey : Prentice-Hall Inc., 1992. pp. 133-140.

5. **Carey, F.A. and Sundber, R.J.** *Advanced Organic Chemistry: Part A: Structure and Mechanisms.* 4th. New-York : Kluwer Academic/Plenum Publishers, 2000. pp. 75-114.

6. *Limonene.* **Thomas, A.F. and Bessiere, Y.** 3, 1989, Nat. Prod. Rep., Vol. 6, pp. 291-309.

7. *Odor incongruity and chirality.* **Friedman L., Miller J.G.** 3987, New-York : s.n., 1971, Science, Vol. 172, pp. 1044-1046.

8. **Barron, L.D.** *Molecular ight scattering and optical activity.* 2nd. New-York : Cambridge University Press, 2004. p. 15.

9. *Thalidomide: The Tragedy of Birth Defects and the Effective treatment of Disease.* **Kim, J. H. and Scialli, A.R.** 1, 2011, Toxicological Sciences, Vol. 122, pp. 1-6.

10. **Strasburger, E. and Fitting, H.** *Strasburger's Text-Book of Botany.* 5th. London : Macmillan, 1921.

11. *The Chromosomes in Heredity.* **Sutton, Walter S.** 5, 1903, The Biol. Bull., Vol. 4, pp. 231-250.

12. *Concerning the Origin of Malignant Tumours by Theodor Boveri.* **Boveri, T.** [ed.] Translated

and annotated by Henry Harris. 1, 2008, Journal of Cell Science, Vol. 121, pp. 1-84.

13. *Genetic implications of the structure of deoxyribonucleic acid.* **Watson, James D., and Crick, F.H.C.** 1953, Nature, Vol. 171, pp. 964-967.

14. *Organisation of subunits in chromatin.* **Carpenter, B.J., et al.** 7, 1976, Nucleic Acids Research, Vol. 3, pp. 1739–1746.

15. *Involvement of histone H1 in the organization of the nucleosome and the salt dependent superstructures of chromatin.* **Thoma, F., Koller, T. and Klug, A.** 1979, Cell Biology, Vol. 83, pp. 402-427.

16. *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* **Davey, C.A., et al.** 2002, J Mol Biol, Vol. 319, pp. 1097-1113.

17. *A new chromosome model.* **Wanner, G., Formanek, H.** 2000, J. of Structural Biology, Vol. 132, pp. 147-161.

18. *3D analysis of chromosome architecture: advantages and limitations with SEM.* **Wanner, G., Schroeder-Reiter, E. and Formanek, H.** 1-3, 2005, Cytogenet Genome Res, Vol. 109, pp. 70-78.

19. *Structural units in chromatin and their orientation on membranes.* **Davies, H.G. and Small, J.V.** 1534, 1968, Nature, Vol. 217, pp. 1122-1125.

20. *Structure of the 3000Å chromatin filament: X-ray diffraction from oriented samples.* **Widom, J. and Klug, A.** 1, 1985, Cell, Vol. 43, pp. 207-213.

21. *Neutron Diffraction of Chromatin in Interphase Nuclei and Metaphase Chromosomes.* **Ibel, Konrad, et al.** 1983, Eur. J. Biochem, Vol. 133, pp. 315-319.

22. *Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure.* **Rhodes, D., Routh, A. and Sandin, S.** 26, 2008, PNAS, Vol. 105, pp. 8872–8877.

23. *EM measurements define the dimensions of the ‘‘30-nm’’ chromatin fiber: Evidence for a compact, interdigitated structure* . **Robinson, P.J.J., et al.** 17, 2006, PNAS, Vol. 103, pp. 6506-6511.

24. *Solenoidal model for superstructure in chromatin.* **Finch, J.T. and Klug, A.** 6, 1976, PNAS,

Vol. 73, pp. 1897-1901.

25. *Structure of the nucleosome core particle at 7 A resolution.* **Richmond, T.J., et al.** 1984, Nature, Vol. 311, pp. 532-537.

26. *Nucleosome arrays reveal the two-start organization of the chromatin fiber* . **Dorigo, B., et al.** 2004, Science, Vol. 306, pp. 1571-1573.

27. *Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions.* **Grigoryev, S. A., et al.** 2009, PNAS, Vol. 106, pp. 13317-13322.

28. *Chromatin structure: does the 30-nm fibre exist in vivo?* **Maeshima, K., Hihara, S., and Eltsov, M.** 2010, Curr. Op. in Cell Biology, Vol. 22, pp. 1-7.

29. *A Method for the In Vitro Reconstitution of a Defined "30 nm" Chromatin Fibre Containing Stoichiometric Amounts of the Linker Histone.* **Van Huynh, A.T., Robinson, J.J and Rhodes, D.** 2005, J. Mol. Biol., Vol. 345, pp. 957–968 .

30. *Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ* . **Eltsov, M., et al.** 50, 2008, PNAS, Vol. 105, pp. 19732–19737 .

31. *Atomic Force Microscope.* **Binnig, G. and Quate, C.F.** 1986, Phys.Rev.Lett., Vol. 56, pp. 930-933.

32. *Structure of human chromosomes studied by atomic force microscopy.* **Tamayo, J.** 3, 2003, J. Struct. Biol., Vol. 141, pp. 198-207.

33. *Atomic force microscopic imaging of 30 nm chromatin fiber from partially relaxed plant chromosomes.* **Sugiyama, S., et al.** 3, 2003, Scanning, Vol. 25, pp. 132-136.

34. *Imaging of Silkworm Meiotic Chromosome by Atomic Force Microscopy* . **Narukawa, J., et al.** s.l. : Wiley Periodicals, Inc. , 2007, Scanning , Vol. 29, pp. 123–127 .

35. **Goldstein, J., et al.** *Scanning Electron Microscopy and X-Ray Microanalysis.* 3rd. New-York : Springer Science+BM, Inc., 2003.

36. *Current SEM techniques for de- and re-construction of centromeres to determine 3D CENH3 distribution in barley mitotic chromosomes* . **Schroeder-Reiter, E., et al.** 2012, J. of Microscopy,

Vol. 246, pp. 96–106 .

37. **Murphy, D.B. and Davidson, M.W.** *Fundamentals of Light Microscopy and Electronic Imaging.* 2nd. Hoboken : John WIley & Sons, Inc., 2013. pp. 199-230.

38. *Visualization and Characterization of High-Order Chromatin Fibers under Light Microscope during Interphase and Mitotic Stages in Plants* . **Liu, J.-W., et al.** 11, 2007, J. of Interactive Plant Biology, Vol. 49, pp. 1634–1639 .

39. **Dehm, G., Howe, J.M. and Zweck, J.** *In-situ Electron Microscopy: Applications in Physics, Chemistry and Materials Science.* 1st. Weinheim : Wiley-VCH Verlag GmbH & Co , 2012. pp. 1-37.

40. **Reimer, L.** *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis.* 2nd. Heidelberg : Springer-Verlag, 1998.

41. **Joy, D.C., Romig, A.D. and Goldstein, J.I.** *Principles of Analytical Electron Microscopy.* New-York : Plenum Press, 1986.

42. *Scanning electron microscopy of chromosomes.* **Wanner, G., Schroeder-Reiter, E.** 2008, Methods in Cell Biology, Vol. 88, pp. 452-454.

43. *Chromosome condensation in mitosis and meiosis of rye (Secale cereale L.).* **Zoller, J.F, Hermann, R. G.,Wanner, G.** 2004, Cytogenetic and Genome Research, Vol. 105, pp. 134-144.

44. *Imaging of DNA in human and plant chromosomes by high-resolution scanning electron microscopy.* **Wanner, G. and Formanek, H.** 6, 1995, Chromosome Research, Vol. 3, pp. 368-374.

45. *Efficient preparation of plant chromosomes for high-resolution scanning electron microscopy.* **Martin, R., et al.** 1994, Chromosome Res, Vol. 2, pp. 411-415.

46. *Refined examination of plant metaphase chromosome structure at different levels made feasible by new isolation methods.* **Schubert, I., Dolezel, J., Houben, A., Scherthan, H., and Wanner, G.** 1993, Chromosoma, Vol. 102, pp. 96-101.

47. *High resolution morphological analysis of in situ human chromosomes* . **Falconi, M., et al.** 2006, Micron, Vol. 37, pp. 146–153 .

48. **Sobolev, V.I.** Convolution of functions. *Hazewinkel, Michiel, Encyclopedia of Mathematics.*

[Online] 2001. [Cited: 14 March 2013.] http://www.encyclopediaofmath.org/index.php?title=C/c026430 . ISBN 978-1556080104 .

49. **Dominguez-Torres, A.** *The Origin and History of Convolution 1: Continuous and Discrete Convolution Operations.* Cranfield Institute of Technology. Bedford : CONACYT, 2012.

50. **Chandler, D.** *Introduction to Modern Statistical Mechanics.* London : Oxford University Press, 1987. pp. 1-288.

51. **Pikulin, V.P. and Pohozaev, S.I.** *Equations in Mathematical Physics: A Practical Course.* [trans.] A. Iacob. Heidelberg : Springer Basel AG, 2001. pp. 170-175.

52. **Stein, E.M. and Weiss, G.** *Introduction to Fourier Analysis on Euclidean Spaces.* New Jersey : Princeton University Press, 1971. pp. 153-172.

53. **Titchmarsh, E.C.** *Introduction to the Theory of Fourier Integrals.* London : Oxford University Press, 1948.

54. **Als-Nielsen, J. and McMorrow, D.** *Elements of Modern X-ray Physics.* 2nd. Chichester : John Wiley & Sons, Ltd, 2011. pp. 113-145.

55. *Density Autocorrelation Function in a Classical Fluid from Initial Correlations.* **Lado, F.** 4, 1970, Phys.Rev. A, Vol. 2, pp. 1467-1473.

56. **Egelstaff, P.A.** *An Introduction to the Liquid State.* 2nd. s.l. : Oxford University Press, 1994. pp. 1-100.

57. **Warren, B.E.** *X-Ray Diffraction.* New-York : Dover Publications, Inc., 1990. pp. 1-15.

58. *Neutron and x-ray diffraction studies of liquids and glasses.* **Fischer, H.E., Barnes, A.C. and Salmon, P.S.** 2006, Rep. Prog. Phys. , Vol. 69, pp. 233–299 .

59. **March, N.H. and Tosi, M.P.** *Introduction to Liquid State Physics.* Singapore : World Scientific Publishing Co. Pte.Ltd., 2002. pp. 75-79.

60. **Sundararajan, D.** *The Discrete Fourier Transform. Theory, Algorithms and Applications.* London : World Scientific Publishing Co. Pte. Ltd., 2001. pp. 225-230.

61. **Brigham, O.E.** *The Fast Fourier Transform and its Applications.* London : Prentice-Hall, Inc.,

1988. pp. 80-106.

62. **Martin, Paul C.** *Measurements and Correlation Functions.* New York : Gordon & Breach, Science Publishers Ltd., 1968. pp. 53-63.

63. *The three dimensional structure of human metaphase chromosomes determined by scanning electron microscopy. In ''Chromosomes and Chromatides''.* **Allen, T. D., Jack, E. M. and Harrison, C. H. J.** 1988, pp. 51-72.

64. **Squires, G.L.** *Introduction to the Theory of Thermal Neutron Scattering.* New-York : Dover Publications, Inc., 1996. pp. 86-88.

65. *The chromosomes in heredity.* **Sutton, Walter S.** 5, 1 April 1903 , The Biological Bulletin, Vol. 4, pp. 231-250.

66. *Involvement of histone H1 in the organization of the nucleosome and the salt dependent superstructures of chromatin.* **Thoma, F., Koller, T. and Klug, A.** 1979, Cell Biology, Vol. 83, pp. 402-427.

67. *Efficient preparation of plant chromosomes for high-resolution scanning electron microscopy.* **Martin, R., Busch, W., Herrmann, R. G., and Wanner, G.** 1994, Chromosome Res, Vol. 2, pp. 411-415.

68. *The three dimensional structure of human metaphase chromosomes determined by scanning electron microscopy. In ''Chromosomes and Chromatides''.* **Allen, T. D., Jack, E. M., and Harrison, C. H. J.** 1988, pp. 51-72.

69. **Morrison, R.T. and Boyd, R.N.** *Organic Chemistry.* 6th. Englewood Cliffs : Prentice-Hall Inc. pp. 133-140.

70. *The Chromosomes in Heredity.* **Sutton, W.S.** 5, 1903, The Biol. Bull., Vol. 4, pp. 231-250.

71. *Chromatin structure: does the 30-nm fibre exist in vivo?* **Maeshima, K., Hihara, S., and Eltsov, M.** 2012, Curr. Op. in Cell Biology, Vol. 22, pp. 1-7.

72. *Three-Dimensional Organization of Drosophila melanogaster Interphase Nuclei. I. Tissue-specific Aspects of Polytene Nuclear Architecture.* **Hochstrasser, Mark and Sedat, John W.** June 1987, J. Cell Biol., pp. 1455-70.

Appendix A.

This is the chromosome preparation according to Wanner's group [44]:

1.Drop-Cryo.

Cold slide were rinsed in distilled water, given a flick. Wiped dry the bottom of the slide (to improve observation of the drop). Positioned on the point of a drop. Taken pipette with 20µl (to 60) of chromosome solution(mixed well beforehand).Let the drop form on the tip of a pipette and let it fall from a height of about 30 cm. (try several times with water for improve of precision).Watched it dry till it evaporated. Then let 2-3 drops of 45% Acetic Acid on the already dried surface. Covered with a cover glass (22*50mm) , pressed firmly. Put facing down onto dry ice for 15 min for cryo purposes [45]

2.Fixation

Next stage dedicated for fixing chromosomes with protein, using glutaraldehyde 2.5% (H=O bonds between dna and proteins). Covered fully both slides with few drops of the solution and shielded with a cover glass for 15-20min. Left in cacodylate buffer for extended period of time (washing time not a matter) to stabilise PH. It should be noted that after cryo (dry ice stage), samples had to be kept wet [46]

3.Staining (&washes)

A sample 0051 is stained with PtBlue 5mM. It was hidden from light in a dark dry space for 30 min. Next stage is washings. For PtBlue sample 0051: Leave for 5 min fully covered in distilled water. Change water, leave again for 10 min. Change water, leave again for 15 min. [17]

4.Drying

Next the sample dehydrated in ethanol: Left for 5 min in 70% solution. After moved to 85% solution for 10 min. Finally left for 15 min in 100% solution.The last stage is to let few drops of 20mkl pipet of HDMS(hexamethyldisilazame) on the sample and leave it to dry. [42]

5. Aftermath:

After the sample was ready the observation of chromosomes in the samples was made on Optical microscope ZEISS. Six distinct spots were identified and marked. The identified spots of chromosomes were next observed on SEM.

Appendix B.

For Direct Autocorrelation: for_correct.m and correct_auto.m scripts

for_correct.m:

```
[FileName,PathName]=uigetfile('.tif','Please select the .tif file you wish to analyse');
% opens user interface to select the tif image needed

%Use the uigetfile for user interface, otherwise paste the full adress of
%the image chosen into fullpath

%fullpath='/Users/kamilasuankulova/Desktop/BSc_Project/MATLAB/FW%3a_Data_for_Correlation_Functions_/0026_Barley_PtBlDapi_06_super_GB-H_SE.tif';

fullpath=fullfile(PathName,FileName);% define a full path

temp = imfinfo(fullpath, 'tif'); %returns info of the tiff file

image = imread(temp.Filename,1); %reads colour image from the temp file

image = double(image); %doubles the precision of the image
pixel_scale=100/128; %this means there are 128pixels for 100nm

nn=8; %this is the number of regions to split the image up into

sz=size(image);% returns the size of the image
xmasksz=round(sz(2)/nn);%define x size of the mask
ymasksz=round(sz(1)/nn);%define y size of the mask

big_out=cell(nn,nn); % creates an empty cell array to store each mask output

figure % call for figure

x=((0:sz(2))*pixel_scale); % make a 'scale' each pixel is 100/128 nm
y=((0:sz(1))*pixel_scale);

imagesc(x,y,image % display the picture with the scale

xlabel('x [ / nm]','FontSize',16) % label x-axis
ylabel('y [ / nm]','FontSize',16) % label y-axis
title(['Image',num2str(round(sz(1)*pixel_scale)),'x',num2str(round(sz(2)*pixel_s
```

```
cale),' nm'],'FontSize',16) % title the image with pixels dimensions defined

 %Now we continue to for loop to analyse each small region

for yyy=1:ymasksz:sz(1)-(nn-1)*(ymasksz)+1;% define function of y parameter

    for xxx=1:xmasksz:sz(2)-(nn-1)*(xmasksz)+1;% define function of x parameter

        %Note!!!!
        %the image is split into nnxnn masked regions. Here its displayed just
        %the first 4 regions (in the top left corner) for demonstration. To
analyse the
    %full image, piece by piece, remove the "-(nn-1)*(ymasksz)" and "-(nn-1)*
                    (xmasksz)" above, but its time consuming.

        mask_points=[xxx, xxx+xmasksz, yyy, yyy+ymasksz];
        %define the mask    dimensions

        [auto smx smy]=correct_auto(image,mask_points,pixel_scale);
         % call for the autocorrelation function of a chosen region

        figure % call for figure
        x=((0:smx-1)*pixel_scale); % make a 'scale' each pixel is 100/128 nm

        plot(x,(auto(round(smy/2),:)-mean(auto(round(smy/2),:))))
        %plot the autocorrelation lineout in x direction
         xlabel('x [ /nm]','FontSize',16)%labels the x-axis
         ylabel('Autocorrelation Function along x-axis','FontSize',16)%labels
the y-axis
         title(['Autocorrelation Function along x-axis',
         num2str(round((smx)*pixel_scale)),'x',
num2str(round((smy)*pixel_scale)),   `        'nm'],'FontSize',16);
          %title the correlation lineout with pixels  defined

         big_out((xxx+xmasksz-1)/xmasksz,(yyy+ymasksz-1)/ymasksz)={auto};
         % populates the cells with each mask output of correlation function

    end
end
```

correct_auto.m:

```
function [auto_extract,smx,smy] = correct_auto (image,mask_points,pixel_scale)%
define a function for autocorrelation

% Now define a mask for easy computation

smy=(mask_points(4)-mask_points(3));  % get y-dimensions of a chosen region
smx=(mask_points(2)-mask_points(1)); % get x-dimensions of a chosen region
x=((0:smx)*pixel_scale); % make a 'scale' each pixel is 100/128 nm
y=((0:smy)*pixel_scale);
% make a zero matrix of a size of a chosen region, magnify the image region by
x2

chopped_image=image(mask_points(3):mask_points(4),mask_points(1):mask_points(2))
; % replace zero matrix with the numbers from the image region chosen,easier
to define it at the +1 corner rather than defining a middle region

figure % call for figure
imagesc(x,y,chopped_image) % display chopped image
```

```
xlabel('x [ / nm]','FontSize',16) % label x-axis
ylabel('y [ / nm]','FontSize',16) % label y-axis
title(['Chopped Image',
num2str(round(smy*pixel_scale)),'x',num2str(round(smx*pixel_scale)),'
nm'],'FontSize',16) % title the chopped image with pixels dimensions defined

% Now perform the autocorrelation function

auto=2*zeros([smy,smx]);%empty array to store values doubled magnification


for yy=1:(2*smy);% define function of y parameter

    for xx=1:(2*smx);% define function of x parameter

          ys=-round(smy)+yy;  %get relative y shift

          xs=-round(smx)+xx;  %get relative x shift
          auto(yy,xx)=sum(sum(circshift(chopped_image,[ys,xs]).*chopped_image));
          % double sum for y and x over the function* shifted function-
          perform  autocorrelation
    end
end

figure%call for figure
imagesc(2.*x,2.*y,auto)%display auto
xlabel('x [ / nm]','FontSize',16)%label x-axis
ylabel('y [ / nm]','FontSize',16)%label y-axis
title(['Repeating autocorrelation
',num2str(round(2*smy*pixel_scale)),'x',num2str(round(2*smx*pixel_scale)),'
nm'],'FontSize',16)%title the image

%Generate extracted correlation region at the centre.
auto_extract=auto(round(smy/2):round(3*smy/2)-1,round(smx/2):round(3*smx/2)-1);
%define borders of the extract
figure% call for pic
imagesc(x,y,auto_extract)% display the extracted image
xlabel('x [ / nm]','FontSize',16)% label x-axis
ylabel('y [ / nm]','FontSize',16)
99
                % label y-axis
 title(['Autocorrelation',num2str(round(smy*pixel_scale)),'x',
num2str(round(smx*pixel_scale)),' nm'],'FontSize',16) % title the image with the
pixels sizes defined

end
```

```
For FFT with mean padding: fast_padded.m (the 'for' function run identical to
for_ correct.m, just call for fast_padded.m instead)
For Fast Fourier transform fast_padded.m is available
fast_padded.m:
function [autoc, smx, smy] = fast_padded (image,mask_points)
%This function use Fourier Transforms to determine autocorrelation function
%on images
%Define mask dimensions

smy=(mask_points(4)-mask_points(3));  %get y-dimensions of a chosen region
smx=(mask_points(2)-mask_points(1)); %get x-dimensions of a chosen region
```

```
ycentre=round(smy/2%define the position of centre for y
xcentre=round(smx/2);%and for x
chopped_image=mean(image(:))*ones(2*[smy,smx]); %make a mean matrix of a size of
a chosen region,
%magnify the image region by x2

%Due to subscript of matrices mismatching due to size of the image chosen 2
types
%of mask points are defined:

%chopped_image((smy-ycentre):(ycentre+smy-1),(smx-xcentre):(xcentre+smx-
1))=image(1:smy,1:smx);
%replace zero
%matrix with the numbers from the image region chosen,easier to define it at the
+1 corner rather
%than defining a middle region

%or

chopped_image((smy-ycentre):(ycentre+smy-2),(smx-xcentre):(xcentre+smx-
2))=image(1:smy,1:smx);
%replace zero matrix with
%the numbers from the image region chosen,easier to define it at the +1 corner
rather than
%defining a middle region

figure%call for figure
imagesc(chopped_image)%display of figure

%The convolution theorem is calculated with Fourier transform of the
%function/image
F=ifftshift(fftn(fftshift(chopped_image))); %fft of the image, with appropriate
shift

P=abs(F).^2;% modulus squared of the convolution/correlation, equivalent
%to power spectrum or intensity


autoc=ifftshift(ifftn(fftshift(P))); %calculate the autocorrelation by applying
inverse fft
%on power spectrum

end
```

For RDF: rdf_comparesz_whole.m is available

rdf_comparesz_whole.m

```
function me2 = rdf_comparesz_whole(image,mask_points)%the radial distribution
function is performed

%The function rdf_comparesz finds and plots the radial distribution function of
%the input image.
smy=(mask_points(4)-mask_points(3));%get y-dimensions of a chosen region
smx=(mask_points(2)-mask_points(1));%get x-dimensions of a chosen region

image=image(1:smy+1,1:smx+1);%define the image region

dr=1;%choose the thickness of the ring
me=zeros(smy,1);%define new array of zeros to store values of the mean

%Now start the loop for ring averaged correlation points
```

```
for         rad=1:round(smx/2)% set radius from 1 to the size of the image
            ring=generate_ring(smx+1,rad,dr);%call for function swiping a ring
of pixels of size
            ring=ring(round((smx-smy)/2+1):round(smy+(smx-smy)/2)+1,:);%cut the
square matrix to fit a rectangulular correlation matrix
            me(rad)=mean(nonzeros(image.*ring));%fill me matrix with the
averaged nonzero values of the matrix multiplication
end

me=me(1:round(smy/2));

%Normalise the 'me' function dividing by the mean value across the whole
correlated region and multiplying by the first value

me2=1/(me(1)/mean(image(:))-1)*(me/mean(image(:))-1);


end

% There are several errors present in the function:
% 1. The function is specialised for the rectangular image.
% 2. The ring function makes a ring of a bigger surface=square,
% which is then cut to fit the rectangle of the image.
% 3. Is the mean same as the number density N, as it was witnessed
% before that the sum of pixels doesnt give us any difference
% if measured on the sample image or on the zero pixels image.
% 4.The generate_ring uses meshgrid, whatever else it could be.
% 5.The normalisation of me2 is done by trial and error, by dividing
% by mean value of the correlation and multiplying by first value.
```

The script to run both mean padded autocorrelation, x-lineout and the RDF on autocorrelation : compare_whole_image.m

compare_whole_image.m:

```
%This script will conduct the fft autocorrelation of the mean padded image
together
%with its x-lineout and the RDF on autocorrelation

pixel_scale=100/128; %this means there are 128 pixels for 100nm


[FileName,PathName]=uigetfile('.tif','Please select the .tif file you wish to
analyse');
%opens user interface to select the tif image needed

%fullpath='/Users/kamilasuankulova/Desktop/BSc_Project/MATLAB/FW%3a_Data_for_Cor
relation_Functions_/0026_Barley_PtBlDapi_06_super_GB-H_SE.tif';


fullpath=fullfile(PathName,FileName);% define a full path

temp = imfinfo(fullpath, 'tif');  %returns info of the tiff file

image = imread(temp.Filename,1);  %reads colour image from the temp file
image = double(image); %doubles the precision of the image

sz=size(image);% returns the size of the image
```

```
mask_points=[1, sz(2), 1, sz(1)];% define mask points

%for fft padded autocorrelation
[auto smx smy]=fast_padded(image,mask_points);% call for the autocorrelation
function of a chosen region


% use a subplot to plot each region analysed on the same figure
% for easy comparison.

x=(0:sz(2))*pixel_scale;    %make a 'scale' each pixel is 100/128 nm
y=(0:sz(1))*pixel_scale;

%We first plot the chopped image that will be analysed:
fig1=figure(1); %name figure
subplot(2,2,1);
imagesc(x,y,image(mask_points(3):mask_points(4),mask_points(1):mask_points(2)))%
define the display region
                xlabel('x [ /nm]','FontSize',12)% label the x and y axis
                ylabel('y [ /nm]','FontSize',12)
                title(['Image Region
',num2str(round(pixel_scale*sz(1))),'x',num2str(round(pixel_scale*sz(2))),'
nm'],'FontSize',12)
                %title of the figure


x=(0:2*sz(2))*pixel_scale;%make a 'scale' each pixel is 100/128 nm
y=(0:2*sz(1))*pixel_scale;

%We now plot the each correlation function of each region:
fig2=figure(1);%name figure
subplot(2,2,2); imagesc(x,y,auto)% define the display region
                xlabel('x [ /nm]','FontSize',12)% label the x and y axis
                ylabel('y [ /nm]','FontSize',12)
                title(['Autocorrelation
',num2str(round(pixel_scale*sz(1))),'x',num2str(round(pixel_scale*sz(2))),'
nm'],'FontSize',12)
                %title of the figure

%We now plot the a lineout of each correlation function (along the x axis),
%remembering to subtract out the mean:
szauto=size(auto);%define size of auto
x=(0:szauto(2)-1)*pixel_scale;%scale the axis
fig3=figure(1);%name the figure
subplot(2,2,3); plot(x,(auto(round(smy/2),:)-mean(auto(round(smy/2),:))))
                xlabel('x [/nm]','FontSize',12)% label the x and y axis
                ylabel('g(r)','FontSize',12)
                title([' X line-out
',num2str(round(pixel_scale*sz(1))),'x',num2str(round(pixel_scale*sz(2))),'
nm'],'FontSize',12)
                axis tight% display option tight graph
                %title of the figure

%We now plot the radial distribution function of each correlation function:
    % (NB we have to run the function first to get the size, to then be
    % able to scale the axis (from px to nm ))
mask_points=[1, 2*sz(2)-2, 1, 2*sz(1)-2];%define mask points
rdf=rdf_comparesz_whole(auto,mask_points);%call for RDF
r=(0:size(rdf)-1)*pixel_scale;%scale the axis (from px to nm )
```

```
fig4=figure(1);% display figure
subplot(2,2,4); plot(r,rdf) %for some reason this seems to plot just half of the
expected radius of each region
                xlabel('r [ /nm]','FontSize',12)% label the x and y   axis
                ylabel('rdf','FontSize',12)
                title(['Radial Distribution Function
(Normalised)',num2str(round(pixel_scale*sz(1))),'x',num2str(round(pixel_scale*sz
(2))),' nm'],'FontSize',12)
                axis tight% display option tight graph
                %title of the figure
```

For generated dots: generate_dots.m

generate_dots.m:

```
%Here is the varranddots code %

nx=160;ny=128;% define size of the image

randoff=10;%some random offset 0=>perfect correlation

dotspacing=10;%spacing between dots

arrayoff0=100*ones([ny+randoff,nx+randoff]);% create a matrix

x=(1:nx);
y=(1:ny);%dimensions of x and y

yoff=round(random('uniform',0,randoff,[ny 1]));% the offsets defined for x and y

xoff=round(random('uniform',0,randoff,[nx 1]));

for qq=1:dotspacing:ny,arrayoff0(y(qq)+yoff(qq),x(qq)+xoff(qq))=110;% the
generated dots for loop
end
figure%call for figure
imagesc(arrayoff0)% display of dots
```